European
Historical
Economics
Society

EHES WORKING PAPERS IN ECONOMIC HISTORY | NO. 114

Missed opportunities? The development of human welfare in Western Europe, 1913-1950

Daniel Gallardo Albarrán

University of Groningen

JUNE 2017

EHES Working Paper | No. 114 |June 2017

Missed opportunities? The development of human welfare in Western Europe, 1913-1950*

Daniel Gallardo Albarrán
University of Groningen

Abstract

Income per capita suggests that the period 1913-1950 is one of missed opportunities for improving living standards in Europe. However, life in Europe during these years improved significantly, as citizens began experiencing dramatic declines in mortality, working time and inequality thanks to (among others) the spread of modern medicine and the introduction of the 8-hour working day. To measure the contribution of these aspects to broader welfare, I apply a new utility-based framework that, contrary to previous composite indices such as the Human Development Index, allows for a welfare analysis directly comparable to GDP across countries and time. The results using the new measure shows that income per capita underestimates welfare growth significantly (up to five percent annually). Moreover, with this indicator cross-country differences in living standards are much larger and more persistent than other composite indices of well-being imply. These findings call for a reappraisal of the evolution of living standards during the period 1913-1950 and, more generally, of the measurement of multi-dimensional welfare in historical contexts.

JEL classification: I100, I3, N30, D63, O53

Keywords: Living Standards; welfare; well-being; twentieth-century; life expectancy; leisure.

* Earlier versions of this paper were presented at seminars at the University of Münster, University of Groningen, University of Wageningen, London School of Economics, and at the Bank of Spain, World Economic History Congress in Kyoto and European Historical Economics Society Conference in Pisa. I thank participants at these meetings as well as Herman de Jong, Robert Inklaar, Marcel Timmer, Martin Uebele, Charles Jones, Ewout Frankema, Jan Luiten van Zanden, Leandro Prados de la Escosura, Daniel Tirado and Alejandra Irigoin.

Notice

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

1 Introduction

The evolution of living standards during the first half of the 20th century in Europe is characterised by sharp contrasts. On one side, this period is often portrayed as one of missed opportunities in that income per capita exhibited historically-low growth rates due to the effect and persistence of (among others) armed conflicts, misguided macroeconomic policies and increasing protectionism (Roses & Wolf, 2010). As a result, material well-being the old continent fell behind other developed economies less affected by these events such as the United States, Canada or Australia. On the other hand, Europeans witnessed unprecedented improvements in other aspects of well-being such as health, leisure or inequality, following the application of germ theory of disease to private and public life, the development of antibiotics or the introduction of the 8-hour day. The rate of progress in these aspects was so remarkable in Europe that in the decades between the dawn of the First World War and the mid-20th century life expectancy at birth increased by 18 years and annual working time declined by more than 600 hours (Riley, 2005; Huberman & Minns, 2007). To jointly consider the different perspectives conveyed by economic and social indicators during this period, an important part of the literature analysing historical human welfare has taken a multi-dimensional perspective on well-being and has employed a composite indicator that aggregates information on income, health and education: the Human Development Index (UNDP, 1990). The application of this index to Europe during the first half of the 20th century indicates that, contrary to income per capita, this period should not be characterized as one of missed opportunities in terms of human development for two reasons. First, European welfare growth was neither low in historical perspective nor significantly slower than in North America. Second, differences across countries narrowed greatly as a result of strong growth in countries with low levels of human development (Crafts, 2002; Millward & Baten, 2010).[1] In other words, HDI-based evidence suggests that opportunities to improve well-being beyond income during the period 1913-1950 in Europe were taken over time and contributed to a more egalitarian distribution of welfare across countries by mid-century.

However, some of the elements underpinning this rather optimistic view of the first half of the 20th century have been recently called into question because of the critiques of the HDI as a suitable measure for understanding human welfare in the past. Prados de la Escosura (2015) argues that the linear transformation that this index applies to its social dimensions (i.e. health and education) introduces a spurious tendency towards convergence and makes comparisons across countries and time difficult. Since these indicators have asymptotic limits, an absolute change in, for example, life expectancy is larger the lower its initial level, thus

[1] See Figure 1 in the Appendix for a graphical representation.

favouring countries with lower levels of HDI. To overcome this problem, he developed the Historical Index of Human Development (HIHD), which gives higher weight to improvements at high levels of the social indicators than improvements of similar magnitude at low levels. With this new indicator, cross-country variation in human development is larger than implied by the HDI. Another aspect of the HDI (and similar indices) that has been criticised is the lack of theoretical basis of the aggregation procedure because it arbitrarily assigns equal weight to its three sub-components (Nordhaus, 2003, 20).[2] Even if this arbitrary weighting scheme reflected individuals preferences in the present, this issue poses a serious obstacle for long-term comparisons because the relative importance of the HDI components is assumed to be constant over time. This disregards the idea that individuals might value different aspects of their lives differently throughout time, as historical evidence for the case of health suggests (Williamson, 1982; Costa & Kahn, 2004). A more fundamental measurement issue for economic historians concerns the interpretation of the HDI in the past and how it compares to income per capita. Costa and Steckel (1997, 73-74) point out that whereas the HDI is a distance measure that shows when modern living standards were achieved, income growth is a measure of velocity that stresses the improvements witnessed by contemporaries. In periods in which income starts from a low level, a modest rate of economic growth (despite being important for contemporaries) has a small relative importance in the index because progress in this aspect represents a very small fraction of modern living standards. On the basis of this argument, the HDI is a retrospective index that may be at odds with contemporaries' views in that the relative importance of its dimensions in history depend on the level of attained living standards in the present. Therefore, although the views provided by this index and income per capita are useful for analysing historical living standards, a comparison between the two should be made with caution. To overcome the weighting issue and provide a closer perspective of contemporaries' well-being, a different strand of the literature has used measurement frameworks grounded in economic theory that are directly comparable to income measures.[3] By combining income, health and leisure, Crafts (1997) reports welfare growth rates for Western Europe during the analysed period that are only one percentage point larger than income and that, contrary to the HDI, imply no welfare convergence during the period 1913-1950. This scattered evidence from different indicators of well-being brings

[2]Ravallion (2012) also argues that the multiplicative formula adopted for the HDI as of 2010 to relax the past assumption of perfect substitutability between its underlying components has serious shortcomings. With the new formula, the weight of longevity in poor countries has been significantly reduced and the value of extra years of schooling is (for most countries) much larger than suggested by their returns in the labour market.

[3]Rijpma (2014) analyses global well-being since 1820 taking a data-driven instead of a theoretically-grounded approach. By using principal component analysis, the weight of each dimension is chosen based on their shared information.

us back to the pessimistic picture portrayed by income per capita because it does not support the high-performing and egalitarian view of European welfare suggested by the HDI. Taken together, this evidence brings into question whether the first half of the 20th century can be characterised as a period of taken or missed opportunities for European living standards once we extent their analysis beyond income per capita.

To answer this question, I partially depart from the dimensions analysed by the HDI and propose a more encompassing framework that better suits the historical context in which Europeans lived by considering four crucial aspects of well-being: material well-being, health, leisure time and inequality. I start looking at income per capita to provide a detailed quantification of the extent to which economic opportunities were missed across regions and time. Note that I will not use this indicator to measure material well-being as is common in the literature, but rather for discussing and motivating the need to expand the study of living standards beyond national income because it partially neglects non-income aspects of well-being. As a more precise (and less used) measure of individuals' economic well-being, I will use household consumption per capita. The advantage of this indicator is that it is more closely related to the level of material living standards of the population because it provides information on the share of national income that is actually consumed by households. Furthermore, household consumption is not affected by some unusual dynamics between government spending and gross domestic product (GDP) during this period. For example, rising political tension between the two wars in some countries such as the United Kingdom resulted in increased military spending and therefore increased GDP per capita which did not, at least directly, benefit the population (Barro & Ursua, 2008). The second aspect of well-being not taken into account in historical calculations of the HDI that will be considered is economic inequality. Especially after 1929, accounting for this element is important since it experiences a long-term decrease until the 1980s (van Zanden, Baten, Foldvari, & van Leeuwen, 2013). Up to 1950, the inequality decline is partially the result of the compression of the earnings distribution in the 1930s and 1940s (Atkinson, 2007). The third aspect of well-being that will be discussed is health. In just a few decades, Europeans' health experienced a revolution due to the application of the germ theory of disease to develop more efficient public health infrastructures and more hygienic practices by individuals, the discovery of antibiotics and improvements in nutrition (Cutler, Deaton, & Lleras-Muney, 2006). As a novelty in this type of historical cross-country studies, I will not use life expectancy at birth but age-specific mortality rates. This is particularly relevant during the analysed period because most of the mortality decrease took place in the youngest part of the population and with these data I can measure the welfare gains from improving health in different parts of the age distribution. The last welfare dimension that will be considered is also not taken into account by

other indicators of well-being despite it having dramatically changed Europeans' lives during the period 1913-1950: working time. Due to the introduction of the 48-hour week and the increase in vacation and national holidays (Evans, 1969; Huberman & Minns, 2007), the amount of time workers could devote to leisure increased substantially in Europe. Moreover, including this aspect in the analysis is not only relevant from a long-term perspective, as the fall in annual hours worked during this time span is unprecedented, but also because other studies have highlighted that welfare gains derived from reduction in hours worked are comparable to those from health improvements (Crafts, 1997; Jones & Klenow, 2016).

The discussion of these four dimensions of well-being will highlight the exceptional character of the first half of the 20th century. Historical opportunities emerged to greatly improve Europeans' lives along with historically-low levels of economic performance which suggests that traditional income measures may be not only significantly underestimating progress in human welfare, but also providing a one-sided perspective on relative levels of living standards -high-income countries did not necessarily perform better in other dimensions, and vice versa. Therefore, to analyse the extent to which welfare opportunities across time and space were missed (or taken), I will apply a new welfare measure grounded in economic theory to a sample of twelve Western European countries and the United States. With this measurement framework, developed by Jones and Klenow (2016), this paper contributes to the literature on comparative historical living standards using the HDI and similar indices in two ways (Crafts, 2002; Prados de la Escosura, 2015). First, in this paper I take into account changes in key dimensions of well-being during the analysed period that are typically not considered in historical HDI-based studies, namely leisure and economic inequality. Second, and more importantly, the new welfare indicator applied in this study tackles the criticisms of the HDI mentioned above (i.e. tendency to convergence, arbitrary weighting scheme and constant weighting scheme over time) by drawing on information about individuals preferences and how, in the case of health, they change over time. Contrary to the HDI, this approach makes possible a direct comparison between the new composite indicator and income per capita since both of them are measured in the same units. This comparison will not only provide a so-far unexplored perspective on historical welfare levels taking a utility approach, but it will also assess the extent to which income per capita underestimates welfare growth by accounting for changes in health, leisure and inequality. This research is also related to another branch of the literature that has analysed welfare growth during the 20th century with utility indicators that can be traced back to the seminal study of Usher (1973). By incorporating gains from mortality reductions to gross national product, Usher concludes that the contribution of health to welfare from the 1910s to the 1960s is substantial since income

growth can be revised up to 40 percent upward (e.g. France).[4] Drawing on Usher's work and Beckerman (1980), Crafts (1997) applies a welfare measure that combines income, health and leisure to a large sample of developed countries for the period 1870-1992. For the first half of the 20th century, he finds that income per capita underestimates annual welfare growth by 1.2 percentage points on average (health contributes to this differential by two thirds and leisure by one third). While the welfare measure I use is similar to that in Crafts (1997) since they both include income, health and leisure (although not inequality), the approach I take differs from his study in several aspects. First, the analytical framework employed in Crafts (1997) only allows for estimating welfare growth rates. This leaves a number of questions unanswered regarding relative well-being levels between different regions in Western Europe (and the United States) that are key to understanding the effect of heterogeneous growth experiences on the distribution of welfare across space. Second, Crafts's analysis does not pay particular attention to the welfare dynamics within the first half of the 20th century (he only considers the benchmarks 1913 and 1950) as I do by including a further benchmark in 1929. Third, I measure material well-being with consumption per capita instead of income. Also, I use Gini coefficients to account for the unequal distribution of the economic status of the population within countries. Fourthly, to measure the contribution of health to well-being Crafts used life expectancy at birth and therefore does not take into account the welfare effect of changes in mortality across ages.[5] For this purpose, I use age-specific mortality rates.

Combining a methodology that allows for measuring consumption, health, leisure and inequality jointly with data for ten European countries and the United States, I show that the period 1913-1950 is one of fast welfare growth in Europe. Taking into account gains in welfare from health and leisure in the new indicator suggests that income per capita underestimates growth in European living standards by four percentage points annually (ranging from 5 to 3.2 percentage points in Italy and Denmark respectively). This differential has important implications for the study of welfare during the analysed period because with the new welfare measure well-being tripled, whereas income grew by less than 30 percent (the differential is still large if changes in the youngest part of the age distribution are ignored). However, progress in human welfare was not equal across space and, contrary to HDI-based evidence, by the end of the period cross-country differences had not narrowed because countries with relatively low welfare levels in 1913 (e.g. Italy or Spain) experienced similar growth rates as those with high levels (e.g. Sweden or the Netherlands). These findings suggest that the opportunities that emerged during the late 19th and early 20th century to improve the lives

[4]Improved versions of Usher's model have been developed and applied exclusively to the United States in Costa and Steckel (1997), Nordhaus (2003) and Murphy and Topel (2006). See also Becker, Philipson, and Soares (2005) for an analysis of world welfare dispersion since 1960 considering income and life expectancy.

[5]Crafts (2007) did consider age-specific mortality, but only for the British case.

of millions were taken unequally and did not result in a more equal distribution of welfare across countries.

2 A period of missed opportunities for well-being?

In this section I will present the data used to measure the four welfare aspects that will be later considered in the composite indicator. By first discussing material well-being, I will establish some basic patterns that will quantify the extent of missed (or taken) economic opportunities across countries and time in Europe. These patterns will then be compared with those of economic inequality, health and leisure to both highlight their mismatch and the need to use a framework that integrates them, if we are to analyse the evolution of European welfare beyond income during the first half of the 20th century.

Material well-being

Before reviewing the growth experience of Western European economies during the period 1913-1950, it is instructive to know their starting levels.[6] In Table 1 (Columns I and II) I present the levels of income per capita relative to the United States for Western Europe and three subregions: Northern Europe, North-Western Europe and Southern Europe (since this paper focuses on citizens' average living standards, all regional averages are weighted by population).[7] At the beginning of the 20th century, the income level of an average Western European was two thirds of the American level.[8] Within the old continent, citizens in the industrial core enjoyed higher levels of material well-being than those living in peripheral

[6]To discuss economic development in Europe I will focus on a ten-country sample: Spain, Italy, France, Germany, Switzerland, Belgium, The Netherlands, the United Kingdom, Sweden, Denmark and the United States. This sample was chosen on the basis of data availability for (at least) 1913 and 1950 on household consumption, age-specific mortality, annual working time and Gini coefficients. For a more comprehensive exposition of European economic performance during the analysed period see Roses and Wolf (2010).

[7]This regional classification will be used in the remaining of the paper and it is defined as follows. Northern Europe refers to Sweden and Denmark. North-Western Europe refers to the United Kingdom, Germany, Belgium, the Netherlands, France and Switzerland. Southern Europe consists of Italy and Spain. The regional aggregates are weighted with population data from Maddison (2006).

[8]Following the literature, the periods for which I report levels and growth rates of income (and consumption) are chosen to avoid the disruptive effect of the world wars and the Great Depression: 1913, 1920, 1929, 1938 and 1950. Moreover, to make these figures more robust to annual fluctuations I constructed these benchmarks by taking averages. For 1913 and 1920, I considered the time spans 1911-1913 and 1920-1922 respectively. For 1929 and 1938, I took 1927-1929 and 1936-1938 to have a benchmark before and during the Great Depression (in the Spanish case, I considered 1933-1935 due to the civil war that took place in the period 1936-1939). 1950 is the average for the years between 1948 and 1952. In a strict sense, some of these averages refer to the year in the middle of the periods considered (i.e. 1912, 1921, 1928 and 1937) and not to the ones I refer to in the tables and the text. However, since these benchmark years are widely used in the literature and taking averages makes the analysis more robust, in the remaining of this paper I will continue to use them, even though strictly speaking they refer to a year earlier or later.

economies. Actually, average income in North-Western Europe almost doubled that of the less-industrialised southern periphery, where more than 50 percent of the labour force was employed in agriculture (Buyst & Franaszek, 2010). Furthermore, differences in economic development were not only substantial between regions, but also within them as illustrated by the 40-percent higher income per capita in the United Kingdom with respect to France (see Table 4 in the Appendix).

The large income gaps with respect to the United States suggest that the growth potential in many European countries was quite substantial. To look at the extent to which this potential was used, Table 1 also shows growth rates for the whole analysed period 1913-1950 (Column V) and two sub-periods: 1920-1929 and 1929-1938 (Column VI and VII respectively). During the first half of the 20th century, material well-being in Europe increased below one percent yearly, whereas at the other side of the Atlantic it grew more than two times faster (if we consider other countries overseas such as Canada or New Zealand instead of the United States, the same pessimistic pattern emerges). These growth rates also stand out negatively in comparison with other historical periods such as the post-1950 decades or the years leading up to World War I when income growth was five and almost two times faster respectively (Bolt & van Zanden, 2014). This evidence is indicative of a period of substantial missed economic opportunities during the period 1913-1950 in Europe from both an international and a long-term perspective.

However, this average should not be taken to summarize this period's economic performance because it masks substantial variation across countries and time. To examine the evolution of income throughout the first half of the 20th century, Columns VI and VII in Table 1 present growth rates for 1920-1929 and 1929-1938. On both sides of the Atlantic, the years of the Great Depression stand in sharp contrast with those of the 1920s. Especially in the United States and Southern Europe, the rate of progress in material welfare came to a halt (or even worsened in North America). In the rest of the old continent, the difference between the two periods was still significant, although much smaller. The growth differentials between these two periods cannot be fully explained by post-war reconstruction in the 1920s because growth was still large even after reaching pre-war income levels and some good-performing countries such as Sweden or Switzerland were not involved in the Great War. Moreover, and more importantly, the efficiency of European economies increased substantially in the 1920s as rapid increases of total factor productivity suggest (Roses & Wolf, 2010). These productivity increases were the result of the diffusion of a number of unused technological opportunities that emerged before and during the war such as the internal combustion engine and the application electricity as the case of the United Kingdom or France illustrate. In these two countries, the number of motor vehicles per inhabitant and electricity production

in the 1920s increased by more than a factor of three and six respectively (Svennilson, 1954; Mitchell, 2003). Partially as a consequence of these factors, economic growth in the interwar period was faster than in the U.S. in all subregions of the old continent. Nevertheless, the overall growth rate for the period 1913-1950 is lower than the American one mainly due to the destructive effect of the Second World War as it can be clearly seen in the large fall in relative income (23 percent with respect to the U.S.) from 1938 to 1950 (see Table 4 in the Appendix). This relative decline is not driven by the specific experience of the United States since it still holds when we consider Canada, Australia or New Zealand (Bolt & van Zanden, 2014).

Did all regions in the old continent lag further behind other countries overseas by mid-century? As Column II and V show, there is a large degree of spatial heterogeneity. For example, material living standards in Northern Europe grew slightly faster than in the United States since countries such as Sweden or Finland industrialised and oriented towards higher value-added industries (Krantz, 1987; Hjerppe & Jalava, 2006). In the remaining regions, despite economic growth during the interwar period was faster than at the other side of the Atlantic, the destructive effect of the Second World War vanished economic progress relative to the United States.

Table 1: Income and consumption per capita in Western Europe and the United States, 1913-1950

	Level (US=100)				Annual growth rate (in %)					
	Income		Consumption		Income			Consumption		
	1913	1950	1913	1950	1913-50	1920-29	1929-38	1913-50	1920-29	1929-38
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)
United States	100	100	100	100	1.7	2.2	-0.7	1.3	2.3	0.2
Western Europe	67	49	61	47	0.8	2.5	0.8	0.6	1.9	0.4
Subregions										
Northern Europe	60	69	65	75	2.0	2.6	2.1	1.7	1.5	2.4
North-Western Europe	77	55	67	51	0.8	2.7	1.0	0.6	2.1	0.6
Southern Europe	42	30	41	32	0.7	2.2	0.0	0.6	1.5	-0.4

Note: income and consumption data were taken from Bolt and van Zanden (2014) and Barro and Ursua (2008) respectively. For detailed information on the countries within each region and the construction of the benchmarks, see footnotes 7 and 8. The figures for Western Europe and its subregions are weighted by population using Maddison (2006). Levels are expressed relative to the United States in that same year and growth rates were obtained by taking the logarithmic difference between benchmark years.

As mentioned above, I will use household consumption instead of income per capita to measure material well-being as a more closely related indicator of the economic status of individuals because it captures the share of the national product that is actually consumed

by households (e.g. clothing, food, rent, etcetera). Furthermore, this measure is not affected by some unusual dynamics between government spending and GDP during this period. For example, rising political tension between the wars in the United Kingdom boosted military spending and income growth (Barro & Ursua, 2008). Similarly, if in the German case we consider consumption instead of income per capita, growth in material living standards are cut in the 1930s from 1.7 to almost zero percentage points.[9] This downward revision is the result of excluding the effects on material living standards of the German rearmament policy based on government debt and low nominal and real wages (Ritschl, 2002).

Are the patterns observed for income in line with those of household consumption? To answer this question, I also include information on household consumption per capita in Table 1 from Barro and Ursua (2008).[10] As Column III indicates, the two economic measures point to a similar situation at the beginning of the analysed period since consumption per capita in the old continent was about two thirds of the American level and the industrial core is the best-performing region (although by a small margin). During the period 1913-1950, material well-being increased slightly less than shown by income on both sides of the Atlantic. This is especially the case in the United States which translates into a less divergent picture by mid-century (see Column IV).

Within the old continent, citizens in Northern Europe were the only ones that experienced more progress in material welfare than the United States during the whole period. The annual growth differential between the two regions of 0.3 percentage points improved the relative position of this region by ten percent with respect to the U.S. at the end of the period. In the remaining regions, consumption growth was only above the United States in the 1930s (with the exception of the southern periphery where material well-being actually worsened). Nevertheless, as for income per capita, progress in relative terms vanished after the Second

[9]Germany's performance in the 1920s is also somewhat atypical because it exhibits a much larger consumption growth rate (four percentage points annually) than any of the other economies considered. This is mainly explained by post-war catch-up because consumption growth after the pre-war level was attained (around 1925) was in line with those of its neighbouring countries.

[10]To test that the trends in these data are the same as in other sources such as the Penn World Table (PWT), in the Appendix I compare the consumption indices in Barro and Ursua (2008) with those of the PWT (Feenstra, Inklaar, & Timmer, 2015). Since the latter source does not provide household consumption in 2011 prices, but household and government consumption put together, the comparison serves a two-fold purpose. First, it provides a useful check on the trends exhibited by the data used in this study for the post-1950 period. And second, given that I do not include government spending for the reasons mentioned above, it is instructive to analyse how a consumption index that includes this component would behave as opposed to one without it. Table 1 shows that for the European aggregate the absolute difference between the consumption index with and without government spending is almost zero. At a regional level, this holds except for Northern Europe (where up to the 1970s the difference between them is around ten points). The reason for this is that in Denmark the ratio of household and government consumption experiences a continued fall until 1982 as a result of a an exceptional doubling in government spending. Taken together, both data sources (and indicators) show the same pattern.

World War and all countries were pushed back to relative levels lower than those in 1913, except for Sweden and Denmark (see Column II and VI in Table 5 in the Appendix). As a result, by mid-century an average Western European had a consumption level lower than half of that in the United States, whereas in 1913 it had been almost two thirds.

Until now, the income and consumption trends I have presented are national averages and do not reflect their distribution within the population. If one aims at measuring the overall level of living standards in a given society, relying on just an average value can be problematic. For example, if the increasing levels of average consumption (or income) reported in Table 1 were the result of the concentration of economic resources by a small group of people, then a higher average income level would not necessarily reflect an overall increase in material well-being in the society. To tackle this potential issue, I will take this into account by using data on a widely used inequality measure: the Gini coefficient.

The source I used is the extensive cross-country dataset created by van Zanden et al. (2013) on income inequality since 1820.[11] In Table 2, I present the Gini coefficients for the regions defined in Table 1 during the period 1913-1950.[12] Throughout the first half of the 20th century, we can distinguish two different time periods. The first time span comprises the years leading up to the Great Depression and it is characterised by a modest rise in inequality across the board (except for Northern Europe). After 1929, income inequality drops substantially without exceptions (compare Columns II and III in Table 2). Atkinson (2007) characterises the decades of the 1930s and 1940s as a period of compression in the distribution of earnings in countries such as the United States, Germany or France. This general declining trend during the period 1913-1950 (that also holds throughout the whole sample, except for the Swiss case, as Table 6 in the Appendix shows) indicates that despite material well-being was growing relatively slow by historical standards, its distribution among the population was becoming significantly more equal. To make sense of the magnitude of this development, we can observe that the decrease in the Gini coefficient in Northern and

[11] Ideally, I would like to have consumption-based Ginis, instead of income-based as reported in van Zanden et al. (2013), since the welfare indicator I will construct later on uses consumption to measure material well-being. However, this should not prevent us from using this dataset since the main goal of including income inequality is rather broad, namely to take into account the distribution of economic opportunities across countries and time (and not necessarily that of income or consumption specifically). Moreover, in the inequality literature Gini coefficients based on consumption and income data are sometimes compared with each other without any adjustment (Lakner & Milanovic, 2015, 6).

[12] It should be noted that in van Zanden et al. (2013) data for Germany are not available. To fill this gap, I draw on the World Income Inequality Database (WIID). However, since the earliest Gini coefficient provided for Germany refers to 1936, I used this value for 1913 and 1929. This can be seen as a very conservative approach since income inequality might have been higher according to the experience of the rest of European countries (van Zanden et al., 2013; Atkinson, 2007). In quantitative terms, this assumption has no impact on the overall trends of North-Western and Western Europe shown in Table 2 because the weighted averages of the Gini coefficients for both regions throughout the period are virtually the same if I include or exclude Germany.

Southern Europe was as substantial as the current difference between Ethiopia and France or South Africa and Australia (van Zanden et al., 2013). Interestingly, this overall trend after 1929 suggests a more optimistic picture than the slowdown (and sometimes worsening) of material welfare discussed before. A further remarkable point concerns comparative levels of economic development and inequality. Comparing Tables 1 and 2, we can observe that there is not an obvious link between inequality and income (or consumption). For example, in 1913 the Gini coefficients for North-Western and Southern Europe are almost the same even though average income in the former region almost doubles the level of the latter.

Table 2: Gini coefficients in Western Europe and the United States, 1913-1950

	1913* (I)	1929 (II)	1950 (III)
United States	51	54	39
Western Europe	47	49	41
Subregions			
Northern Europe	51	48	39
North-Western Europe	48	50	41
Southern Europe	44	46	40

Note: the sources are van Zanden et al. (2013) and WIID for Germany (see footnote 11). See Table 1 for the country composition of each subregional category. *Data for 1913 refers to 1910.

If we look at the Gini coefficients at both sides of the Atlantic, the data point to an inequality reversal during the analysed period. Until the eve of the Great Depression, Americans lived in a more unequal society than Europeans as implied by the four- and five-point Gini differential in 1913 and 1929 respectively. By mid-century though, Western Europe led the way in this aspect due to the larger fall of the Gini coefficient in the United States. This significant decrease in inequality was the result of declining unemployment and narrowing wage structure after the great depression and the war (Goldin & Margo, 1992).

Health

Since the late 19th century, Europeans' lives were undergoing a rapid and unprecedented improvement in terms of mortality. The diffusion of the germ theory of disease created a whole array of new opportunities to reduce mortality from a number of infectious diseases.[13] At the household level, the application of this knowledge allowed the adoption of more hygienic

[13]The number of studies tackling the topic of the causes of the mortality decline is unsurprisingly large. See Cutler et al. (2006), Millward and Baten (2010), Leonard and Ljungberg (2010) or Costa (2015) for a more comprehensive review.

practices that reduced contagion of infectious illnesses. Especially in the field of child care, mothers received better advice and education on the risks of certain feeding practices and the importance of a clean environment for the survival of children (Mokyr & Stein, 1996). In the public sphere, the exposure to illnesses declined substantially as a result of improvements in water supply and sewerage systems when public authorities embraced the germ theory of disease and rejected alternative explanations such as the miasma theory. In the United States, cities that implemented water filtration and chlorination experienced large declines in mortality of waterborne infectious diseases, and (to a lower extent) airborne illnesses due to the diffuse effects of water quality. By reducing the incidence of diseases such as typhoid, that greatly weakened those who suffered from it, the population became less vulnerable to other illnesses (Cutler & Miller, 2005; Ferrie & Troesken, 2008). In the old continent, the establishment of water and sewage systems intensified in the late 19th century, especially in England, France and Germany. In the south of Europe, they would be implemented only some decades later (Bell & Millward, 1998; Leonard & Ljungberg, 2010).

To account for the extent to which health opportunities were taken across countries and time, I will use data on age-specific mortality rates that measure the probability of dying at a certain age. With these data at hand, it is possible to calculate survival rates (i.e. the probability of a person reaching a certain age given the mortality rates of all ages in that same year) or life expectancies (i.e. the expected life time of an individual at a given age, for example at birth, given the mortality rates across all subsequent ages in that same year).[14] For most of the countries in the sample, the main source is the Human Mortality Database (HMD) that provides annual life tables with the exception of Germany, the United States and the United Kingdom, for which the series only start in 1956, 1933 and 1922 respectively. To fill these gaps, data prior to 1923 for the United Kingdom refers to England and Wales. In the case of the United States and Germany, I used the Human Life-Table Database (HLTD).

Considering the age distribution of mortality it is useful to isolate mortality changes (and therefore health improvements) in different parts of the age spectrum, especially at the youngest ages since these have a large effect on widely-used measures in other studies such as life expectancy at birth. This is particularly important during the analysed period because infant and child mortality declined enormously as a result of progress against airborne diseases like influenza or pneumonia for children less than five, and water-borne diseases such as gastroenteritis for infants (Millward & Baten, 2010). To provide an illustration of the magnitude of the mortality decrease below the age of five, consider the British case.

[14]Following similar studies in the literature, the mortality measures I will use are based on period instead of cohort life tables. The difference between them is that the former conveys information on the mortality experience of a fictitious cohort in a given year (or period), and the latter provides information on the mortality experience of an actual cohort from birth.

Whereas at the beginning of the period 22 percent of a given cohort were expected to die before reaching the age of five (given the mortality rates in that same year), in 1950 this figure dropped to only three percent. The impact of this development on life expectancy at birth is great since its increase during the period 1913-1950 is much higher than if we consider life expectancy at age five (15 and eight years respectively). Thus, the impression of health improvements will be very different if we consider the perspective of a newborn or the one of a person at a different age. At this point, a question arises concerning the perspective that I should take while looking at health differences across countries and time since these can differ significantly depending on the age of reference. To put it in another way, should I take the perspective of a newborn (as common in the literature) or that of someone at a different age? Choosing a particular age threshold depends on the health aspect of the population that the researcher wants to analyse and on these grounds there is not a unique and desirable perspective. To avoid arbitrary choices, in this study I will consider the whole age spectrum, although (for the sake of simplicity) in the following I will discuss the health status of countries across time and space by taking the perspective of a child at birth, five and ten. Table 3 shows these data for 1913 and 1950.[15]

Table 3: Life expectancy at birth, five and ten in Western Europe and the United States, 1913-1950

	Life Expectancy					
	At birth		At 5		At 10	
	1913 (I)	1950 (II)	1913 (III)	1950 (IV)	1913 (V)	1950 (VI)
United States	54	68	57	66	53	61
Western Europe	50	66	57	66	53	61
Subregions						
Northern Europe	58	71	60	68	56	63
North-Western Europe	51	67	57	66	53	61
Southern Europe	45	64	56	65	52	60

Note: see the text for the sources. See Table 1 for the country composition of each subregional category.

In contrast with material well-being, the improvement of citizens' health during the period 1913-1950 is quite remarkable. In less than four decades, the years a newborn was expected to live in the old continent increased by 16 (Column I and II). This development retains

[15]As for income and consumption, I took three- and five-year averages in 1913 and 1950 to make the benchmarks more robust to unusual figures in a certain year. Given the data availability, this could be done for all countries but Germany (although the German life tables refer to the periods 1910-1911 and 1949-1951).

its exceptional character even after ignoring health progress below the age of five and ten, in which case life expectancy gains were eight and nine years respectively.[16] As highlighted with the British example, the increase in life expectancy at birth is larger than at other ages because of the great reduction in mortality of children before the age of five. This point can also be seen by comparing life expectancy at different ages in a given year and seeing how this difference evolves over time. For example, in 1913 the expected lifetime of a five-year-old in Europe was seven years larger than that of a newborn due to the disproportionately-high mortality rates during the first years of life (in Southern Europe, this was particularly extreme since surviving the fifth birthday increased life expectancy by eleven years). By mid-century, after decades of child mortality decline this unusual pattern for today's standards disappeared everywhere, but in the south of Europe.

The improvement in life expectancy at birth is not only substantial on its own, but it also stands out historically. To continue with the British example, the increase of the first half of the 20th century accounts for more than 60 percent of the total increase during the 20th century (below the age of five and ten this figure is roughly 50 percent).[17] Contrary to income, progress was widespread in the countries considered since, as Table 3 shows, citizens' health in Western Europe did not fall behind the United States at the end of the period. Despite the negative effects of political tensions, increasing protectionism and the world wars on economic welfare, the increasing trend of citizens' health did not come to a halt over this period and by 1950 European newborns were expected to live just two years less than their American counterparts. This apparent lack of correlation with income measures also holds true during the Great Depression and the 1940s when the upward trend of health continued steadily, whereas material well-being stagnated or worsened in many countries (see Table 7 in the Appendix) partially due to the emergence of new medical therapies (e.g. new vaccines and, most importantly, the coming of antibiotics). In the United States between 1937 and 1943, sulfa drugs account for up to a third of the decline of maternal mortality and pneumonia; and up to two thirds of scarlet fever (Jayachandran, Lleras-Muney, & Smith, 2010).

Finally, if we compare health and income levels, an interesting pattern emerges. Europeans with the longest expected life span at birth are not necessarily living in the richest region (i.e. the industrial core), but rather in the northernmost area. Northern Europe even outperforms the United States despite having much lower levels of income per capita through-

[16]The sensitivity of this indicator to changes in the age of reference drops substantially after the age of five, which indicates that the mortality decline afterwards is more equally distributed than between the age bracket 0-4.

[17]Had this increase been linear, the analysed period would only account for 43 percent of the whole increase. These calculations were made by taking life expectancy at different ages for the countries in my sample in 1913, 1950 and 2000 from the HMD.

out the whole period (this also holds if we look at life expectancy at five and ten). However, the link between health and income should not be dismissed since the escape from chronic malnutrition contributed to improve the physical condition of human bodies and their ability to deal with diseases (Fogel, 2004).

Paralleling what we observed in terms of life expectancy increases differing at birth and other ages, Table 3 shows that the impression of health gaps differs notably depending on the perspective we take. If we consider life expectancy at birth, Southern Europe has a life expectancy differential of nine years with the United States (Column I). Changing the age of reference in the indicator to five (or ten) reduces this difference to only year (Column III). However, despite taking the view of a five-year-old individual provides a more egalitarian perspective, the patterns observed are still the same. Citizens in Northern Europe have the best health status followed by those in the industrial core and then Southern Europe. Also, by the end of the period life expectancies converge across the board.

Working time

An aspect of people's lives that changed dramatically during the first half of the 20th century (not considered in the HDI) is working time or leisure.[18] Using information on annual hours of work from Huberman and Minns (2007), we can see in Table 4 that Europeans witnessed a sizeable decline in annual working time of roughly 650 hours throughout the period 1913-1950.[19] Paralleling the evolution of mortality, this development stands out on its own and historically for it was the most pronounced decline in market work since 1870 (17 hours annually).[20] Most of the decline (90 percent) happened between 1913 and 1929 (Column I and II). This can be explained by the increasing number of vacations and national holidays during the interwar period as the annual paid vacation became a reality in many countries (Huberman & Minns, 2007, 546). Interestingly, most progress in this respect was achieved in the 1930s when economic circumstances were less buoyant when the idea of the right to

[18]Obviously, non-working-time is not the same as leisure. To account for changes in the latter more precisely, one would have to distinguish between activities that involve leisure and home production. However, given the data availability and the focus of this study on other dimensions of well-being (not only leisure), I will provide an approximation of the contribution of leisure and home production to welfare by looking at non-market work as done in Crafts (1997). Therefore, despite strictly speaking I am not measuring leisure but a combination of the two of them, I will refer to this as leisure for the sake of simplicity.

[19]The unit of measurement is annual hours of full-time production (male and female) workers engaged in non-agricultural activities. Huberman and Minns (2007) interpret their series as the 'approximate usual or normal hours the representative production worker would have been engaged for during the year' (page 543).

[20]I obtained this figure by using the information for Europe in Table 3 from Huberman and Minns (2007, 548), row 'Old W. (weighted)', and splitting the period 1870-2000 into four subperiods: 1870-1913, 1913-1950, 1950-1973 and 1973-2000. The yearly decrease measured for the four subperiods (in this order) is 7, 17, 12 and 11 hours annually. If we measure the decline in percentage terms, the same pattern emerges as the decreases are 0.25, 0.70, 0.60 and 0.66 percent per annum.

leisure gained support from international institutions such as the International Labour Organisation (ILO), work science journals and public authorities (Cross, 1989). Besides this, an element that changed workers' lives even more dramatically was the reduction of the working week with the introduction of the 48-hour week (or 8-hour day). This is a process that can be traced back to the 19th century when social reformers advocated the 8-hour day and their claims turned into one of the main demands of the Workingmen's Association at its first congress in 1866. By 1913, only a few occupations had achieved this (e.g. miners in the United Kingdom) since industrial workers in France, Germany, the United Kingdom or the United States typically worked ten hours a day. When the First World War came to an end, workers' requests were heard and governments promised to introduce the 8-hour day after working long hours to meet the demands of the war economy had proofed to be not very productive (Evans, 1969). Also, the ILO embodied this idea and incited countries to adopt it at its first conference in 1919 (Huberman & Minns, 2007). At the end of that year, countries such as Denmark, France, Italy, Netherlands, Norway, Spain and Switzerland had passed 8-hour laws; by 1922 the 48-hour week in industry was common throughout Europe and by 1930 it consolidated (Evans, 1969).[21]

Table 4: Annual hours of work in Western Europe and the United States, 1913-1950

	1913 (I)	1929 (II)	1950 (III)
United States	2900	2316	2008
Western Europe	2783	2200	2140
Subregions			
Northern Europe	2740	2206	2032
North-Western Europe	2769	2192	2209
Southern Europe	2829	2222	1989

Note: data were taken from Huberman and Minns (2007), Table 3 (page 548). See Table 1 for the country composition of each subregional category.

If we compare Western Europe with the United States, we see that workers in the old continent worked fewer hours (117 hours per year) at the beginning of the period, although this situation was reversed by 1950 (Columns I and IV). Only after the 1970s would working time be lower again in Europe (Huberman & Minns, 2007). Across regions, there seems to be another reversal. Whereas southern Europeans worked longer hours than in the rest of the continent before the outbreak of the First World War, by mid-century annual working

[21]The international character of this process is even more clearly visible at the country level as Table 8 in the Appendix shows.

time was the shortest in this region. In 1950, workers in the industrial core had the lowest amount of time for leisure activities.

The analysis of European material well-being at the beginning of this section pointed out that the period 1913-1950 can be characterised as one of missed opportunities in material living standards across time and space because income (and consumption) growth were historically low and significantly lower than in other countries overseas (e.g. United States, Canada or New Zealand). If we extend the analysis of citizens' well-being beyond their economic status, these patterns do not match. Taking France as an example, by mid-century citizens lived in a drastically different society than they (or their counterparts) lived in 1913. In terms of health, citizens lived and worked in a much safer environment as more hygienic practices were adopted and efficient central water supply and waste disposal systems greatly reduced the exposure of the population to infectious diseases. Moreover, with the coming of antibiotics illnesses that had taken the lives of many for decades such as tuberculosis could finally be treated. In terms of leisure, the demands for shorter working weeks by 19th-century social reformers became a reality and French workers spent almost 900 annual hours less in the workplace than at the beginning of the century. What did these changes mean to French citizens in 1950? If they had been confronted with the choice of living in the same society as their counterparts did in 1913, how would they have valued the achievements in human welfare since then? How does this view compare to the more traditional view provided by income measures? To answer these questions, in the next section I will present a new measurement framework developed by Jones and Klenow (2016) that will be the basis of the welfare calculations.

3 Methodology

In the spirit of the veil of ignorance emphasized in Rawls (1971), the framework developed in Jones and Klenow (2016) theoretically confronts an individual with a lottery. She does not know in which country she will live, the level of consumption she will enjoy, or whether her life will be expected to be long and full of leisure. What is the proportion of her yearly consumption living in the U.S. that would make her indifferent between living there and, say, in France? Equivalently, for time comparisons we could also ask: what is the proportion of her yearly consumption living in France in 1950 that would make her as well off as her counterpart in 1913? The answer to these two questions is a consumption-equivalent measure of the standard of living and this is what I will use to calculate welfare differences across

countries and over time.[22]

To make welfare comparisons, this methodology draws on information about individual preferences. Using these provides a solution to two important critiques of the HDI. First, the consumption-equivalent indicator avoids assuming an arbitrary weighting scheme because empirical evidence about people's preferences on trade-offs between different components of well-being gives information about their relative importance. Second, assuming a fixed weighting scheme over time, as the HDI does, has the implication that the relative importance of each dimension does not differ over time. In the case of income and health, this can be problematic because, as Costa and Kahn (2004) shows, the value individuals assign to health has increased between 1940 and 1980 in the United States as measured with the compensating differential for job risk. Therefore, assuming constant weights over long time spans can be troublesome in the case of health because its weight in the indicator will be very different if we take the perspective of an individual today or in the past. To take this into account, the methodology used in this paper considers a representative individual with a given set of preferences in 1950.[23] Following Jones and Klenow (2016), the lifetime welfare of our individual is determined as follows:[24]

$$U = \mathbb{E} \sum_{a=1}^{100} \beta^a u(C_a l_a) S(a) \tag{1}$$

where $S(a)$ is the probability that the individual is alive up to age a, β is the discount factor, C is annual consumption and l is leisure. In this framework, well-being is given by the expected value (due to uncertainty) of things that matter to our individual such as the amount of goods and services consumed and the number of hours spent working throughout her lifetime. At this point, it is important to clarify the aspect of health that is being measured with survival rates. Murphy and Topel (2006) argues that health-related knowledge can affect the quality and (or) the quantity of life. With $S(a)$, I am mainly measuring improvements that affect the quantity of life (i.e. mortality). On the other hand, one could also argue that mortality rates at young ages can also capture some aspects of the quality of life because

[22]Following Jones and Klenow (2016) I will also refer to the consumption-equivalent indicator as *welfare*, *well-being* or *living standards* for the sake of simplicity. Also, strictly speaking, the example above refers to one of the two ways of calculating this measure: the equivalent variation. To ease the explanation of the methodology in this section, I will focus on this variation. See the Appendix for further elaboration on the differences between the two.

[23]Following the literature, preferences across countries are kept constant. Jones and Klenow (2016, 2429) notes that a similar issue arises in cross-country comparisons using income per capita, since this indicator requires a set of common prices. Moreover, these comparisons become more complicated as we consider countries in very different stages of economic development. Given that income in my sample differs at most by a factor of four, in my analysis this issue is less problematic than in other studies where income can differ by a factor of more than 40.

[24]The mathematical notation is the same as in their article.

children are more vulnerable than adults to the disease environment, and a bad disease environment leads to high morbidity. Nevertheless, measuring improvements in the quality of life is beyond the scope of this paper and for the discussion of the results I will consider that lower mortality rates enhance well-being by increasing the prospects of citizens to have a longer and richer life in terms of consumption and leisure.[25]

To assess differences in living standards across countries with this methodology, we first need to choose a benchmark with which welfare comparisons will be made. By doing this, welfare levels are expressed in relative terms in the same manner as the difference in income between two countries can be expressed in percentage terms. Since one of the main purposes of this paper is to put the European welfare experience into perspective with that of other developed countries that were less affected (in economic terms) by the disruptive happenings of this period, the benchmark for welfare comparisons will be the United States.[26] With this country as a benchmark, we can extend Equation 1 as follows:

$$U_i(\lambda) = \mathbb{E}_i \sum_{a=1}^{100} \beta^a u(\lambda C_{ai}, l_{ai}) S_i(a) \tag{2}$$

where i indexes countries and λ multiplies consumption at every age. Given that our individual lives behind the veil of ignorance and she does not know in which country she will live, what her consumption level, health and leisure she will have: What proportion of the individual's annual consumption in the United States would make her indifferent between living in the United States and France? The answer to this question (λ in Equation 2) can be formally expressed as follows:

$$U_{us}(\lambda_i) = U_i(1) \tag{3}$$

To apply this theoretical framework, we first need to choose the function that will determine an individual's welfare:

[25] See Murphy and Topel (2006) and Hickson (2014) for such an attempt.

[26] As I discussed in the previous section, European economic performance was not only below that of the United States, but also of other developed economies such as New Zealand, Canada, Australia (or the weighted average of these four countries). Therefore, choosing the United States as a benchmark for welfare comparisons does not misrepresent the relative performance of Europe throughout this period. Furthermore, it is worth highlighting that this choice does not affect comparisons between countries and regions that do not involve the benchmark country (e.g. the percentage income difference between Spain and Denmark is independent of whether we choose the United States or Canada as a reference country). And, more importantly, using the United States to calibrate the model makes my results comparable to other studies in the literature that have used American data to calibrate their indicators (Becker et al., 2005; Murphy & Topel, 2006; Hickson, 2014; Jones & Klenow, 2016).

$$u(C_i, l_i) = \bar{u} + log\ C_i + v(l_i) \tag{4}$$

where $log\ C_i$ is log-transformed consumption in country i (this form reflects diminishing marginal returns as common in other welfare indices such as the HDI or the HIHD), $v(l)$ is the value of non-working time and $\bar{u}$ is a constant.[27] From a pure materialistic point of view, an individual would only care about consumption and therefore annual well-being in the previous equation would only depend on C. However, by introducing $\bar{u}$ and $v(l)$ Equation 4 also captures other aspects of the individual's life. To take into account economic inequality, this framework incorporates the uncertainty concerning the unequal distribution of consumption among the population.[28] Jones and Klenow (2016) includes this channel by considering that expected consumption decreases when its variance increases: $\mathbb{E}\ [log\ C] = log\ c - \sigma^2/2$.[29] Taken together, the expected lifetime utility for an individual is given by:

$$U_i = [\sum_{a=1}^{100} \beta^a S_i(a)] \cdot (\bar{u} + log\ c_i + v(l) - \frac{1}{2} \cdot \sigma_i^2) \tag{5}$$

It is interesting to highlight the interaction between survival rates and the rest of the dimensions. Suppose there is a certain improvement in survival rates. If living conditions in country i are characterised by low levels of material-well being and leisure, then the increase in welfare (U_i) would be lower than if the individual lived in a rich country with high levels of leisure. As a result, and similar to the HIHD developed by Prados de la Escosura (2015), health improvements at high levels of other dimensions of well-being represent higher achievements than at low levels. Finally, to calculate the consumption-equivalent measure (λ) we can combine Equations 5 and 3:

[27] Note that the observed level of consumption in a certain year is assumed to be the same throughout the individual's life. Assuming this makes the indicator more intuitive because it provides a point-estimate of expected welfare given the level of consumption, health, leisure and inequality in a given year. In a way, this approach resembles that of period life tables in that they present the expected health outcomes of a fictitious cohort according to the mortality experience of the whole population in one specific year. Moreover, this assumption does not change the conclusions of this study as shown in the robustness tests performed in the Appendix where I present a richer version of the model that considers different consumption paths.

[28] It should be noted that the effects of inequality on welfare are not trivial and in this paper I just account for one mechanism through which it may affect individuals. Further research is needed for accounting for this well-being dimension in composite indicators.

[29] This holds if we assume that consumption is log-normally distributed (with arithmetic mean c_i and a variance of log consumption of σ_i^2) and independent of mortality and age.

$$
\begin{aligned}
log\ \lambda_i = \sum_{a=1}^{100} \frac{\beta^a [S_i(a) - S_{us}(a)]}{\sum_{a=1}^{100} \beta^a S_{us}(a)} [\bar{u} + log\ c_i + v(l_i) - \frac{1}{2} \cdot \sigma_i^2] \\
+ log\ c_i - log\ c_{us} \\
+ v(l_i) - v(l_{us}) \\
- \frac{1}{2}(\sigma_i^2 - \sigma_{us}^2)
\end{aligned} \tag{6}
$$

This equation shows that welfare in country i relative to the United States consists of four components. The first measures the contribution of differences in health to relative welfare, the second term captures differences in consumption and the last two measure differences in leisure and inequality.

For comparisons of welfare over time, Equation 6 is also used. However, instead of comparing countries with the United States in a given year, a country is compared with itself in two points in time. If we, for example, want to calculate welfare growth in France during the period 1913-1950, λ_i in Equation 6 would answer the question: by what percentage must the individual's consumption be adjusted in France in 1950 so that she is as well off as her counterpart in 1913? In other words, by which factor do we have to adjust our individual's consumption in 1950 so that she is willing to give up on working much shorter hours and living in a society where infectious diseases have almost disappeared?

One of the strengths of the presented methodology is that it allows for a direct comparison with GDP. In terms of growth rates, it is possible to calculate the extent to which income per capita underestimates welfare growth and provide a decomposition of the sources for the discrepancy between the two measures (i.e. consumption, health, leisure or inequality). If we perform this exercise in terms of levels, it would show whether the welfare measure shows a higher or lower level of well-being in a certain country than income and, as before, the contribution of each dimension to the difference between them.

To implement the welfare calculations using Equation 6, I will use the data on consumption, health, working time and inequality presented in the previous section.[30] It is worth noting that by excluding government consumption, the indicator is missing public spending on education or health that can be an important part of people's welfare. For 1950, robustness tests calculating welfare levels including and excluding government spending to-

[30]See the Appendix for country-specific data except for age-specific mortality rates since these would take too much space (instead, I provide life expectancy at birth). Note that, as common in the literature, the Gini coefficients are converted into the standard deviation of log consumption inverting the formula suggested in Aitchison and Brown (1957): $G = 2\Phi\frac{\sigma}{\sqrt{2}} - 1$ (where $\Phi(.)$ is the cumulative distribution function of the standard normal distribution).

gether with a comparison between the trends in total and household consumption show that the quantitative effect of this issue on the results is negligible (see Tables 1 and 3 in the Appendix). A further issue related to the use of household consumption instead of income per capita is that it ignores people's decisions on delaying present consumption to save and increase it in the future. As a result, some countries might exhibit relatively low levels of consumption because of high investment levels. While this issue has to be taken into account when interpreting the level of material well-being across countries, not including investment better serves the purpose of this paper because it aims at capturing the expected welfare of individuals in a specific point in time. If the level of investment, say, in 1913, was relatively high in a certain country, then a higher level of consumption would be observed in 1950.

Moreover, it is necessary to calibrate the model according to the preferences of a representative individual in 1950 to value changes in leisure and health.[31] For calculating the contribution of leisure to welfare, the first step is to create and indicator to measure it. From working-time evidence, we can infer the amount of time that workers spent in non-market work annually with respect to their total time endowment. To illustrate how this calculation is made, consider the following example. In 1913, annual hours worked in the United States were 2900. Given that workers' time endowment (ignoring sleeping time) is 5840 hours (16 hours per day multiplied by 365 days), annual leisure amounts to 50 percent of their time. The second step is, similar to what I did in Equation 5, to choose a functional form for leisure. Following Jones and Klenow (2016), I use a form that implies that the percentage change in hours worked due to a percentage change in wages is constant (keeping the marginal utility of consumption fixed), and calibrate it in the same way by assuming that the elasticity of labour supply (i.e. how working hours respond to wage changes) is one with the observed annual working time observed in 1950 for the United States. [32]

As we can see in the first term of Equation 6, health differences across countries (and over time) depend on annual consumption, leisure, inequality and a constant to be calibrated ($\bar{u}$). A value for this parameter (and similar ones in other studies) has been traditionally chosen on the basis of empirical evidence from trade-offs involving risk and money. This particular trade-off exists when a certain decision involves a health risk. Since such risk is undesirable, the individual has to be compensated in some other aspect (e.g. money) to accept that choice; from this compensation we can infer the value individuals put into mortality risk

[31] Also, a value has to be chosen for β to measure preference for present and delay satisfaction. For the main results I will assume that there is no discount rate. The reason for this choice is that by assuming that $\beta = 1$, Equation 5 becomes much more intuitive in that welfare in a given country is determined by the accumulated expected value of realising the consumption and leisure levels observed in one particular year (the conclusions are the same if discounting rates are applied as the robustness tests in the Appendix show).

[32] See the Appendix for a more comprehensive explanation of this parameter as well as several robustness tests using different values.

to get estimates of the so-called value of a statistical life (VSL). A historical example of this trade-off has been examined by Williamson (1982). In this study, the author observed that British workers during the industrial revolution had to be "bribed" (or compensated in monetary terms) in order to move from small market towns with low population density and relatively low mortality rates to densely-populated industrial towns where mortality was much higher. To obtain an estimate for $\bar{u}$, I have used the same benchmark VSL ($6 million) as in Murphy and Topel (2006) and Jones and Klenow (2016), which in turn is in line with the range suggested by Viscusi and Aldy (2003). If we instead consider the range provided by Costa and Kahn (2004) using American data for the period 1940-1980, we can conclude that this figure is a conservative one. Despite there is some agreement in the literature concerning this estimate, drawing on current evidence of individuals preferences to choose a value for $\bar{u}$ in 1950 can be problematic because ample evidence in the literature points out that the value that people put into health has changed along with increasing income levels over time (Costa & Kahn, 2004).[33] To use a VSL representative of workers' preferences in 1950 I use an income elasticity of 1.3 that is in the middle of the estimate provided by Costa and Kahn (2004) and Becker and Elias (2007).[34]

4 Results

How did an average Western European compare with her American counterpart at the beginning of the 20th century? Are differences in living standards larger using income or the consumption-equivalent measure? These questions can be answered with the information provided in Table 5. In the left panel, I provide relative levels of living standards using the consumption-equivalent measure (Column I), income per capita (Column II) and consumption per capita (Column III). To understand why the new indicator revises welfare levels upwards or downwards with respect to income, I present a welfare breakdown in the right panel. Column IV shows the difference between the two measures (in logarithmic points) and the remaining columns the extent to which each well-being aspect contributes to this difference (the sum of these add up to Column IV). To make sense of the sign and magnitude of their contribution to this difference, I provide the underlying raw data for each aspect considered at the regional level: life expectancy at birth for health, consumption relative to the United States for material well-being, annual working time for (the lack of) leisure and the Gini coefficient for inequality.

[33] This is also supported by the VSLs calculated for England and the United States in the late 19th and beginning of the 20th century (Fishback, 1992; Kim & Fishback, 1993; Williamson, 1982).

[34] The analysis in Becker et al. (2005) also supports this choice (see the Appendix for testing the sensitivity of the results to changes in the VSL).

To interpret the results in Table 5, an example can be illustrating. Welfare in Northern Europe is significantly higher than measured with income per capita; actually, with the new welfare measure the gap in living standards between this region and the U.S. almost disappears. Since the consumption-equivalent indicator is higher than income, the difference between the two of them is positive (see Column IV). What drives this difference? First, taking into account that Northern Europeans were expected to live four years more increases welfare around 13 percent. Secondly, switching to consumption to measure material well-being instead of income increases welfare because relative consumption is eight percent higher than measured with income (see Columns II and III).[35] Third, leisure further widens the welfare difference between the regions considered (20 logarithmic points) because citizens in Northern Europe worked 160 hours less annually. This dimension together with health are the most important in improving the relative position of Northern Europe as they account for more than 80 percent of the upward revision (inequality barely contributes to this because the Gini coefficients are almost the same in both regions). Taken together, the new welfare measure indicates that better health and more leisure in the northern periphery largely compensate for much lower consumption so that overall welfare is the same in both regions. The remaining continental regions exhibit slightly lower levels of welfare than suggested by income due to relatively low levels of health in Southern Europe and consumption in the industrial core. Interestingly, with the welfare measure the best-performing region is not North-Western Europe (as shown by income), but the northern periphery followed by the central and southern part of the continent. At the country level, the United States is outperformed by countries such as Denmark, the United Kingdom and Switzerland.

If we look at the European continent as a whole, the welfare of an average citizen is around two thirds of the American level, exactly the same as what income per capita shows. The reasons for this is that lower levels of consumption and health are compensated for higher leisure and equality. This upward revision could have been 13 percent larger (as in Northern Europe), if life expectancy had not been four years lower in Europe. However, in this counterfactual scenario the gap with the United States would not have closed entirely because less inequality and working time would not have fully compensated for a much lower consumption level in the old continent.

[35]This figure was obtained by taking the logarithmic ratio of consumption and income (with respect to the United States) in Northern Europe: 64.83/60.20=1.08 (this ratio for the United States is zero by definition). Calculating this ratio with logarithms might be a source of confusion because we can arrive at the same figure with the ratio itself. However, for other regions (or countries) this might not be the case. In North-Western Europe, for example, the ratio of consumption to income (see Table 5, Columns II and III) is 0.88 (67.29/76.55). If we take the logarithmic ratio we arrive at the figure reported in the table: -0.13. Also, note that in Table 5 decimals for income and consumption levels are not reported and therefore calculating these ratios with these figures may give slightly different results.

Table 5: Living standards in 1913

	Well-being indicators			Welfare decomposition				
Countries	Welfare (I)	Income (II)	Consumption (III)	Difference (IV)	Health (V)	Consumption (VI)	Leisure (VII)	Inequality (VIII)
United States	100	100	100	0	0 *54*	0 *1*	0 *2900*	 *51*
Sweden	68	53	57	0.25	0.12	0.08	0.19	-0.14
Denmark	136	74	79	0.60	0.15	0.06	0.20	0.19
Northern Europe	89	60	65	0.39	0.13 *58*	0.07 *1.08*	0.19 *2740*	-0.01 *51*
United Kingdom	145	93	93	0.45	-0.02	0.00	0.29	0.17
Germany	55	68	50	-0.21	-0.15	-0.32	0.21	0.04
Netherlands	69	76	63	0.10	0.06	-0.20	-0.05	0.08
Belgium	86	81	85	0.06	-0.08	0.05	0.07	0.02
France	49	66	62	-0.29	-0.10	-0.06	-0.04	-0.09
Switzerland	149	137	94	0.09	-0.01	-0.38	0.23	0.24
North-Western Europe	76	77	67	0.00	-0.09 *51*	-0.13 *0.88*	0.16 *2769*	0.06 *48*
Spain	50	39	42	0.25	-0.43	0.06	0.35	0.27
Italy	33	43	41	-0.25	-0.18	-0.05	-0.07	0.04
Southern Europe	40	42	41	-0.05	-0.26 *45*	-0.01 *0.99*	0.09 *2829*	0.13 *44*
Western Europe	67	67	61	-0.01	-0.13 *50*	-0.10 *0.90*	0.14 *2783*	0.08 *47*

Note: see the text for the data sources. Figures in Column I and II are expressed in percentage terms with respect to the United States. The estimates in Column I and the welfare breakdown in from Column IV to VIII were obtained by taking the geometric average of the outcomes in Equations 6 and 6 (the latter in the Appendix). Below the contribution of each well-being dimension to the logarithmic difference between income and the welfare measure (Column IV), I provide data on life expectancy at birth, consumption divided by income (Column III and II), annual hours worked and Gini coefficients.

Comparing the aggregates for Western or Southern Europe (Columns I and II) one might be inclined to think that the new indicator does not add much to the information provided by GDP per capita since they show very similar results. However, drawing this conclusion on the basis of aggregated data can be misleading as country-specific differences between them can be not only very significant but also of opposite sign. For example, whereas well-being is revised upwards in Northern European countries and some states from the industrial core (e.g. the United Kingdom or Belgium), others such as Germany, the Netherlands or France are revised downwards. The reasons for these variations are also specific to each country. For instance, in the United Kingdom less annual hours worked and lower inequality more than compensate for poorer health levels than in the United States. In the Swedish case, the main

factors for its upward shift in the welfare ranking are health and leisure despite its high Gini coefficient. By taking an average for Europe, these variations of different sign cancel out and result in an average welfare level that is very close to that using income per capita.[36]

Having shed light into relative welfare levels across countries at the beginning of the analysed period with the new indicator, we can now address the question whether the first half of the 20th century in Western Europe can be characterised as one of missed or taken opportunities in terms of broader well-being. For this purpose, I present annual growth rates of welfare using the consumption-equivalent measure in Table 6 (Column I). To put these into perspective, I also provide income and consumption growth rates in Columns II and III. Moreover, in the right panel I provide a welfare breakdown to look at the reasons why welfare and income per capita growth differ (the only difference with Table 5 is that Column VI is now interpreted as the percentage growth difference between income and consumption). Does income growth underestimate welfare growth in Western Europe during the period 1913-1950? According to the consumption-equivalent measure, the answer is clearly affirmative and its magnitude is sizeable. To continue with the previous example of France, if in our measure of living standards we take into account that French citizens in 1950 not only lived in a country where the incidence of infectious diseases was nowhere near that in 1913, but also spent almost 900 hours less in the workplace, welfare growth is four times faster than what income per capita suggests. This remarkable case was not an isolated case since considering the European aggregate, living standards as measured with income per capita grew below a percentage point annually, whereas the welfare measure suggests a yearly increase of almost five percentage points. The implications of this difference are very substantial: in almost four decades income per capita only grew by 25 percent and welfare growth points to a doubling of well-being every 15 years. This difference is mainly explained by welfare gains from the decline of mortality and the expansion of leisure. The 16-year increase in life expectancy and the almost 900-hour decline in annual working time add 1.9 and 1.8 percentage points to European welfare growth respectively (Column V and VII). The decrease in inequality also contributed positively, but to a much lower extent. Across countries, growth experiences differ substantially as the case of Germany and Sweden illustrate. While in the former welfare grew at 3.6 percentage points yearly because progress in material well-being was near zero, in

[36]In Figure 4 in the Appendix, I look at this in a more comprehensive way by calculating welfare levels for several benchmark years (1913, 1920, 1929, 1938 and 1950) and plotting these against income per capita in the same years. Note that those estimates are much weaker than those used in the text (and therefore they are not reported) because data for working time and inequality are not available for all benchmarks. As one can clearly see in that figure, welfare levels are revised upwards and downwards indistinctly across the whole income spectrum. This indicates two things: the new measure conveys information that is not considered in income per capita and country-specific experiences can deviate significantly from the more aggregated regional experience.

the latter country citizens saw the four well-being aspects considered improve and contribute to a yearly growth rate of six percentage points. As a result, the underestimation of welfare growth by income differs greatly in these two countries. If we consider the whole sample, the growth differential between the consumption-equivalent indicator and income ranges from 3.2 to five percentage points annually in Belgium and Italy respectively.

These results suggest that income per capita is significantly underestimating welfare growth during this period and that it provides a one-sided view of welfare development by not taking into account changes in health and leisure (and to some extent inequality). If we extend the concept of living standards beyond its material component to include the gains from the advent of modern medicine and the rise in leisure, welfare in Western Europe by 1950 had more than doubled with respect to 1913, despite the negative effects that armed conflicts, financial crises and the like had on material well-being. According to the new welfare measure, opportunities to increase well-being in health and leisure were taken over time, and their impact on people's lives was sizeable.

Given the similarity between the welfare measures used in this paper and in Crafts (1997), a comparison between the two is interesting. In his 16-country sample, average welfare growth (i.e. a combination of income, health and leisure) is 2.4 percentage points and it ranges from 1.2 to 3.3 percentage points (in Germany and Sweden respectively). In my calculations, the growth average and range are double as high. Despite the country sample is different in both studies and my welfare measure accounts for inequality, the growth rates in Table 6 are larger because of the magnitude of the welfare gains from health and leisure. If we consider health, there are two reasons why the consumption-equivalent measure I use gives more importance to welfare gains. First, the VSL used in Crafts (1997) is lower than the one in this study. Second, Crafts makes a 25-percent downward adjustment to his calculations in order to avoid double counting because part of health improvements during this period might be due to income growth, which are already accounted for by GDP per capita. In other words, this adjustment is aimed at taking into account only those health improvements that are exogenous to income. While this issue is a relevant one in studies using any type of composite index, the lack of agreement in the literature on this matter makes such an adjustment very difficult (and to some extent arbitrary). For this reason, I do not adjust my estimates and I interpret them as an upper-bound estimate of the contribution of health to welfare. In the case of leisure, there are also two reasons why my estimates are larger. One is that the functional form used in this paper gives more value to this aspect and the other is that the underlying data for working time from Maddison (1995) exhibit a lower decrease in annual hours worked than the source used in this study (Huberman & Minns, 2007).

Moreover, Crafts used hours worked per person and not per worker to measure leisure.[37]

Table 6: Welfare across time in Western Europe and the United States, 1913-1950

	Well-being indicators			Welfare decomposition				
Countries	Welfare (I)	Income (II)	Consumption (III)	Difference (IV)	Health (V)	Consumption (VI)	Leisure (VII)	Inequality (VIII)
United States	5.8	1.7	1.3	4.2	1.4 *54-68*	-0.4 *1.7-1.3*	2.5 *2900-2008*	0.6 *51-47*
Sweden	6.0	2.4	1.9	3.6	1.2	-0.5	2.0	0.9
Denmark	4.6	1.5	1.3	3.2	1.3	-0.2	1.8	0.2
Northern Europe	5.5	2.0	1.7	3.4	1.2 *58-71*	-0.4 *2.0-1.7*	1.9 *2740-2032*	0.6 *51-39*
United Kingdom	4.4	1.0	0.8	3.4	1.8	-0.3	1.5	0.4
Germany	3.6	0.2	0.1	3.4	2.0	-0.1	1.0	0.4
Netherlands	5.1	1.1	0.8	4.1	1.5	-0.3	2.3	0.5
Belgium	3.8	0.7	0.4	3.2	1.6	-0.3	1.3	0.6
France	4.5	1.1	0.6	3.4	1.6	-0.5	2.5	-0.2
Switzerland	4.3	0.7	1.1	3.6	1.7	0.4	1.7	-0.1
North-Western Europe	4.3	0.8	0.6	3.6	1.8 *51-67*	-0.2 *0.8-0.6*	1.6 *2769-2209*	0.3 *48-41*
Spain	4.7	0.3	0.2	4.3	3.0	-0.1	1.5	0.0
Italy	6.0	1.0	0.8	5.0	2.0	-0.2	2.8	0.3
Southern Europe	5.5	0.7	0.6	4.7	2.3 *45-64*	-0.2 *0.7-0.6*	2.3 *2829-1989*	0.2 *44-40*
Western Europe	4.6	0.8	0.6	3.9	1.9 *50-66*	-0.2 0.8-0.6	1.8 *2783-2140*	0.3 *47-41*

Note: see Table 5. The regional raw data in Column VI refers to income and consumption growth (also in Columns II and III)

Welfare convergence

Large welfare gains derived from lower mortality and working time indicate that European states took a number of crucial opportunities during the period 1913-1950. However, the extent to which these opportunities were taken across space and how they affected relative welfare levels throughout this period has not been considered yet.

Beginning with a comparison between the United States and the European aggregate, we can see in Table 6 that the old continent exhibits slower welfare growth (1.2 percentage points). This growth differential that might appear small became very significant over a period of almost four decades. Actually, as Table 7 shows, Europe's relative position fell more than 40 percent by 1950. This process was the result of lower consumption growth

[37]In the Appendix, I test these issues and show that the main conclusions remain unaltered.

and a bigger rise in leisure and inequality at the other side of the Atlantic. With respect to mortality, progress was faster in Europe due to its lower starting level, although this could not compensate for a worse performance in the rest of the dimensions considered. This divergence did not take place throughout the period, but only after 1929. Up to the Great Depression, larger welfare gains from declining mortality and inequality resulted in slightly faster welfare growth in Europe (Northern Europe even surpassed the United States in 1929). However, this trend is dramatically reversed in the 1930s and 1940s as Europe's relative welfare level declines by almost a half after the Second World War mainly due to the negative impact of the war on consumption per capita.

Despite this pattern is shared by all subregions in the sample, some regions experienced a more dramatic relative decline than others as the industrial core illustrates. In the period 1929-1950 its relative welfare level fell from 81 to 41 of the American level due to a slowdown in material progress, leisure and inequality. The peripheries lost less ground with respect to the United States due to substantial reductions in infant and child mortality rates in the south and good economic performance in the north of the continent.

Table 7: Welfare levels in Western Europe and the United States, 1913-1950

	1913 (I)	1929 (II)	1950 (III)
United States	100	100	100
Western Europe	67	71	39
Subregions			
Northern Europe	89	105	84
North-Western Europe	76	81	41
Southern Europe	40	45	28
CoV (welfare)	0.48	0.39	0.52
CoV (HDI)	0.12	n.d.	0.05

Note: own calculations (see Table 5). The calculations of the coefficient of variation (CoV) are based on the HDI values reported in Crafts (2002).

What can we conclude from all these different developments? If we take Europe and the United States as the units of analysis, we clearly observe no convergence. But, what if we look at country-level experiences? Was the distribution of welfare more equal after the analysed period. Answering these questions would shed light on whether historical achievements in medicine, sanitation and workers rights created a more equal Europe in terms of welfare as the HDI-based evidence suggests. Unfortunately, a direct comparison between the HDI

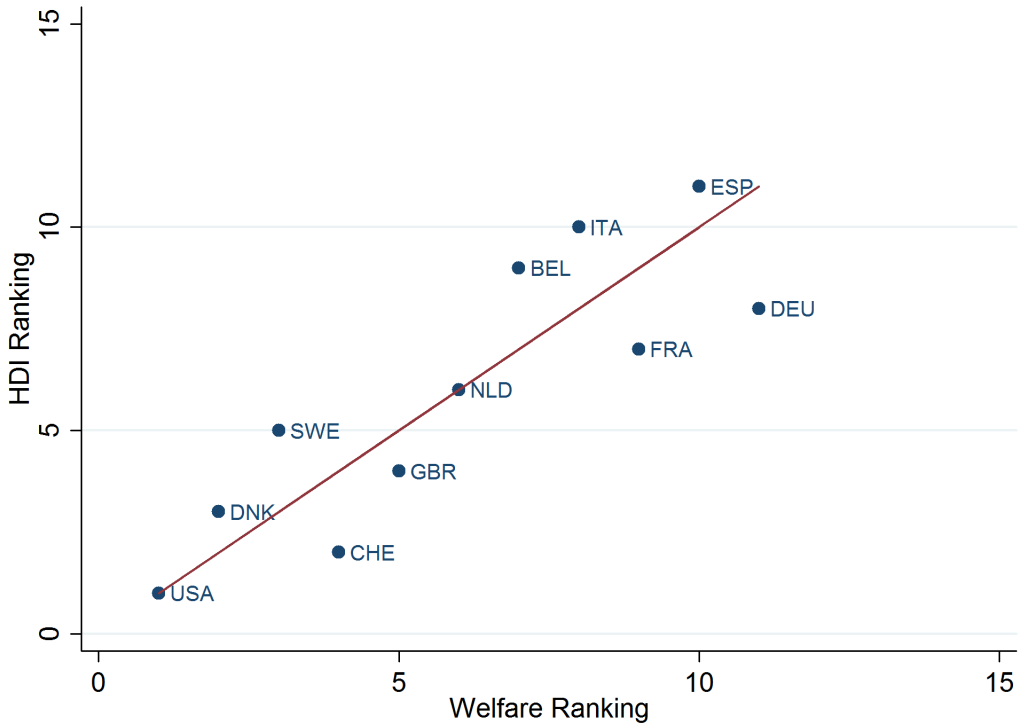
and the consumption-equivalent measure (i.e. in terms of absolute levels) is problematic to tackle the previous questions since they are measured in different units. However, we can explore the implications of each measure with regards to convergence in living standards by analysing the extent to which countries welfare differ using either of them. For this purpose, we can use the coefficient of variation and measure the degree to which welfare differs across countries (relative to the mean of a given indicator) and how this dispersion evolved during the period 1913-1950. In the last two rows of Table 7 I report the calculations of this spread measure using the consumption-equivalent measure and the HDI estimates of Crafts (2002). Considering the evolution of living standards spread using the HDI, we can see that by mid-century cross-country differences had greatly narrowed as a result of countries with low levels of human development in 1913 experiencing faster growth than those with higher levels. In line with the idea that this indicator tends to show convergence (Prados de la Escosura, 2015), welfare spread declined by more than a half. If we consider the results for the consumption-equivalent measure, two points stand out. First, the coefficient of variations for the welfare measure in 1913 and 1950 are four and ten times larger than those of the HDI. Second, the strong convergence in living standards observed using the HDI is not supported by the new indicator since the coefficient of variation in 1950 is almost the same as in 1913. This implies that differences across countries were not only much larger than the HDI shows, but they persisted throughout the analysed period.

The different implications of the two measures for our understanding of the period are somewhat discouraging for the study of cross-sectional welfare in history because indicators that aim at measuring similar things (i.e. human welfare) show very different patterns. However, it is worth highlighting that the previous analysis was specifically concerned with relative levels and in other studies the HDI is used with a different purpose, namely to rank countries. Therefore, an interesting cross-check for the two measures would be to study how its ranking performance compare since if both convey information on a number of crucial aspects of human life, they should show similar results.

In Figure 1 I present the outcome of ranking countries in 1950 (one is the highest and 11 the lowest). If there was a perfect correlation between the two of them, we would expect all points to lie along a 45-degree line because the position of a certain country in terms of the consumption-equivalent metric (horizontal axis) would be the same as with the HDI (vertical axis). It is very remarkable that despite the large differences in terms of data, methodology and dimensions considered in these indices, the correlation between the two is quite high. Both measures recognise that welfare is highest in the United States and Denmark; and lowest in Southern Europe and some parts of the industrial core (i.e. Germany and France). We can test this correlation with a larger sample by calculating further welfare and HDI

levels for more benchmark years. In a simple Ordinary Least Squares framework, a perfect correlation between the two (i.e. 45 degree line in Figure 1) would yield a coefficient and R-squared of one and 100 percent respectively.[38] The results for this sample yield a coefficient of 0.78 and R-squared of 0.60 which indicate that the two of them are highly correlated and convey very similar information (see Table 2 in the Appendix). These results are very encouraging in that they do not only support the usefulness of both measures for ranking countries (without taking into account the preference of the researcher for one or the other), but they also reinforce the use of the new welfare indicator for analysing historical welfare.

Figure 1: Country ranking using the HDI and the consumption-equivalent measure in 1950



Note: own calculations. Countries are ranked in an ascending order by their level of welfare.

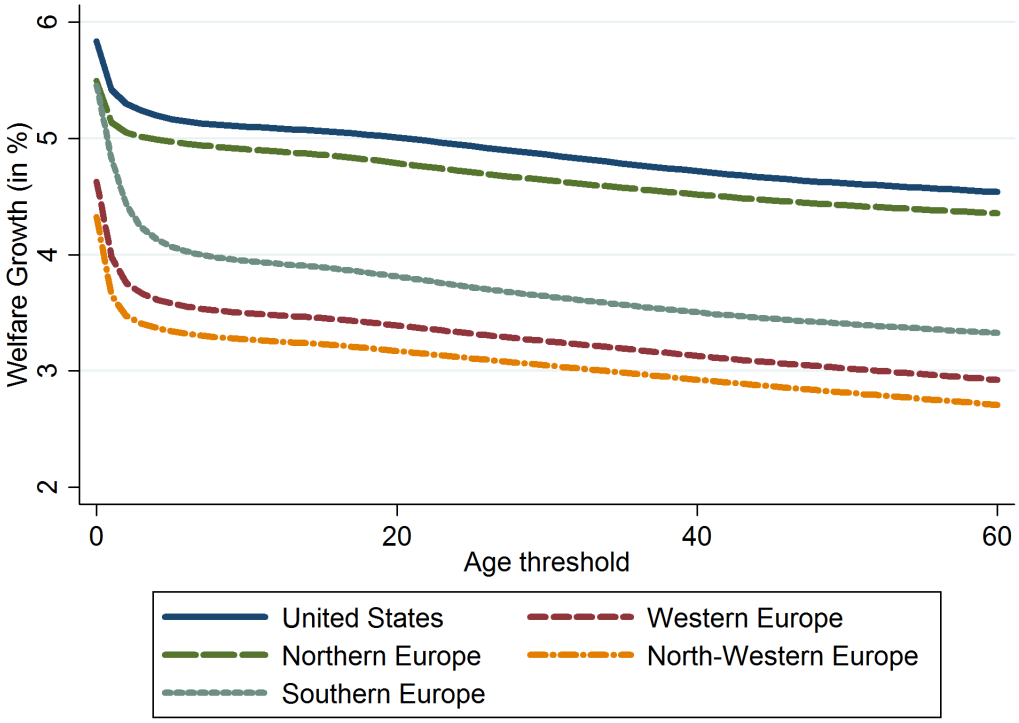
Age-specific mortality rates

As highlighted before, the decline in mortality did not happen in the same way across the age distribution. Instead, progress against diseases such as pneumonia or gastroenteritis

[38]Note that data availability for the additional benchmarks are much more limited and several assumptions had to be made to perform the calculations using the consumption-equivalent indicator (see footnote 36). However, given that these assumptions make the estimations of my welfare measure less realistic than those of the HDI for which data are complete, finding a strong correlation becomes more difficult.

disproportionately increased the survival probabilities of children between their birth and fifth birthday. To look at this in more detail, I will quantify the effect of health improvements in different parts of the age distribution and test the robustness of the patterns described until now. For this purpose, I can take advantage of the richness of the age-specific mortality data to perform the same exercise as in Table 6 but considering the counterfactual that mortality improvements only happened after certain age thresholds. The results of this exercise are presented in Figure 2 where I provide welfare growth rates for the period 1913-1950 when considering mortality changes from birth until age 60 for the United States, Western Europe and its subregions. Unsurprisingly, welfare growth falls as the age threshold increases because mortality declined less at older ages. The drop in welfare growth is specially the case in the age range 0-5. In fact, according to these estimates, not taking into account progress in child mortality lowers welfare growth between 0.5 and 1.4 percentage points annually in Northern and Southern Europe respectively. While these growth differentials are important, they do not override the idea that the period 1913-1950 is one of strong welfare growth for Europeans and Americans. Moreover, the figures at the regional level broadly support the view that the U.S. was performing better than any region in Western Europe and that within the continent citizens in the industrial core witnessed a lower improvement than their counterparts from the peripheries. The only difference with the previous results is that Southern Europe is not on a par with the northernmost region if we take the perspective of someone older than one.

Figure 2: Welfare growth across the age distribution, 1913-1950



The counterfactual exercise of the growth analysis can be also performed in a cross-sectional setting by assuming idential mortality rates across countries below a given age threshold. Given that Columns III to VI in Table 3 indicate that cross-country health patterns barely change after age five, in Table 8 I report levels in 1913, 1929 and 1950 taking the perspective of a five-year old. The results reported in this table broadly confirm the evolution of relative welfare exposed in the previous section. First, we can distinguish a first period between 1913 and 1929 of catching up (and forging ahead in the case of Northern Europe) followed by years of relative decline in the 1930s and 1940s. The relative decline of Western Europe is very similar taking the perspective of a newborn or that of a five-year-old child (42 and 46 percent respectively). The main difference between the results in Tables 8 and 7 is that relative welfare levels are more compressed because mortality levels are more homogeneous after the age of five. This affects our view of some regions in different ways. For example, ignoring child mortality rates reduces relative welfare in Northern Europe and it increases that of Southern Europe together and the industrial core (specially in 1913 when mortality at these ages was higher). Comparing the implications for the growth rate and level analyses, it is interesting to note that taking one perspective or the other does not benefit regions (or countries) in the same way. For example, the most penalised regions in terms of growth rates by ignoring health improvements below the age of five were North-Western

and Southern Europe, while these two are the most benefited in the level analyses as their welfare gap with respect to the U.S. decreases substantially.

If we look at welfare dispersion, the coefficient of variations in the last row of Table 8 supports the previous findings that welfare spread is much higher using the consumption-equivalent indicator than using the HDI. Moreover, and most importantly, considering the counterfactual that cross-country mortality did not differ below the age of five does not support the idea of convergence in welfare that the HDI calculations imply. Instead, a slight rise in welfare dispersion across countries rejects the egalitarian view of welfare and implies that cross-country differences were large and persistent throughout the analysed period.

Table 8: Welfare levels in Western Europe and the United States taking the perspective of a five-year-old, 1913-1950

	1913 (I)	1929 (II)	1950 (III)
USA	100	100	100
Western Europe	76	80	41
Subregions			
Northern Europe	84	105	82
North-Western Europe	85	87	43
Southern Europe	50	59	31
CoV (welfare)	0.44	0.33	0.48

Note: own calculations (see Table 5).

5 Conclusions

The lenses through which we look at society greatly influence our perception and narratives of the past. In the study of historical living standards, this is particularly relevant and challenging when a focus on different indicators present wildly different paths of human development. The study of European living standards during the period 1913-1950 is a perfect example of this: while historically-low rates of economic growth would characterise this period as one of missed opportunities for improving people's welfare, unprecedented achievements in levels of health and leisure due to the spread of modern medicine, the implementation of large-scale sanitation infrastructures or the introduction of the 8-hour working day indicate the opposite .

To examine whether the idea of missed opportunities for living standards beyond pure

material component is supported by a larger basis of measures, I have applied a new composite indicator of welfare that integrates material well-being, health, leisure time and inequality. Contrary to previously-used composite indicators that look at this period, this measure is more comprehensive in the aspects of well-being considered and, more importantly, it is grounded in economic theory which allows for welfare calculations across countries and time that are directly comparable to GDP. The main findings show that income per capita underestimates welfare growth significantly -up to five percentage points yearly- and that therefore many of the opportunities that emerged during the first half of the 20th century to improve people's well-being were not missed. However, these opportunities were not taken (or did not arise) equally across countries since, as opposed to what other measures of human welfare such as the HDI imply, differences across countries were large and more persistent throughout this period.

References

Aitchison, J., & Brown, J. (1957). *The lognormal distribution with special reference to its uses in economics*. Cambridge University Press: London.

Atkinson, A. B. (2007). The long rung earnings distribution in five countries: "remarkable stability," u, v, or w? *Review of Income and Wealth*, *53*(1), 1-24.

Barro, R. J., & Ursua, J. F. (2008). Macroeconomic crises since 1870. *Brooking Papers on Economic Activity*, *Spring*, 255-350.

Becker, G. S., & Elias, J. J. (2007). Introducing incentives in the market for live and cadaveric organ donations. *Journal of Economic Perspectives*, *21*(3), 3-24.

Becker, G. S., Philipson, T. J., & Soares, R. R. (2005). The quantity and quality of life and the evolution of world inequality. *The American Economic Review*, *95*(1), 277-291.

Beckerman, W. (1980). Comparative growth rates of 'Measurable Economic Welfare': some experimental calculations. In R. Matthews (Ed.), *Economic growth and resources.* London: Macmillan.

Bell, F., & Millward, R. (1998). Public health expenditures and mortality in England and Wales, 1870-1914. *Continuity and Change*, *13*(2), 221-249.

Bolt, J., & van Zanden, J. L. (2014). The Maddison project: collaborative research on historical national accounts. *The Economic History Review*, *67*(3), 627-651.

Buyst, E., & Franaszek, P. (2010). Sectoral developments, 1914-1945. In S. Broadberry & K. H. O'Rourke (Eds.), *The Cambridge economic history of modern Europe. Volume 2: 1870 to the present.* Cambridge: Cambridge University Press.

Costa, D. L. (2015). Health and the economy in the United States from 1750 to the present. *Journal of Economic Literature*, *53*(3), 503-570.

Costa, D. L., & Kahn, M. E. (2004). Changes in the value of life, 1940-1980. *The Journal of Risk and Uncertainty*, *29*(2), 159-180.

Costa, D. L., & Steckel, R. H. (1997). Long-term trends in health, welfare, and economic growth in the United States. In R. H. Steckel & R. Floud (Eds.), *Health and welfare during industrialization* (p. 47-89). Chicago: University of Chicago Press for National Bureau of Economic Research.

Crafts, N. (1997). The Human Development Index and changes in standards of living: Some historical comparisons. *European Review of Economic History*, *1*(3), 299-322.

Crafts, N. (2002). The Human Development Index, 1870-1999: Some revised estimates. *European Review of Economic History*, *6*(3), 395-405.

Crafts, N. (2007). Living standards. In I. G. N. Crafts & A. Newell (Eds.), *Work and pay in twentieth century britain* (p. 11-34). Oxford: Oxford University Press.

Cross, G. (1989). *A quest for time. the reduction of work in Britain and France, 1840-1940.* Berkeley: University of California Press.

Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). The determinants of mortality. *Journal of Economic Perspectives*, *20*(3), 97-120.

Cutler, D., & Miller, G. (2005). The role of public health improvements in health advances: the twentieth-century United States. *Demography*, *42*(1), 1-22.

Evans, A. A. (1969). Work and leisure, 1919-1969. *International Labour Review*, *99*(1), 35-59.

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the Penn World Table. *American Economic Review*, *105*(10), 3150-3182.

Fernández-Villaverde, J., & Krueger, D. (2007). Consumption over the life cycle: Facts from consumer expenditure survey data. *The Review of Economics and Statistics*, *89*(3), 552-565.

Ferrie, J. P., & Troesken, W. (2008). Water and Chicago's mortality transition, 1850-1925. *Explorations in Economic History*, *45*(1), 1-16.

Fishback, P. V. (1992). *Soft coal, hard choices: The economic welfare of bituminous coal miners, 1890-1930.* New York-Oxford: Oxford University Press.

Fogel, R. W. (2004). *The escape from hunger and premature death, 1700-2100. Europe, America, and the third world.* Cambridge: Cambridge University Press.

Goldin, C., & Margo, R. A. (1992). The great compression: The wage structure in the United States at mid-century. *Quarterly Journal of Economics*, *107*(1), 1-34.

Hickson, K. J. (2014). The GDP value of twentieth-century health improvements in developed economies: initial estimates for England. *The Review of Income and Wealth*, *60*(2), 385-399.

Hjerppe, R., & Jalava, J. (2006). Economic growth and structural change - a century and a half of catching-up. In J. E. J. Ojala & J. Jalava (Eds.), *The road to prosperity: An economic history of Finland.* Helsinki: SKS.

Huberman, M., & Minns, C. (2007). The times they are not changin': Days and hours of work in old and new worlds, 1870-2000. *Explorations in Economic History*, *44*(4), 538-567.

Human Life-Table Database. (2013). http://www.lifetable.de. accessed november 2013.

Human Mortality Database. (n.d.). University of california, berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Retrieved June 9 2015, from www.mortality.org.

Jayachandran, S., Lleras-Muney, A., & Smith, K. V. (2010). Modern medicine and the twentieth century decline in mortality: Evidence on the impact of sulfa drugs. *American*

Economic Journal: Applied Economics, *2*(2), 118-146.

Jones, C. I., & Klenow, P. J. (2016). Beyond GDP? Welfare across Countries and Time. *American Economic Review*, *106*(9), 2426-2457.

Kakwani, N. C. (1993). Performance in living standards. an international comparison. *Journal of Development Economics*, *41*(2), 307-336.

Kim, S.-W., & Fishback, P. V. (1993). Institutional change, compensating differentials, and accident risk in American railroading, 1892-1945. *Journal of Economic History*, *53*(4), 796-823.

Krantz, O. (1987). Schweden, Norwegen, Danemark, Finnland, 1914-1970. In H. K. I. M. W. Fischer J. A. van Houtte & F. Vittinghoff (Eds.), *Handbuch der Europäischen Wirtschafts- und Sozialgeschichte* (Vol. 6). Stuttgart: Klett-Cotta.

Lakner, C., & Milanovic, B. (2015). Global income distribution: from the fall of the Berlin wall to the great recession. *World Bank Economic Review*, *30*(2), 203-232.

Leonard, C., & Ljungberg, J. (2010). Population and living standards, 1870-1914. In S. Broadberry & K. H. O'Rourke (Eds.), *The Cambridge economic history of modern Europe. Volume 2: 1870 to the present.* Cambridge: Cambridge University Press.

Maddison, A. (1995). *Monitoring the world economy, 1820-1992.* Paris: OECD.

Maddison, A. (2006). *The world economy: A millennial perspective (vol. 1). historical statistics (vol. 2).* Paris: Development Center of the OECD.

Millward, R., & Baten, J. (2010). Population and living standards, 1914-1945. In S. Broadberry & K. H. O'Rourke (Eds.), *The cambridge economic history of modern europe. volume 2: 1870 to the present.* Cambridge: Cambridge University Press.

Mitchell, B. R. (2003). *International historical statistics: Europe, 1750-2000.* London: Macmillan.

Mokyr, J., & Stein, R. (1996). Science, health, and household technology: The effect of the pasteur revolution on consumer demand. In T. F. Bresnahan & R. J. Gordon (Eds.), *The economics of new goods* (p. 143-205). Chicago: University of Chicago Press.

Murphy, K. M., & Topel, R. H. (2006). The value of health and longevity. *Journal of Political Economy*, *114*(5), 871-904.

Nordhaus, W. D. (2003). The health of nations: The contribution of improved health to living standards. In K. M. Murphy & R. H. Topel (Eds.), *Measuring the gains from medical research: An economic approach* (p. 9-40). Chicago: University of Chicago Press.

Prados de la Escosura, L. (2015). World human development: 1870-2007. *Review of Income and Wealth*, *61*(2), 220-247.

Ravallion, M. (2012). Troubling tradeoffs in the Human Development Index. *Journal of*

Development Economics, *99*(2), 201-209.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Rijpma, A. (2014). A composite view of well-being since 1820. In J. L. van Zanden et al. (Ed.), *How was life? global well-being since 1820* (p. 249-269). OECD Publishing.

Riley, J. C. (2005). Estimates of regional and global life expectancy, 1800-2001. *Population and Development Review*, *31*(3), 537-543.

Ritschl, A. (2002). *Deutschlands Krise und Konjunktur, 1924-1934. Binnenkonjunktur, Auslandsverschuldung und Reparationsproblem zwischen Dawes-Plan und Transfersperre*. Berlin: Akademie-Verlag.

Roses, J. R., & Wolf, N. (2010). Aggregate growth, 1913-1950. In S. Broadberry & K. H. O'Rourke (Eds.), *The Cambridge economic history of modern Europe. Volume 2: 1870 to the present.* Cambridge: Cambridge University Press.

Svennilson, I. (1954). *Growth and stagnation in the european economy*. Geneva: United Nations.

UNDP. (1990). *Human development report 1990*. New York: Oxford University Press.

UNDP. (2010). *Human development report 2010. the real wealth of nations: Pathways to human development*. New York: Oxford University Press.

Usher, D. (1973). An imputation to the measure of economic growth for changes in life expectancy. In D. Moss (Ed.), *The measurement of economic and social performance* (p. 193-223). New York: National Bureau of Economic Research.

Usher, D. (1980). *The measurement of economic growth*. New York: Columbia University Press.

van Zanden, J. L., Baten, J., Foldvari, P., & van Leeuwen, B. (2013). The changing shape of global inequality 1820-2000; exploring a new dataset. *The Review of Income and Wealth*, *60*(2), 279-297.

Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, *27*(1), 5-76.

WIID. (2015). UNU-WIDER, 'world income inequality database (wiid3.3)', september 2015.

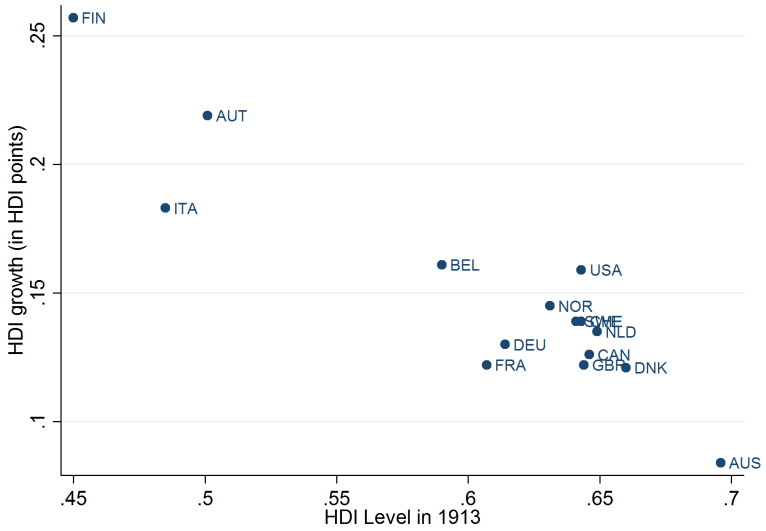
Williamson, J. G. (1982). Was the industrial revolution worth it? disamenities and death in 19th century British towns. *Explorations in Economic History*, *19*(3), 221-245.

Williamson, J. G. (1984). British mortality and the value of life. *Population Studies*, *38*(1), 157-172.

6 Appendix 1

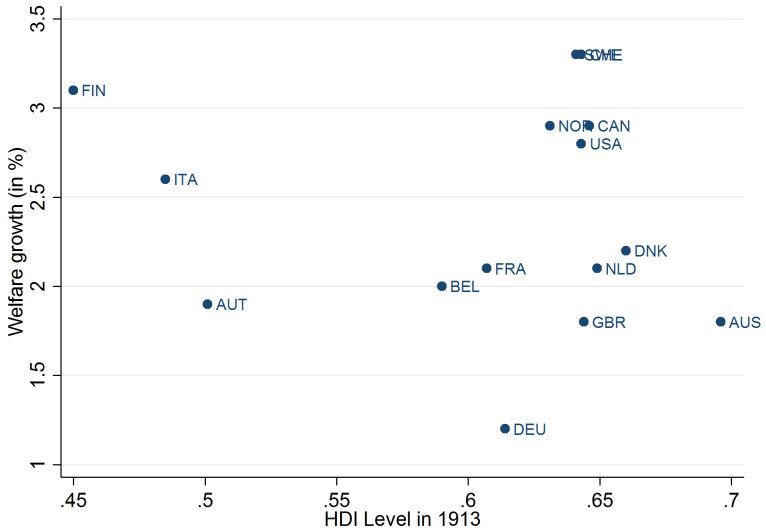
6.1 Support Material

Figure 1: Initial levels of human development and HDI growth, 1913-1950



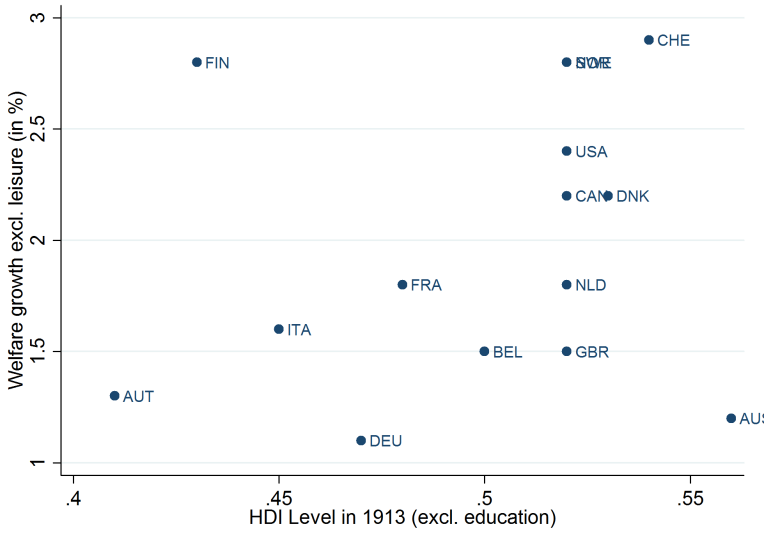
Source: Crafts (2002).

Figure 2: Initial levels of human development and welfare growth, 1913-1950



Source: Crafts (2002) for the HDI and Crafts (1997) for the utility-based measure.

Figure 3: Initial levels of human development (excl. education) and welfare growth (excl. leisure), 1913-1950

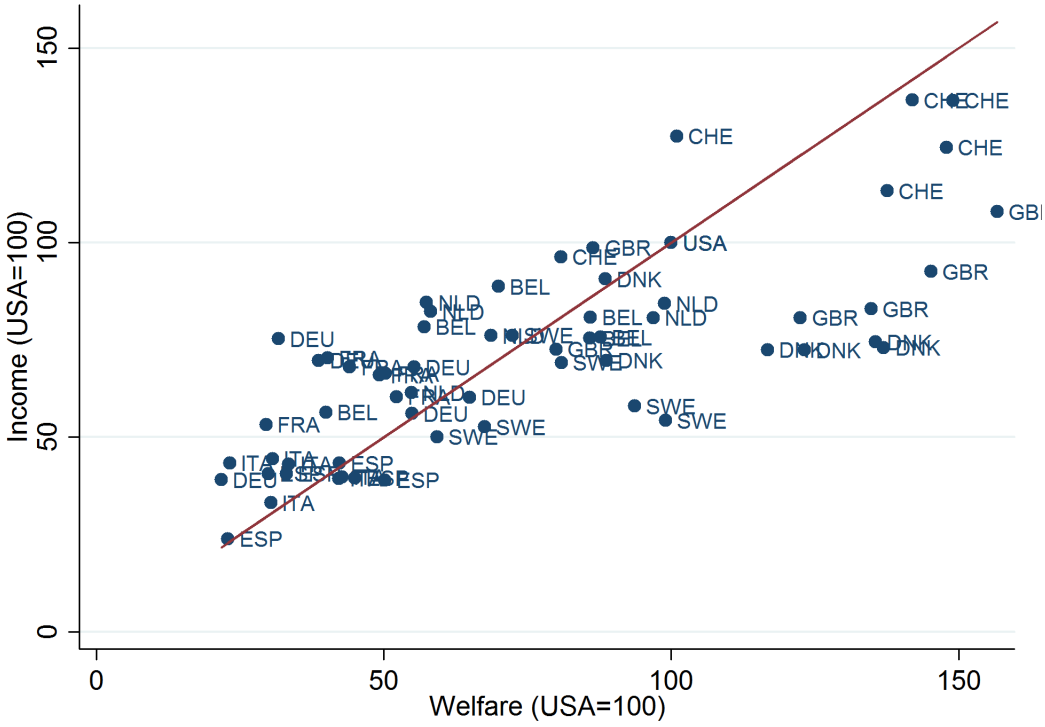


Note: I calculated HDI levels keeping education constant in 1913 (at the U.S. level) from Crafts (2002) and took the growth rates of the utility measure in Crafts (1997) excluding leisure.

Table 1: Absolute difference between the consumption indices reported in Barro and Ursua (2008) and PWT 9.0

	1950-59	1960-69	1970-79	1980-89	1990-1999	2000-06
United States	4	4	3	2	2	0
Sweden	4	4	2	2	2	1
Denmark	13	13	10	4	1	2
NE	7	7	5	2	2	1
United Kingdom	1	0	0	1	2	1
Germany	1	1	1	0	1	2
The Netherlands	2	3	6	3	4	3
Belgium	0	1	3	4	2	1
France	0	0	1	1	3	1
Switzerland	1	1	0	1	2	1
NWE	1	1	1	1	2	1
Spain	0	1	4	1	1	0
Italy	3	1	2	2	1	0
SE	2	1	3	2	1	0
WE	1	1	2	1	2	1

Figure 4: Welfare and income levels, 1913-1950



Note: the source for income is Bolt and van Zanden (2014). For welfare, I performed country-specific calculations for 1913, 1920, 1929, 1938 and 1950. Given the lack of data on annual working time for 1920, I assumed annual hours of work were the same as in 1929. Regarding Gini coefficients, levels for 1920 and 1938 were obtained by interpolating between benchmarks.

Table 2: Correlation between countries' ranking using the HDI and the consumption-equivalent measure

	(1)
	rank_HDI
rank_Welfare	0.780
	(9.07)
Constant	1.320
	(2.26)
Observations	55
Adjusted R^2	0.601

t statistics in parentheses

Note: the HDI and welfare rankings were obtained by calculating these with the data presented in the text. Data for education was obtained from Clio-Infra (2014).

6.2 Methodological extension

The aim of this section is to complement the explanation of the methodology sketched in the text with two extensions. First, I will provide a more formal exposition of the differences between the compensating and the equivalent variation, the estimation of welfare growth rates and how the consumption-equivalent measure is related to income per capita. Second, I will discuss the calibration of the model in more detail.

In the framework developed by Jones and Klenow (2016), the utility of an individual behind the veil of ignorance is the expected value of flows derived from consumption and leisure:

$$U = \mathbb{E} \sum_{a=1}^{100} \beta^a u(C_a l_a) S(a) \tag{1}$$

where S are survival rates up to age a, β is the discount factor, C is consumption and l is leisure. To perform welfare calculations, we define $U_i(\lambda)$ as expected lifetime utility in a certain country if consumption is multiplied by λ across ages:

$$U_i(\lambda) = \mathbb{E}_i \sum_{a=1}^{100} \beta^a u(\lambda C_{ai}, l_{ai}) S_i(a) \tag{2}$$

where i indexes countries. After choosing the United States as a reference country, we can perform the calculation of welfare across countries in two ways. The first is the one presented in the text (i.e. equivalent variation) and it answers the question: by what factor (λ^{ev}) must an individual's consumption be adjusted in the United States to make him indifferent between living there and in country i? The second is the compensating variation and it calculates the factor (λ^{cv}) by which an individual's consumption in country i would have to be adjusted so that she is as well off as in the United States. In terms of the questions addressed by these variations, their difference in formal terms is:

$$U_{us}(\lambda_i^{ev}) = U_i(1) \tag{3}$$

$$U_{us}(1) = U_i(\lambda_i^{cv}) \tag{4}$$

To implement the welfare calculation using the previous equations, Jones and Klenow (2016) proposes the following form for lifetime utility:

$$U_i = [\sum_{a=1}^{100} \beta^a S_i(a)] \cdot (\bar{u} + log\ c_i + v(l) - \frac{1}{2} \cdot \sigma_i^2) + g \cdot \sum_{a=1}^{100} \beta^a S_i(a) \tag{5}$$

where $\bar{u}$ is a constant, σ is the variance of log-transformed consumption and g is consumption growth. Remember that the notion of consumption inequality has been introduced by assuming that consumption in country i follows a logarithmic distribution and it is independent of age and mortality with mean c_i and variance σ_i^2. If this applies, then expected consumption is: $\mathbb{E}[log\ C] = log\ c - \sigma^2/2$. In other words, behind the veil of ignorance the individual is inequality-averse and therefore rising inequality lowers his expected consumption. Contrary to the form used in the text, Equation 5 does not imply that consumption remains constant over the life cycle (the sensibility of the results to this are tested in the next section).

Once we calculate U_i for every country, we can implement the calculation in Equations 3 and 4. If we assume that $g = 0$ (for the sake of easing notation and making this section comparable to the text), the compensating variation can be obtained as follows:

$$\begin{aligned} log\ \lambda_i^{cv} = \sum_a \frac{\beta^a[S_i(a) - S_{us}(a)]}{\sum_a \beta^a S_i(a)} [\bar{u} + log\ c_{us} + v(l_{us}) - \frac{1}{2} \cdot \sigma_{us}^2] \\ + log\ c_i - log\ c_{us} \\ + v(l_i) - v(l_{us}) \\ - \frac{1}{2}(\sigma_i^2 - \sigma_{us}^2) \end{aligned} \tag{6}$$

This way of calculating welfare differs with with Equation 6 in the text in one important aspect: the first term is multiplied by the utility flow in country i instead of the United States.[39] As a result, the equivalent variation gives lower weight to differences in survival rates for low-income countries since these are usually weighted by lower annual utility flows. In the case of the compensating variation, the opposite is true because health differences are weighted by typically-higher American utility levels. To avoid arbitrarily choosing either of them I take the geometric average between the two.[40]

Why do the consumption-equivalent measure and income per capita differ? This methodology provides a clear answer to this question since both are measured in the same units. Formally, we can compare them as follows:

[39] Also, this term is divided by the cumulative discount rate of the survival rates in country i instead of the United States.

[40] In any case, the effects on the main results are minor when using the different variations. For example, the average level difference between the two variations for the whole sample is only two percent. If we consider Western Europe in 1913, I obtain a level of 68 and 66 percent relative to the United States using the equivalent and compensating variation respectively.

$$
\begin{aligned}
log\ \frac{\lambda_i^{cv}}{\bar{y}_i} = \sum_a \frac{\beta^a[S_i(a) - S_{us}(a)]}{\sum_a \beta^a S_i(a)} [\bar{u} + log\ c_{us} + v(l_{us}) - \frac{1}{2}(\sigma_{us}^2)] \\
+ log\ c_i/y_i - log\ c_{us}/y_{us} \\
+ v(l_i) - v(l_{us}) \\
- \frac{1}{2}(\sigma_i^2 - \sigma_{us}^2)
\end{aligned} \tag{7}
$$

where $\bar{y}_i$ is the ratio of income per capita in country i to the United States. With this transformation, the logarithmic ratio of welfare differences measured with the consumption-equivalent measure and income per capita can be decomposed into four elements: differences due to consumption, health, leisure and inequality.[41]

To make comparisons over time, we can repeat the previous exercise but instead of looking at two countries, we consider a country in two different points in time. For example, if we want to calculate French welfare growth during the period 1913-1950, we just have to replace the United States and country i in Equation 6 France in 1950 and 1913 respectively. In this way, we would calculate the factor by which the annual consumption of a French citizen in 1913 would have to be adjusted so that she is as well off as her counterpart in 1950. Note that, as for cross-country comparisons, the the same issue between the equivalent and the compensating variation arises. In this case, whereas the former weights health differences by utility flows in 1913, the latter does it by taking utility flows in 1950.[42] Once both variations have been calculated, I obtain the welfare growth rate with the following formula:

$$
\Delta_i^{\lambda} \equiv -\frac{1}{t_1 - t_0}\ log\ \lambda_i \tag{8}
$$

where t_1 and t_0 are the end and starting year of the analysed time period and $log\ \lambda_i$ is $\sqrt{log\ \lambda_i^{ev} \cdot log\ \lambda_i^{cv}}$. Also, we can directly compare countries performance using the consumption-equivalent measure and income per capita terms as before:

[41] If we look at the consumption component, this formulation implies that part of the difference between income per capita and the consumption-equivalent measure is due to differences in consumption over GDP. In this study, I interpret this term in a different way as in (Jones & Klenow, 2016) by expressing both income and consumption relative to the United States. In this way, $log\frac{c_{us}}{y_{us}}$ is zero by definition and $log\frac{c_i}{y_i}$ becomes the percentage difference between relative material well-being as measured with consumption and income. Note that the only effect this transformation has on the results is of qualitative nature in interpreting them since quantitatively the welfare calculations do not differ at all.

[42] In this case, the differences between the two variations are not very significant either (although a bit larger than in the cross-sectional setting). The average welfare growth difference in the sample between the two is 1.3 percentage points; if we consider welfare growth in Western Europe for the period 1913-1950 using the equivalent and compensating variation gives four and 5.3 percent respectively.

$$
\begin{aligned}
log\ \frac{\Delta_i^{\lambda}}{\Delta_i^{y}} = \sum_a \frac{\beta^a[S_{t0}(a) - S_{t1}(a)]}{\sum_a \beta^a S_{t0}(a)}[\bar{u} + log\ c_{t1} + v(l_{t1}) - \frac{1}{2}(\sigma_{t1}^2)] \\
+ log\ c_{i,t0}/y_{i,t0} - log\ c_{i,t1}/y_{i,t1} \\
+ v(l_{i,t0}) - v(l_{i,t1}) \\
- \frac{1}{2}(\sigma_{i,t0}^2 - \sigma_{i,t1}^2)
\end{aligned} \tag{9}
$$

where Δ_i^y is the percentage annual increase in income per capita during the analysed period. As before, with this formula we can look at the factors accounting for the differences in well-being growth as calculated with the welfare indicator and income per capita (i.e. mortality, differences in consumption and income growth, leisure and inequality).

Model calibration

Before calculating welfare across countries and time using Equation 6 and Equation 6 (in the text), we need to choose a functional form to calculate the value individuals put in leisure and calibrate the model according to the preferences of American individuals. Using a utility function separable in consumption and hours worked, a form is used that, keeping the marginal utility of consumption constant, implies a constant Frisch elasticity of labour supply, namely $v(l) = -\frac{\theta\epsilon}{1+\epsilon}(1-l)^{\frac{1+\epsilon}{\epsilon}}$. This elasticity ($\epsilon$) measures how hours worked respond to wage changes abstracting from consumption. In this paper, I follow Jones and Klenow (2016) that, after surveying the literature, chooses a unitary value. Given that the decision on supplying labour depends on the wage of individuals, the calibration of θ (i.e. disutility from working) is carried out by solving $\theta = w(1-t)(1-l)^{\frac{-1}{\epsilon}}/c$ where t is the marginal tax rate. For the marginal tax rate, I used a value of 0.2 that was obtained by using information on marginal tax rates in 1950 from the database of individual income tax rates from the Tax Foundation.[43] Given that leisure in 1950 in the United Sates is 0.6562 and the w/c ratio is 1.56 (I obtained this ratio using income and consumption data for 1950 from PWT), θ in my setting is 3.63.

The next parameter to be calibrated is $\bar{u}$. As for leisure, I do not take the value reported by Jones and Klenow (2016) since they calibrate their model with data referring to the 2000s. To assign a value to this parameter, I use evidence from the literature on the money-risk trade-offs that is consistent with other studies, such as Murphy and Topel (2006), that

[43]Their database is available online at: http://taxfoundation.org/article/us-federal-individual-income-tax-rates-history-1913-2013-nominal-and-inflation-adjusted-brackets To look for the appropriate income bracket I used income per capita (in nominal terms) in 1950 using data on total GDP from the Bureau of Economic Analysis and population from PWT 9.0.

imply a VSL of $6 million (in 2007 prices). Then, since ample evidence points that the value individuals put into health has been increasing over time proportionally more than income, I apply an income elasticity of the VSL consistent with the literature. As highlighted in the text, given the uncertainty around such elasticities, I have chosen a value that is in the middle of the estimates provided by Costa and Kahn (2004) and Becker and Elias (2007): 1.3. In the robustness tests, I use the elasticities reported in these studies to provide upper and lower bounds for my calculations. The last step in calibrating $\bar{u}$ is choosing a value for this constant so that a 40-year old in 1950 with the welfare function specified earlier (that is with consumption and leisure uncertainty) equals a value of remaining life of $0.81 million (in 1990 prices). In the following section, I will test the robustness of the main results in the text to the main assumptions and parameters considered in the text.

6.3 Other composite indices of human welfare

This section is aimed at complementing the introduction with a more formal exposition of the attempts by Crafts (1997) and Prados de la Escosura (2015) to overcome some of the measurement issues of the HDI.

Beginning with the HDI, this indicator was developed in 1990 and since then it has been published in the Human Development Report by the United Nations Development Programme. This measure aggregates information at the country level on life expectancy, educational attainment and income. To make these elements comparable, they are linearly-transformed and put on a common (0,1) scale as follows:

$$I_x = \frac{x - x^{min}}{x^{max} - x^{min}} \tag{10}$$

where x is life expectancy (LE) or educational attainment (E), x^{min} is their minimum observed value and x^{max} their maximum. From UNDP (2010) onwards, educational attainment is measured considering mean and expected years of schooling (previously literacy and gross enrolment rates were employed). For the dimension of income, gross national income (GNI) is log-transformed and rescaled:

$$I_y = \frac{lnY - lnY^{min}}{lnY^{max} - lnY^{min}} \tag{11}$$

where Y is GNI per capita (GDP per capita was used previously). For each variable, the maximum and minimum values are goalposts that determine the upper and lower bounds of

the sub-indices. Finally, once the three dimensions have been scaled the three of them are aggregated by taking their geometric average:

$$HDI = I_{LE}^{1/3} I_{E}^{1/3} I_{Y}^{1/3} \tag{12}$$

Note that this aggregating procedure is different from the pre-2010 one since it uses a geometric and not an arithmetic average. This change was made in order to avoid a perfect substitutability between its components.[44]

Prados de la Escosura (2015) argues that the variables used for measuring health and educational attainment have asymptotic limits and that therefore an identical increase (in absolute terms) in these variables is lower, the higher their starting level. If we compare life expectancy in countries at very different stages of development, we can observe that mortality declines take place in different parts of the age distribution. In poor countries, life expectancy improvements come from mortality reductions among the youngest and in rich countries among the oldest. Therefore, if a change of similar magnitude is observed in life expectancy and this receives a larger weight in the less developed the country, we are arbitrarily giving more weight to saving the life of younger than older people. By linearly transforming these variables, the author argues, cross-country differences become smaller which introduces a spurious tendency for convergence. In order to correct for this, Prados de la Escosura (2015) draws on Kakwani (1993) which uses a function that allows for increases at higher starting levels to represent larger achievements than at lower starting levels:

$$f(x, x^{min}, x^{max}) = \frac{(x^{max} - x^{min})^{1-\epsilon} - (x^{max} - x)^{1-\epsilon}}{(x^{max} - x^{min})^{1-\epsilon}}, for\ 0 < \epsilon < 1 \tag{13}$$

where x, as before, is life expectancy or educational attainment; the function is a convex function of x. Similar to Equation 10, the index ranges from 0 to 1 (if $x = x^{min}$ and $x = x^{max}$ respectively). Note that if $\epsilon = 0$ the function becomes identical as the form used in the HDI. The formula used in Prados de la Escosura (2015) to create the sub-indices for life expectancy and educational attainment is obtained by considering that $\epsilon = 1$:

$$I_{x,HIHD} = \frac{log(x^{max} - x^{min}) - log(x^{max} - x)}{log(x^{max} - x^{min})} \tag{14}$$

[44]The HDI figures reported in Crafts (2002) were obtained by following a different version of the HDI presented here because it applied the then-used aggregating procedure, literacy and enrolment rates to measure educational attainment and GDP instead of GNP per capita.

For income per capita, Equation 11 is used (since income does not exhibit an asymptotic upper bound, relaxing the property of diminishing returns would drive the development of the indicator).

While Prados de la Escosura (2015) tackles the problems associated with the linear transformation of the non-income variables in the HDI and their tendency towards convergence, it still does not deal with the weighting issue because the HIHD gives the same importance to each dimension. To address this type of concerns, Crafts (1997) used a utility-based indicator drawing on Usher (1980). With this methodology, welfare gains from mortality changes were imputed to income in the following way:

$$\frac{\Delta Y^*}{Y^*} = \frac{\Delta Y}{Y} + (\frac{\Delta L}{L})/\beta \tag{15}$$

where Y^* is GDP adjusted for mortality, L is an age-structure weighted average of discounted life expectancies and β is the elasticity of utility with respect to consumption. To value changes in mortality (the second component in the right-hand side of Equation 15), Crafts uses a value of five percent for the discount rate and 0.25 for β. This calibration is consistent with wage premia for working in urban environments with high mortality levels (Williamson, 1984). Besides health, Crafts accounts for changes in non-market work drawing on Beckerman (1980). By using average wage rates, a positive (or negative) imputation is made for a reduction (or increase) in working time. Contrary to the methodology used in my study, the indicator applied in Crafts (1997) does not allow for cross-country comparisons.

6.4 Robustness Tests

Table 3 presents a number of tests that are aimed at analysing the robustness of the main findings of this paper (see the first row). One of the features of the benchmark results is that I assumed that β=1 (i.e. no discounting except for the inherent one because of mortality). If we set β at 0.98 or 0.96, the results are broadly the same. The only difference is that welfare growth decreases slightly because welfare gains are now discounted. If we allow for consumption to change over the life cycle at a growth rate of two percent or assuming a consumption path as suggested in Fernández-Villaverde and Krueger (2007), Table 3 shows that welfare growth is slightly revised upwards.

The second set of tests shows the robustness of the main estimates to changes in the VSL (and therefore $\bar{u}$) as a result of choosing different VSL elasticities or benchmark values. With respect to the elasticities, I chose a value of 1.3 for being in the middle of the calculations by Costa and Kahn (2004) and Becker and Elias (2007). To get an idea of the extent to

which the estimates differ applying the different elasticities, I took the values reported in these studies and used them for calculating upper and lower bounds of the main findings. As we can see in Table 3 for the United States and Western Europe the difference between the upper and lower range is rather small: 0.9 and 1.2 percentage points respectively. In terms of levels, the welfare ranking stays the same with the difference that using a low VSL elasticity puts Northern Europe almost at the same level of the U.S. in 1913 because health is valued more. Concerning welfare spread, the results are unchanged. A second way of testing the relative importance of health in the model is if, assuming the chosen elasticity is correct, we consider different values for the benchmark VSL. For this purpose, I used the range suggested by Viscusi and Aldy (2003) and, as before, despite small upward and downward revisions in the growth rates and levels, they largely support the main conclusions of the paper. These even hold if I take the VSL used in Becker et al. (2005) which is much lower than the one applied in my study.

The third aspect I will test concerns the way leisure is measured. For this, the most important parameter is the Frisch Elasticity of substitution. Using the range discussed in Jones and Klenow (2016), I use two alternative set of estimates using a value of two and 0.6. Choosing the former gives less relative importance to non-working time and as a result growth rates decline by less than one percent across regions. In terms of levels, the relative differences between regions are very similar and their ranking is the same. If we use a Frisch Elasticity of 0.6, growth rates and levels change more then before, although qualitatively they support the main findings. For instance, growth rates are revised upwards by, on average, two percentage points. However, given the results in the rest of the tests, this specification does not seem very plausible. Finally, if we test the values assigned to the marginal tax rate and the wage to consumption ratio to estimate θ, the results do not change.

The last points that I will test are three. First, in the text I provide welfare levels taking the perspective of a five-year-old. As Table 3 shows, levels (and also growth rates) are unchanged if we consider an age threshold of ten or 20. Second, I show that using data on household and government consumption in 1950 from PWT does not alter the results in that year. This test also provides support to the data on consumption used in this article. Third, one might argue that the way inequality is accounted for might be to some extent at odds with health and leisure in that the last two are calibrated with empirical evidence on individuals' preferences. For this reason, in the last row of the Table 3, I recalculate welfare growth rates and levels excluding inequality from the equations presented in the previous section. The results do not change.

Overall, the first conclusion that the first half of the 20th century is characterised as one of rapid welfare growth is not overturned by these robustness tests. Moreover, these Table

3 also supports the idea that the United States is ahead of the European aggregate around one percent and that the peripheries perform better than countries in the industrial core. Second, Europe has a welfare level of (roughly) two thirds and 40 percent with respect to the United States in 1913 and 1950 respectively; and the best-performing region is typically the northernmost periphery followed by North-Western and Southern Europe. The last finding that implies that by mid-century welfare differences are much larger and persistent than what HDI-evidence suggests also holds since the coefficient of variation for 1950 never falls below the 1913 level.

Table 3: Robustness Tests

	Welfare Growth (1913-1950)					Welfare in 1913					Welfare in 1950					CoV	
	USA	WE	NE	NWE	SE	USA	WE	NE	NWE	SE	USA	WE	NE	NWE	SE	1913	1950
Benchmark results	5.8	4.6	5.5	4.3	5.5	100	67	89	76	40	100	39	84	41	28	0.48	0.52
General parameters																	
Discounting (β=0.98)	5.5	4.3	5.2	4.0	5.0	100	67	85	76	40	100	39	80	41	28	0.48	0.51
Discounting (β=0.96)	5.3	4.0	5.0	3.7	4.7	100	67	83	76	41	100	39	78	40	28	0.48	0.51
Consumption Growth (2 %)	6.2	5.1	5.8	4.8	6.0	100	67	89	76	39	100	39	84	41	27	0.49	0.52
Consumption path	5.9	4.7	5.5	4.4	5.5	100	67	89	76	39	100	42	84	41	27	0.49	0.52
Changes in the base VSL																	
Upper Bound (elast. 1.12)	6.3	5.2	5.9	4.8	6.1	100	63	94	74	35	100	38	87	41	26	0.51	0.53
Lower Bound (elast. 1.50)	5.4	4.0	5.1	3.8	4.8	100	71	84	79	45	100	39	81	41	29	0.46	0.50
Base VSL \$5.5 million	5.6	4.4	5.3	4.1	5.2	100	68	87	78	42	100	39	83	41	28	0.47	0.51
Base VSL \$7.5 million	6.4	5.3	6.0	5.0	6.3	100	62	95	73	34	100	38	87	41	26	0.52	0.53
Base VSL \$2.5 million	4.5	3.0	4.4	2.8	3.5	100	78	76	85	56	100	41	77	42	31	0.41	0.48
Leisure parameters																	
Frisch Elasticity 2.0	4.7	3.9	4.6	3.7	4.5	100	60	83	69	35	100	41	85	44	27	0.47	0.50
Frisch Elasticity 0.6	8.2	6.2	7.3	5.7	7.5	100	85	106	97	52	100	35	83	35	28	0.55	0.56
Ratio Income/Consumption = 0	5.0	4.1	4.9	3.9	4.7	100	62	85	71	36	100	41	85	44	27	0.47	0.50
Marginal Tax Rate = 0	6.4	5.0	5.9	4.6	6.0	100	70	92	80	42	100	38	84	39	28	0.49	0.53
Others																	
Age threshold at 10	5.1	3.5	4.9	3.3	3.9	100	76	84	85	50	100	42	82	43	32	0.43	0.48
Age threshold at 20	5.0	3.4	4.8	3.2	3.8	100	77	85	85	51	100	42	82	43	32	0.43	0.48
Using PWT (C+G) from NA	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	100	42	81	43	34	n.d.	n.d.
Excluding inequality	5.5	4.6	5.1	4.2	5.6	100	60	93	70	32	100	39	84	42	27	0.43	0.49

Note: WE=Western Europe, NE=Northern Europe, NWE=North-Western Europe and SE=Southern Europe. See Table 1 for the country composition of each subregional category. The last two columns show the coefficient of variation (CoV) of welfare levels in the sample for 1913 and 1950.

6.5 Data

Table 4: Income per capita in Western Europe and the United States, 1913-1950

	Level (US=100)					Annual growth rate (in %)		
	1913 (I)	1920 (II)	1929 (III)	1938 (IV)	1950 (V)	1920-29 (VI)	1929-38 (VII)	1913-1950 (VIII)
United States	100	100	100	100	100	2.2	-0.7	1.7
Sweden	53	54	58	76	69	3.0	2.3	2.4
Denmark	74	73	72	91	70	2.1	1.8	1.5
Northern Europe	60	61	63	82	69	2.6	2.1	2.0
United Kingdom	93	83	81	99	73	1.9	1.5	1.0
Germany	68	56	60	75	39	3.0	1.7	0.2
The Netherlands	76	81	84	85	61	2.7	-0.7	1.1
Belgium	81	76	75	78	56	2.2	-0.3	0.7
France	66	60	66	70	53	3.3	-0.1	1.1
Switzerland	137	113	124	127	96	3.3	-0.5	0.7
North-Western Europe	77	68	70	82	55	2.7	1.0	0.8
Spain	39	41	40	41	24	1.9	-0.4	0.3
Italy	43	39	40	43	33	2.3	0.2	1.0
Southern Europe	42	40	40	42	30	2.2	0.0	0.7
Western Europe	67	60	62	72	49	2.5	0.8	0.8

Note: see Table 1 in text.

Table 5: Consumption per capita in Western Europe and the United States, 1913-1950

	Level (US=100)					Annual growth rate (in %)		
	1913 (I)	1920 (II)	1929 (III)	1938 (IV)	1950 (V)	1920-29 (VI)	1929-38 (VII)	1913-1950 (VIII)
United States	100	100	100	100	100	2.3	0.2	1.3
Sweden	57	69	65	82	72	1.6	2.7	1.9
Denmark	79	96	86	100	80	1.1	1.8	1.3
Northern Europe	65	79	73	89	75	1.5	2.4	1.7
United Kingdom	93	100	91	101	76	1.3	1.4	0.8
Germany	50	41	47	47	32	4.0	0.2	0.1
The Netherlands	63	65	61	62	52	1.7	0.4	0.8
Belgium	85	87	82	84	60	1.8	0.4	0.4
France	62	68	63	61	48	1.6	-0.3	0.6
Switzerland	94	101	100	106	86	2.2	0.7	1.1
North-Western Europe	67	68	66	69	51	2.1	0.6	0.6
Spain	42	50	50	48	28	2.4	-0.2	0.2
Italy	41	49	43	41	34	1.0	-0.6	0.8
Southern Europe	41	49	46	44	32	1.5	-0.4	0.6
Western Europe	61	64	61	63	47	1.9	0.4	0.6

Note: see Table 1 in text.

Table 6: Gini coefficients in Western Europe and the United States, 1913-1950

	1913 (I)	1929 (II)	1950 (III)
United States	51	54	39
Sweden	57	51	40
Denmark	41	43	36
Northern Europe	51	48	39
United Kingdom	42	43	30
Germany	49	49	40
The Netherlands	47	42	36
Belgium	50	52	38
France	55	62	58
Switzerland	38	39	41
North-Western Europe	48	50	41
Spain	35	36	35
Italy	49	51	43
Southern Europe	44	46	40
Western Europe	47	49	41

Note: source is van Zanden et al. (2013). See footnote 12 for Germany.

Table 7: Life Expectancy at birth in Western Europe and the United States, 1913-1950

	1913 (II)	1920 (III)	1929 (IV)	1938 (V)	1950 (VI)
United States	54	57	59	61	68
Sweden	58	60	62	65	71
Denmark	58	60	62	64	70
Northern Europe	58	60	62	65	71
United Kingdom	53	57	59	62	69
Germany	49	57	59	61	67
The Netherlands	56	59	63	67	71
Belgium	51	54	57	60	66
France	50	53	55	59	66
Switzerland	53	57	60	63	68
North-Western Europe	51	56	58	61	67
Spain	42	42	49	52	62
Italy	47	48	52	56	65
Southern Europe	45	46	51	55	64
Western Europe	50	54	57	60	66

Note: see Table 3 in text.

Table 8: Annual hours worked in Western Europe and the United States, 1913-1950

	1913 (I)	1929 (II)	1938 (III)	1950 (IV)
United States	2900	2316	1756	2008
Sweden	2745	2152	2131	2009
Denmark	2731	2301	2203	2071
Northern Europe	2740	2206	2158	2032
United Kingdom	2656	2257	2200	2112
Germany	2723	2128	2187	2372
The Netherlands	2942	2233	2281	2156
Belgium	2841	2229	2196	2404
France	2933	2198	1760	2045
Switzerland	2704	2281	2085	2092
North-Western Europe	2769	2192	2093	2209
Spain	2601	2342	2030	2052
Italy	2953	2153	2162	1951
Southern Europe	2829	2222	2114	1989
Western Europe	2783	2200	2101	2140

Note: see Table 4 in text.

European
Historical
Economics
Society

EHES Working Paper Series

Recent EHES Working Papers

2017

EHES 113 Portugal's wine globalization waves, 1750-2015
Pedro Lains

EHES.112 Effects of reforms and supervisory organizations: Evidence from the Ottoman Empire and the İstanbul bourse
Elmas Yaldız Hanedar, Avni Önder Hanedar and Ferdi Çelikay

EHES.111 Resource endowments and agricultural commercialization in colonial Africa: Did labour seasonality and food security drive Uganda's cotton revolution?
Michiel de Haas and Kostadis J. Papaioannou

EHES.110 Currency Valuations, Retaliation and Trade Conflicts: Evidence from Interwar France
Thilo Albers

EHES.109 Measuring Extractive Institutions: Colonial Trade and Price Gaps in French Africa
Federico Tadei

EHES.108 Between war and peace: The Ottoman economy and foreign exchange trading at the Istanbul bourse
Avni Önder Hanedar, Hatice Gaye Gencer, Sercan Demiralay and İsmail Altay

EHES.107 Why did Argentina become a super-exporter of agricultural and food products during the *Belle Époque* (1880-1929)?
Vicente Pinilla and Agustina Rayes

EHES.106 Plague and long-term development: the lasting effects of the 1629-30 epidemic on the Italian cities
Guido Alfani and Marco Percoco

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