Capital, Productivity, and Human Welfare Since 1870

Daniel Gallardo-Albarrán

Contents

Abstract

This chapter reviews the proximate factors of human welfare since 1870 by discussing two strands of the economic history literature and identifying various key areas for further research. The first strand focuses on level accounting studies that attribute between-country economic inequality to differences in capital and productivity. I argue that most income gaps in the late nineteenth century were due to variation in physical and human capital endowments, while widening productivity differentials account for most of rising cross-country inequality during the twentieth century. These patterns are likely explained by waves of skill- and capital-biased technological innovation, but additional research is needed to underpin these findings in, at least, three ways: capital and income series should be deflated by appropriate price indices, samples should include

I am grateful to Thomas Westland, who read an earlier version of this manuscript and provided useful comments. All errors are, obviously, my own.

D. Gallardo-Albarrán (✉)
Wageningen University, Wageningen, The Netherlands
e-mail: daniel.gallardoalbarran@wur.nl

C. Diebolt, M. Haupert (eds.), *Handbook of Cliometrics*,
https://doi.org/10.1007/978-3-031-35583-7_88

many more lower-income countries and methodologies could explore more realistic production functions. The second strand of the literature I review considers the measurement of long-run human development. Three approaches are popular among practitioners (capability, data-driven and utility frameworks), although there is still no consensus on which one to use. This makes it challenging to interpret broad trends in human welfare, as different well-being indices show contrasting patterns of growth and inequality. I argue that the field needs a more solid theoretical foundation to guide our choice of measurement frameworks. In this respect, utility-based indicators may be especially useful, as they address relevant issues raised in the literature, such as how to weight different dimensions, how individuals trade-off between them, and how to interpret the results.

Keywords

Capital · Economic history · Human development · Productivity · Well-being

Introduction

One of the major goals of economic history is to understand the evolution and drivers of prosperity in the long run. Societal progress in a number of key areas has led to unprecedented levels of human well-being, as a result of improvements in both income and nonpecuniary aspects of life, such as education or health. This chapter focuses on the proximate factors of human welfare – henceforth I will use the following terms indistinctively: welfare, well-being, broader living standards, and human development – by discussing two different, though related, strands of the economic history literature. The first examines the relative role of capital and productivity in accounting for economic inequality across countries since 1870. More specifically, this chapter considers studies performing development-accounting exercises to understand how the production possibilities of countries are influenced by technological change, the organization of labor, and processes of capital accumulation. The second branch of the reviewed literature pays particular attention to how human welfare has been measured in history and what the main (proximate) sources of welfare growth have been since 1870, as one considers nonmonetary dimensions of human lives. Each section concludes with a discussion of how the field could move forward in a few key areas.

Capital and Productivity During "Divergence, Big Time"

The unequal spread of industrialization and modern economic growth ushered in an era of both unprecedented increases in material living standards and rising economic inequality between countries. Widening income gaps became particularly noticeable during the twentieth century as some economies had been profiting from steady

productivity growth and structural transformation for decades, while others remained largely agricultural. The rising trend of economic inequality between countries was so far-reaching and steep during the twentieth century that this period has been referred to as one of "divergence, big time" (Pritchett 1997). Between 1870 and 1990, the income ratio between the United States and countries at the lower end of the distribution, such as Chad, has soared from about 8 to 45.

How can we make sense of the drivers of this divergence? One popular approach in economics and economic history is the use of accounting techniques that allow decomposing income differences, across countries and time, into factor inputs (e.g., labor and capital) and a residual that is often interpreted as a measure of productivity and technological progress. These are usually referred to as "proximate" determinants of economic development. In the following, this section presents this framework and then reviews the literature using it to summarize our current knowledge on the proximate drivers of rising economic inequality during the twentieth century.

Level Accounting

The literature often considers the following production function as a starting point:

$$Y = A\,f(L, K), \tag{1}$$

where Y is total output that is determined by total factor productivity (A) and two measurable inputs of the production process, namely, labor (L) and productive capital (K). It is then common to consider the average "quality," or human capital, of the workforce in the production process and to use a Cobb-Douglas production to model how factor inputs contribute to total output so that the previous equation becomes:

$$Y = A\,H^{1-\alpha}K^{\alpha}, \tag{2}$$

where α refers to the elasticity of output with respect to productive capital and $H = Lh$, where h is average human capital, typically, in the form of education. This formulation implies that factor inputs exhibit diminishing returns and constant returns to scale, and that technology is factor neutral. A final transformation is applied to link changes in factor inputs with changes in economic growth that consist of dividing both sides of the equation by the amount of raw labor (L), so that each term is expressed in per unit of labor: $y = Ah^{1-\alpha}k^{\alpha}$.

Before proceeding further, it is useful to distinguish between two different, though closely related, strands of the literature. The first refers to growth-accounting studies that focus on economic growth, so they apply the previous equations to a given country and period of time, to determine the relative importance of productivity, human capital, and productive capital in accounting for *temporal changes* in gross domestic product (GDP) per worker (e.g., Crafts and Woltjer 2021). The focus of this chapter, however, is on differences in *levels* of income so the previous

equations will be employed to determine the relative importance of *cross-country differences* of factor inputs in accounting for differences in output per unit of labor in a given year. Accordingly, the following equation can be proposed:

$$y_i = A_i {k_i}^{\alpha} {h_i}^{1-\alpha}, \tag{3}$$

where i indexes countries and each component is expressed in per-worker terms for simplicity, except for A. Although there is no subscript for time, Eq. 3 can be applied to any year or period. In this setting, total factor productivity (TFP) can be interpreted as a measure of technological development because it captures the effectiveness with which inputs are combined to obtain output. However, we should do this with a certain degree of caution because A is a very imperfect metric of technology. Strictly speaking, TFP refers to that aspect of the production process that is left unexplained after measuring the contribution of observed factor inputs to output. As explained below, this indicator is sensitive to the inclusion of additional inputs (e.g., health as a part of human capital) or to assumptions of the production function (e.g., the output elasticity of productive capital).

Equation 3 is intuitive and widely used in the literature, but it does not separate the impact that exogeneous changes in productivity have on factor input accumulation. More specifically, an exogeneous improvement in productivity, assuming constant investment rates, leads to higher output and capital per worker. While in this example the increase in output is due to higher TFP, Eq. 3 attributes parts of this to capital accumulation. To overcome this, one can express Eq. 3 in capital-output ratios so that changes in capital-output ratios allow for per-worker capital to vary:

$$y_i = A_i^{\frac{1}{1-\alpha}} \left(\frac{K_i}{Y_i} \right)^{\frac{1}{1-\alpha}} h_i. \tag{4}$$

To obtain an estimate of unobserved TFP, we can use data on y, k and h and then calculate the implied value for A for any country in a given year. Additionally, if we express these terms relative to a certain country, we can compute how much more/less productive a country is relative to the chosen benchmark. It is also possible to obtain a summary metric of the relative importance of TFP and factors inputs for a whole sample, and not just country pairs, by performing a variance decomposition exercise. There are several approaches in the literature that answer different questions, such as that of Caselli (2005, p. 687): If all countries have the same TFP level, what would income dispersion in the sample be? Or that of Klenow and Rodríguez-Clare (1997, p. 80): when we see 1% higher income in one country relative to the sample mean, how much higher is our conditional expectation of ${k_i}^{\alpha} {h_i}^{1-\alpha}$. Alternatively, one can look at countries at different parts of the income distribution, such as the 90th versus 10th income percentiles, 90th versus 50th income percentiles, or 50th versus 10th income percentiles.

The simplicity of this framework comes with a very important caveat: The results cannot be interpreted causally. Investment in productive capital over the long term

can causally explain economic development because of the mere accumulation of capital and its gradual technological upgrading. In addition, both factor accumulation and productivity can ultimately depend on the same underlying driver (e.g., institutions protecting property rights). Also, some of the assumptions laid out above might be unrealistic at times. These limitations aside level (and growth)-accounting exercises allow us to refine our narratives of economic development and to provide a hint at what should be examined causally. Whether findings from causal empirical frameworks support those obtained from accounting exercises or not, understanding the reasons behind their differences provides additional insights into the process of economic development.

The Proximate Determinants of Income Inequality Since 1900

Productivity and Efficiency

What is the relative importance of factors inputs and TFP in accounting for income differentials over the long term? Development-accounting studies from the economics literature can only answer this question for recent periods, because they have predominantly focused on recent years, mostly due to data availability. They indicate that productive capital differentials account for about one-third of cross-country income variation, and TFP accounts for the rest (Caselli 2005; Jones 2016). Is this pattern representative of longer periods of time? An early assessment of the proximate determinants of relative levels of economic development in 1910 was provided by Clark and Feenstra (2003) for about 20 countries. They find that, as common for recent benchmarks, TFP explains a significant fraction of income gaps. However, these results must be interpreted with caution because the authors made two methodologically convenient assumptions to estimate a version of Eq. 3 "*without having reliable data on capital stock across countries*" (Clark and Feenstra 2003, p. 284): full capital mobility and equalization of the rental cost of capital due to international capital flows.

Gallardo-Albarrán and Inklaar (2021) took a different approach and constructed a dataset with productive capital stocks for 33 countries since 1913, with uninterrupted investment series starting (in some cases) in the nineteenth century building on Gallardo-Albarrán (2018b). They show that the relative importance of TFP has been far from constant over time. In fact, it accounted for about 40% of income variation in the early decades of the twentieth century, as compared with about 70% in the early twenty-first century. The rise of TFP happened mostly in two periods: 1930s–1950s and 1970s–1980s. What do these developments tell us about productivity change? To answer this question, I follow Gallardo-Albarrán and Inklaar (2021), who draw on secondary literature to argue that two factors likely played an important role: diffusion patterns of new technologies and the efficiency with which these were adopted and applied.

Beginning with the diffusion of technological change, the rising wave of TFP during the interwar era can be explained as the result of rapid technological change

in the United States due to improvements in technologies stemming from the second industrial revolution, such as electricity or the internal combustion engine (Gordon 2016). These technologies crossed the Atlantic and Pacific oceans after the Second World War prompting TFP increases in Europe and Japan (Bergeaud et al. 2016), so the adoption lag pushed up the relative importance of TFP after 1930. In other parts of the world, such as East Asia, investment-led convergence during the 1960s and 1970s pushed down the relative importance of productive capital (Young 1995).

The second factor that can explain why TFP has grown much more important over the twentieth century relates to the "appropriateness" of new technologies. If most technological change happens at high levels of productive and human capital, then they are designed to solve problems present in such contexts and are thus not suitable for low-end technological environments. Comin and Mestieri (2018) provide some evidence supporting this, since they observe that while initial adoption lags of new technologies have dramatically gone down (e.g., Internet vs railways), the intensity with which they are used has become more unequal. We can get a better sense of the relative importance of efficiency in accounting for income differences by moving beyond the development account framework above. More specifically, we can draw on nonparametric exercises to decompose TFP differences into differences in technology and efficiency. Jerzmanowsky (2007, p. 2098) shows that the world technology frontier has predominantly moved up at high levels of income per capita since 1960, and that many countries have been falling behind in terms of efficiency, as the percentage of income variance accounted for by efficiency has risen from 28 in 1960 to 43 in 1995. If we focus on Europe exclusively, we find that TFP differentials (relative to the United States) were relatively large by 1960 and mostly due to efficiency differences (Crafts and Woltjer 2021, p. 19). Do we observe something similar before the mid-twentieth century? Allen (2012) provides an answer to this question by estimating frontier production functions capable of tracking local technical change. According to this result, rich countries have consistently defined the world production frontier since 1820 at relatively high levels of capital per worker, so it seems this phenomenon has become more pronounced over time. Although not everyone agrees that efficiency plays a crucial role – Ziebarth (2013) argues that resource misallocation in the late-nineteenth-century United States is comparable to current levels in China and India – the economic history literature suggests this is likely to be the case.

If efficiency is so important to explain TFP and between-country income inequality, what then determines efficiency in the first place? Management practices are one likely candidate, since they impact how labor and capital are allocated within firms. For the United States, Alexopoulos and Tombe (2012) argue that managerial innovations during the twentieth century are associated with substantial aggregate output and productivity increases. However, evidence from India during the first half of the twentieth century indicates the opposite, as managerial practices do not explain low economic performance in the textile sector (Wolcott and Clark 1999). The field would benefit substantially from more long-term comparative research to assess these claims in other contexts.

Human Capital

Human capital in the form of education is an important input of the production process in the framework presented above (see Eq. 4). Several growth-accounting studies point out that it plays an important role, though not overwhelming over the very long term. Crafts and O'Rourke (2014, pp. 285–286) estimate that the percentage of growth attributed to human capital rarely exceeds 20% in a number of industrialized economies after 1960. If we focus on the first half of the twentieth century, the authors find that in places like France and Germany, human capital added between 0.36 and 0.22 percentage points per annum, respectively, to labor productivity growth, which in turn grew annually at 2.01 and 1.05 percentage points. Gallardo-Albarrán and Inklaar (2021, p. 962) consider a larger number of countries that span most of the income distribution and show that the relative contribution of human capital to income differences rose from slightly more than 20% in 1913 to almost 40% in 1935, and only after the mid-1990s did it fall below 20%.

The estimates above might look disappointing to readers placing more emphasis on human capital. However, it is important to take into account what can (and cannot) capture. Two points are particularly important. First, h in Eq. 4 simply refers to the (private) economic returns of attaining different levels of education without considering the role that human capital plays in the generation and diffusion of new ideas, since an educated labor force is better able to make use of cutting-edge technologies and adopt them to local contexts. In addition, education interacts with demographic trends that have a deep impact on economic development, such as mortality and fertility.

Second, returns to years of schooling are usually measured rather crudely in accounting exercises, since it is often assumed that they are constant across countries and time. More specifically, $h_i = e^{\mu_i(s)}$, where s is average years of schooling and $\mu_i(s)$ is a piecewise linear function with different slopes given observed returns in world regions with different levels of schooling time (also education quality is often ignored). The problem with this approach is that education returns are determined by technology (typically) increasing demand for skilled workers and the supply of education determining the relative scarcity of workers with different skills. Goldin and Katz (2000) refer to the changing relative importance of these countervailing forces as a *race* between education and technology: When education is winning this race, then education returns decrease, and the opposite happens when technology takes the lead. In the United States, the interaction of these two forces have resulted in high returns at the turn of the twentieth century followed by a decline by midcentury and an additional rise as education demand has outpaced supply.

An aspect of human capital that has been comparatively neglected concerns workers' health. This is unfortunate given that health enhances the development of cognitive skills, mental concentration, and energy available for work. Fogel (1994) argues that low nutritional status in historical Great Britain hampered the productive capacity of the labor force significantly, especially among the poorest who barely had any energy to perform demanding physical activities. After 1790, two elements vastly improved the amount of hours that people could work as well as

their productivity: rising calories available for work and the improved efficiency of human bodies to convert energy into work output due to richer diets, better clothing and shelter, and a lower incidence of infectious diseases. These two mechanisms, account for about 50% of British economic growth between 1790 and 1980 (Fogel 1994, p. 388). Arora (2001) expanded Fogel's efforts to ten industrialized countries since the late nineteenth century and found that health improvements have increased the long-term growth path of these countries by about 30–40%. In other words, health changes are not merely the outcome of income changes, but rather they are growth enhancing. Some likely important mechanisms are that higher life expectancy allows for more accumulation of learning and work experience. Further, lower mortality means a greater pool of individuals and ideas stemming from them that can end up shifting up the knowledge frontier leading to faster growth.

These studies, while informative, do not tell us much about the potential impact of spatial health differences on levels of economic development. They focus on the temporal aspect of the health-income relationship, and their samples consist of industrialized countries. Gallardo-Albarrán (2018a) overcomes these two limitations by analyzing a sample of 36 high-, medium-, and low-income countries. Using an extension of the level-accounting framework presented above, he uses a human capital composite that includes both education and health to calculate the percentage of cross-country income dispersion that is accounted for by health differentials between 1900 and 2008. He finds that the relative importance of health, as measured by life expectancy, has greatly changed during the twentieth century. Around 1900, before major declines in mortality happened, health differentials account for almost a fifth of cross-country income inequality. This number rose to 26% by midcentury, and then it declined until 1990 as a result of international health convergence; since then, failure to further convergence has kept the contribution of health gaps to income differentials at about 10%.

Productive Capital

How have cross-country differences in productive capital evolved during the twentieth century? Intuitively, the results by Gallardo-Albarrán and Inklaar (2021, p. 962) pointing at rising TFP imply that cross-country capital gaps must have gone down. Indeed, their show that during the first half of the twentieth century its variance declined by two-thirds indicating that some investment-driven income convergence took place then. After the mid-twentieth century, this had little potential to lead to additional income convergence, though, because capital only accounted for about 20% of income variance, especially after the 1980s. Another interesting from this dataset concerns the accumulation of productive capital. They observe that capital-output ratios since 1913 have exhibited a constant upward trajectory, a trend that does not line up with one of the stylized facts of Kaldor (1961). They find that average capital-output ratios in their sample have risen from 1.7 to 3.4 in three waves during the periods 1910s–1950s, 1970s, and 2000s–2010s. In addition, as one might expect from their TFP results, the largest increases in capital-output ratios have taken place in lower-income countries.

These trends resonate with earlier evidence on Latin America by Tafunell and Ducoing (2016, p. 60), who find that the capital-output ratio in Chile rose from 1.3 to 1.9 between 1890 and 1991, while that of Mexico increased from 0.5 to 2.1 during the same period. Also, the argument that capital is becoming more important as a percentage of aggregate output has been made by Piketty and Zucman (2014). Although their capital measure and methodology are different than those used by the aforementioned studies, the consistent pattern of these trends is remarkable, at least for high-income countries.

Moving Forward

This section highlights a few aspects of the literature that need attention in the future. The first concerns price indices. In growth accounting, the income and productive capital series are expressed in constant prices *over time* so that inflation does not influence the calculations. In development accounting, prices should be constant *across space* so that levels of income and capital are comparable across countries with different price levels. Long-run monetary indicators in economic history are typically provided at internationally comparable prices that are constant both over time and across countries (e.g., 1990 Gheary-Khamis dollars), as pioneered by the work of Angus Maddison (1995). However, this approach has limitations when prices are difficult to track (e.g., during military conflicts, inflationary periods, or without a solid statistical foundation) or when they exhibit measurement error and biases. In these cases, resulting income levels from backward extrapolation can be unreliable. Ward and Devereux (2021, p. 235) have overcome this problem for a small subset of Western countries, in 1872 and 1910. The differences between GDP per capita using this method and backward extrapolation following Maddison (1995) are striking. For instance, the latter results in about 80% higher GDP per capita in the case of Switzerland, and 40% lower Norwegian GDP per capita in 1910. Also, the economic performance of the United Kingdom looks much less exceptional in 1872 when using current-price data.

The construction of accurate price deflators is necessary not only for income datasets, but also for productive capital. To be more specific, capital-output ratios in Eq. 4 should be adjusted for relative prices of capital and income. Ideally, $\frac{K_i}{Y_i}$ should be:

$$\frac{K}{Y} = \left(\frac{p^k K}{PPP^K}\right) / \left(\frac{p^y Y}{PPP^Y}\right) \equiv \frac{p^K K}{p^Y Y} \times \frac{PPP^Y}{PPP^K}, \tag{5}$$

where PPP^Y and PPP^K are the purchasing power parity for GDP and capital, respectively, and p^Y and p^K are the price deflators of GDP and capital, respectively. The development accounting estimates discussed above by Gallardo-Albarrán and Inklaar (2021) do not take into account differences in the relative levels of purchasing power parities for income and capital. Instead, they use current price capital-output ratios $\left(\frac{p^K K}{p^Y Y}\right)$. Although their evidence indicates that the PPP ratio $\left(\frac{PPP^Y}{PPP^K}\right)$ is close to one from 1950 onward and thus their estimates would not be biased (at least

since then), proper capital PPPs are necessary to obtain accurate development-accounting estimates. An additional issue to consider when building price deflators for productive capital concerns the relative prices of different assets, since they have evolved differently over time. Technological improvements have increased productivity in the manufacturing sector to a larger extent than in the construction sector. Therefore, the output price of manufacturing relative to construction in, say, 2011 will be much smaller than before technological innovations took place. This relative price effect will result in an overestimation of past capital stocks, as the weight of machinery and equipment will be underestimated in favor of buildings. Consider the case of Britain: A chained index leads to 25% lower capital stocks as compared to using a deflator calculated in 2011 relative prices (Gallardo-Albarrán 2018b, pp. 88–89). In sum, more research using historical prices is needed to construct current-price GDP and productive capital figures to fully comprehend their implications for our narratives of economic development and its proximate drivers. Similar data improvements are needed in terms of adding additional asset-specific information, natural resources, or intangible capital.

The second way in which accounting exercises can be improved relates to departing from traditional production functions, which exhibit limitations with regard to some of their assumptions, such as neutral technological change or constant returns to scale. In this respect, analyses employing nonparametric functions that allow inferring the most appropriate production function from data can be an important step forward. One possible approach is to use Data Envelopment Analysis (DEA) to create global best-practice productivity frontiers using the highest levels of output per unit of work across different levels of capital intensity. Then, one could use this to attribute labor productivity differences to the extent to which a country can realize the maximum observed potential and different levels of capital per worker. A productivity decomposition based on a smooth Cobb-Douglas production function will tend to overstate the role of TFP differentials when the marginal product of capital computed using DEA is higher than that implied by a smooth function. This is a case that happens very often when technological changes are *local*, i.e., specific to a given (typically high) level of capital intensity. The benefits of this approach have been used by Timmer et al. (2016) to study the labor productivity gap between the United States and Germany. They argue that it emerged due to the inefficient assimilation of major American innovations that required substantial learning and adaptation to local circumstances in Germany. This created a reserve productivity potential that was only realized after the Second World War. Similarly, Woltjer (2013) argues that the United Kingdom did not realize the productivity potential of American technologies stemming from the second industrial revolution until a process of learning by doing reduced the impact of inefficient assimilation of foreign techniques. A different approach involves using a constant elasticity of substitution production function to understand the direction and bias of technical change. Kukić (2021) applies this to a number of planned economies after 1950 and finds that, unlike others suggest, socialist economies experienced rapid increases in labor efficiency, converging with the United States. Therefore, the reason why their overall economic performance was disappointing has to do with the declining

efficiency of capital. In sum, these studies demonstrate that applying new methods to a dataset with internationally comparable data on productive capital might revise some of current knowledge about the proximate determinants of economic development during the twentieth century.

Another important limitation of the literature concerns data availability for low-income countries. African countries are massively underrepresented in growth- and level-accounting exercises before the mid-twentieth century. To the extent this is possible, the field would greatly benefit from the inclusion of economies that have faced a different set of political, geographic, and demographic constraints. The development of more comprehensive datasets on GDP per capita is a first step (e.g., Broadberry and Gardner 2022), but similar information should be available in terms of productive capital, price indices, and educational attainment. This point applies to some Latin American and Asian economies as well for which historical national accounts are not available.

The Rise of Human Welfare During the Twentieth Century

The Limitations of Economic Indicators

The previous section has focused on the role of capital and productivity on economic development and material living standards. But human lives have improved in so many respects that we should go beyond an analysis of economic measures exclusively, if we are ultimately interested in how individuals and societies flourish. This is not to say that escaping poverty, rising purchasing power, and the enjoyment of new products and services are unimportant. Rather, my argument is that a more encompassing approach to the measurement of living standards has clear benefits toward understanding the causes and consequences of overall human progress. Since an overemphasis on economic indicators limits our perspective in several ways. First, it provides a one-sided view of what progress really means for individuals, and it ignores a significant part of economic history research that has valuable insights into the long-run condition of human societies. Certainly, say, a fivefold increase in purchasing power is indicative of human flourishing, but it misses much of the human experience in terms of suffering due to ailments affecting ourselves, relatives, and friends; our intellectual development; the ability to participate in civil society and political processes; social and economic discrimination; etcetera. Second, although it is often argued that income correlates with aspects of people's lives that they tend to value, there are many instances in history when changes in economic performance do not do any justice to improvements in people's lives, such as the early phase of the industrial revolution or the first half of the twentieth century (more on this below). The same applies to cross-sectional comparisons. As discussed above, between-country income inequality has markedly increased during the twentieth century, while life expectancy and education exhibited the exact opposite trend: substantial convergence, especially after 1950 (Neumayer 2003). The third reason why purely economic indicators limit our perspective on human

flourishing relates to data availability and the challenges that price indices face to capture the impact of changing relative prices, quality improvements, and the emergence of new products and services. While these issues are not exclusive to economic history research, the emphasis of the field on long-run comparisons makes them more salient. In this respect, other indicators can give us clues about historical living conditions (e.g., infant mortality, life expectancy, or educational attainment) that are more widely available and less prone to suffer from the aforementioned comparability issues. A related issue concerning comparability emerges when making GDP comparisons between countries with substantially different levels along noneconomic dimensions. Lower- and medium-income countries today have similar levels of GDP per capita as industrialized countries in the past, but their health and education performance tend to be (in many cases) comparatively much better. Income levels in this case say little about human development.

In sum, GDP per capita and similar metrics are not good (proxy) measures of broader living standards across countries and time. For this purpose, we need different frameworks that put multidimensional well-being at the center of the analysis. The next section presents some of the main attempts to measure well-being used in economic history. Before this, however, it is important to motivate the use of comprehensive frameworks that aggregate a number of variables expressed in different units. This is particularly relevant for those arguing that a dashboard approach might be preferred (i.e., presentation of a number of indicators measuring different aspects of citizens' lives), on the basis that aggregation procedures are difficult and involve value judgment about the relative importance of each component. Although this is a valid critique that will be considered below in greater detail, it is worth highlighting that value judgments are not an inherent feature of composite measures *exclusively*. A reader presented with a dashboard of metrics needs to prioritize which ones will receive more or less attention, and this process will involve value judgments as well. This is particularly the case during periods when different indicators present markedly different trends, such as the early phase of English industrialization (ca. 1750–1850). Up to the Napoleonic wars, stagnating (or even deteriorating) real wages were coupled with rising working time and economic inequality, while life expectancy increased substantially (Gallardo-Albarrán and de Jong 2021, p. 11). After that, purchasing power improved and mortality declines came to a halt. Given these contrasting trends, how should we then assess the consequences of industrialization for the English population? In this, and similar historical cases, a dashboard approach might yield more questions than answers on the topic of the causes and consequences of long-term well-being.

Approaches to Well-Being in Economic History

The economic history literature has approached the concept of well-being in a number of different ways. This section mostly focuses on approaches using objective indicators of well-being instead of subjective ones, such as happiness or life satisfaction (Hills et al. 2019; Lack 2021), because the former group has received most

attention. Also, an important part of the field has used anthropometric information to study biological well-being, but this literature is vast, and for reasons of space it is not covered here (e.g., Floud et al. 2011; Harris 2021).

The first major framework employed in the field to understand and measure the proximate determinants of human flourishing is the capability approach by Sen (1999), which argues that human development is a process of enlarging the choices that people can make. Therefore, attaining different levels of human development depends on the extent to which individuals lead the lives that they want to lead. *Capability* within this framework is the freedom to choose between different *functionings*, which Sen defines as things that people value doing or being, such as being nourished and healthy or taking part in society. This perspective, for instance, captures the crucial difference between a person that decides to cut down food consumption voluntarily by fasting, and another who is forced into such a situation unwillingly. The ability to choose is essential in Sen's theory.

The rich theoretical framework by Amartya Sen is not easy to bring to practice; however, the attempt first presented in the 1990 Human Development report has received a lot of attention: the Human Development Index (UNDP 1990). This metric was initially used for ordinal (rather than cardinal) purposes, and it focuses on three dimensions that provide a concise snapshot of people's lives: health, education, and income. Indicators expressed in different units relate to these dimensions: GNI per capita for income; life expectancy at birth for health; and expected years of schooling and mean years of schooling for education. To make these variables comparable across countries and time, standardized indices are built ranging from 0 to 1 as follows:

$$\text{Index} = \frac{\text{observed value} - \text{min. value}}{\text{max. value} - \text{min. value}}, \tag{6}$$

where minimum and maximum values are goalposts representing lowest and highest attainable values (e.g., for years of schooling, these are 0 and 18, respectively). *Index* in Eq. 6 strictly refers to health and education only. For GNI per capita, the same basic equation is used but each term on the right-hand side is log-transformed (UNDP 2022). Given that each underlying component of the HDI proxies capabilities in its individual dimension, the logarithmic transformation reflects that there are diminishing returns into how income translates into capabilities (Anand and Sen 2000, p. 100). To put it another way, the extent to which freedom to choose a life without material deprivation increases is larger at low levels of income than at high levels. Ever since the inception of the HDI, similar indices have been developed that adjust it for economic or gender inequality, or how each dimension is captured. These alternatives fall within the capability approach and will be discussed below.

Although the field of economic history quickly embraced the HDI and similar indices (Crafts 1997; Steckel and Floud 1997), it has received criticisms from studies in both economics and economic history. The equal-weighting scheme is regarded as arbitrary (Nordhaus 2003, p. 20), and the implied trade-offs between dimensions are seen by some as "unacceptable" (Ravallion 2012b, p. 208). An additional limitation of a fixed weighting scheme is that various aspects of well-being might be valued

differently at different times, as implied by historical evidence from tradeoffs between health and income (Gallardo-Albarrán 2019, pp. 65–66). Even irrespective of the weighting scheme (and the dimensions included), the HDI has been considered equivalent to a paternalistic social welfare function that merely represents the ethical systems of researchers (Amendola et al. 2023). Finally, Costa and Steckel (1997, pp. 73–74) argue that the interpretation of the HDI in the past in relation to income is not straightforward. They indicate that the HDI should be seen as a *distance* index conveying information on when modern levels of human development were achieved, while income is a *velocity* measure reflecting material improvements by contemporaries. A modest rate of income growth was certainly important to historical populations, but in the HDI it presents a relatively small achievement because it is just a small fraction of modern income levels. Consider the following example: A country with a level of income of 20,000 dollars scores 0.81 in the income index of the HDI, assuming that the minimum and maximum goalposts are 100 and 75,000, respectively. If we reduce the maximum goalpost to, say, 50,000, then the income index becomes 0.87;[1] note that this point also applies to the other dimensions. This feature of the HDI suggests that comparisons with income metrics should be made carefully. Fleurbaey (2009) provides a more elaborated discussion on other critiques.

An alternative approach to the HDI is data driven. In this framework, choices are made by each author according to data considerations (e.g., availability, reliability) or the specific focus of research. After considering a number of indicators deemed relevant to people's lives, their aggregation can be made according to different principles. One consists of standardizing each component and then taking an average assuming equal weights, and another uses a latent variable model that informs the researcher about relative weights. This latter method delivers relative weights for each well-being dimension on the basis of how correlated the different dimensions are. Highly correlated measures will obtain higher weights given that they exhibit the same underlying information, assuming that this correlation is caused by a common latent variable (Slottje 1991). This method is particularly useful if the researcher accepts that these indices proxy for the same underlying concept. However, this is an assumption that does not need to be true (or desirable), since some variables can reflect different things and thus may not be correlated with each other (Ravallion 2012a, p. 12). Principal component analysis avoids some of the arbitrary weighting issues mentioned above, although it can be criticized for lacking theoretical underpinning. Also, latent variable models can be sensitive to the inclusion of different variables as well as data availability and quality.

A third approach that deals with some of the issues mentioned above draws on utility theory. In this framework, individual well-being depends on the amount of (expected) utility they can derive from living a healthy life, consuming goods and services, having leisure time, and so on. To be more specific, consider the following utility function proposed by Jones and Klenow (2016, p. 2430):

[1] $\frac{\log(22,000)-\log(100)}{\log(75,000)-\log(100)} = 0.81$; $\frac{\log(22,000)-\log(100)}{\log(50,000)-\log(100)} = 0.87$.

$$u(C, l) = \overline{u} + \log C + v(l), \tag{7}$$

where the utility (or well-being) of an individual depends on her consumption (C) and leisure time (l). In this type of framework, the focus is placed on trade-offs between different aspects of people's lives, and from these one can infer the relative weight of each dimension. To be more specific, the relative weight assigned to leisure in Eq. 7 depends on the functional form of $v(l)$, which Jones and Klenow (2016) calibrate using information on the Frisch elasticity of labor supply; this parameter shows how workers respond to wage changes. In a similar manner, we could assign a value to $\overline{u}$ by using information on the monetary value that people assign to various risk environments. A historical example of trade-offs between income and health risks comes from the decisions that workers made when migrating to urban areas in England. Williamson (1990) argues that they received a wage premium in places with high levels of infant mortality, which he interprets as evidence that they had to receive a monetary compensation to withstand the disamenities that came with living in industrializing towns. Unlike the HDI which considers income, health, and education as equally important, a utility-based framework requires information on how to weigh each of the dimensions (Becker et al. 2005; Jones and Klenow 2016). In addition, the use of individual preferences provides a closer perspective of contemporaries' well-being. Certainly, this method does not avoid altogether value judgments, since there are important choices to be made concerning the specific form of the utility functions employed or the set of preferences used. For instance, should everyone have the same utility function? Can individual preferences (or well-being levels) be aggregated across groups? However, these choices are assessed and underpinned within a theoretical framework that has a long tradition in both economics and economic history (e.g., Usher 1973; Williamson 1984).

Well-Being in Historical Perspective

What can we learn from well-being indicators about the evolution and drivers of human flourishing? To answer this question, this section focuses on the consequences of the English industrial revolution between 1750 and 1850 for workers' welfare, and the period of "divergence, big time" (1870–2000) considered at the beginning of this chapter.

The Industrial Revolution

The positive impact of industrialization for long-run human development in England is hardly disputed, but the same does not apply to the period 1750–1850. In fact, the consequences of industrialization for workers' lives have been fiercely debated since the days of Karl Marx, Friedrich Engels, and Charles Dickens. The literature is too vast to summarize it here (see Gallardo-Albarrán and de Jong 2021; Taylor 1975), but suffice it to say that two distinct positions can be identified. On one side, a number of studies hold a positive view of the period by arguing that real wages rose substantially during the period, offsetting any negative consequences that

industrialization and urbanization might have brought about. Other authors, instead, have a more pessimistic stance and argue that stagnating health and rising inequality and working time were not compensated by rising purchasing power.

Given these disparate trends, what can we infer about the evolution of English workers' well-being during industrialization? Crafts (2002, p. 625) puts together different estimates of the HDI, and other indices that are not discussed here, and found a clear rising pattern conveying a rather optimistic view of aggregate well-being during the period. More specifically, his index rose by almost 50% between 1760 and 1850, with most of this increase happening after 1800. Adjusting the HDI by the level of inequality yields a slightly lower improvement in human development, although still very substantial (between 42% and 48%). Voth (2004) counters Crafts' optimistic assessment of this period by creating a pseudo-HDI that replaces life expectancy at birth with average statures. This version of the HDI presents an almost 20% increase, which (almost) exclusively takes place before 1820. This new index provides a more optimistic view of the last decades of the nineteenth century.

One problem with using the HDI to assess workers' well-being during industrialization is that it does not take into account inequality and working time trends. These evolved rather negatively, since according to Allen (2019, pp. 110–111), the Gini coefficient increased from 0.53 to 0.6 between 1759 and 1798, and it remained high until the mid-nineteenth century. Similarly, annual hours worked rose from about 2600 to 3350 between 1760 and 1830, and then they slightly declined by 1850 (Voth 2001, p. 1078). To take these trends into account, Gallardo-Albarrán and de Jong (2021) take a utility approach to create a more encompassing metric that jointly considers income, health, leisure, and inequality. Their findings are more pessimistic than those of earlier work using the HDI (and similar) indices. In particular, they obtain decreasing levels of human welfare between 1760 and 1800, as a result of rising working time and inequality and lower purchasing power. After this year, or rather after the Napoleonic Wars, overall well-being likely increased as real wages grew and the rising trend of working time and inequality came to a halt. The plateau reached by life expectancy barely contributed to changes in well-being given that it remained at about 40 years (Gallardo-Albarrán and de Jong 2021, p. 12). The trends in human welfare shown by these authors are also more pessimistic than those by Williamson (1984, p. 168), who created a metric combining income and health changes, and calculated that English well-being rose between 1% and 1.15% during the period 1781–1851. Without considering hours worked and inequality, a study of workers' broader living standards during industrialization is likely to yield overly optimistic outcomes.

Past research quantifying the broader impact of industrialization has predominantly focused on economic and demographic outcomes. However, other variables are likely important in assessing the impact of industrialization, such as biodiversity or environmental degradation. Also, there are interactions between different dimensions of well-being. For instance, air pollution in industrial areas greatly influenced people's health during the nineteenth century. According to Beach and Hanlon (2018), coal use was so detrimental to people's health that it accounts for about one-third of urban mortality penalty by the middle of the nineteenth century. Although the health impact of pollution might be captured to some extent by life

expectancy estimates, there are more reasons why one would care about having a clean environment. The same applies to river pollution as a result of industrial and sewage waste. However, more research is needed into understanding how these dimensions interact in creating an environment that people value.

Another point to consider in future work looking at human welfare during industrialization concerns the reference or benchmark we use to value changes in well-being. Previous studies have considered a representative individual (or average person), but the focus on averages masks substantial variation across different regions and groups of people. In fact, taking a perspective different from the "average" can have a great impact on estimated well-being trends. Gallardo-Albarrán and García-Gómez (2022) demonstrate this with the case of a leading industrial city during Spanish industrialization (Alcoy). Using detailed wage and mortality information from this town, they provide utility-based calculations of well-being by sector, age, and gender. They find that human welfare stagnated in the agricultural sector, while that of workers in the industrial and service sectors rose significantly. In terms of health, they quantify the welfare impact of taking different age perspectives to establish the contribution of health changes to well-being, by considering a measure that mostly reflects children's health (i.e., life expectancy at birth) and another that influences adults' health (life expectancy at age 15). Their results show that adults gained 50% less well-being than children, since most health improvements during the epidemiological transition come from reductions in infant and child mortality. Gallardo-Albarrán (2019) provides a more refined calculation of the impact of considering mortality changes for different age groups. More specifically, he calculates utility-based welfare growth rates for a number of industrialized countries between 1913 and 1950 by assuming the counterfactual that mortality improvements only took place after a given age threshold. His results point out that well-being growth rates tend to decline rapidly in the 0–5 age bracket and that not considering these mortality improvements pushes down welfare growth rates between 0.5 and 1 percentage points annually (Gallardo-Albarrán 2019, p. 71). In sum, future work should engage more deeply with distributional issues of overall living standards along different dimensions, such as income, occupation, gender, age, or location (e.g., urban or rural).

Global Well-Being During the Period of "Divergence, Big Time"

The industrial revolution gradually spread from England to a number of countries after the mid-nineteenth century. This process ushered in an era of rapid income growth in some parts of the world that led to increasing between-country income inequality after 1870. At the same time, changes in other dimensions of well-being took place, such as health, education, or working time. When we put together these two developments since 1870, the following two questions emerge: the first has a temporal focus and asks how much lives have changed; the second has a cross-sectional focus and relates to whether well-being diverged or converged across countries.

Beginning with the former question, we can consider the HDI estimates by Crafts (2002) to get a sense of human progress since 1870. As mentioned above, HDI scores are better interpreted as a distance measure, so direct comparisons with income (related to velocity) will not be provided. Considering the case of the United

States, he finds that most HDI points (ca. 70%) between 1870 and 1999 were achieved before 1950; the same applies to some European countries. If we consider a country with a lower initial level of human development, such as India, the opposite pattern emerges: Most human progress (ca. 75%) during the period 1913–1999 took place after 1950. If we consider the case of Mexico, the most (positively) impactful period was 1913–1975, as 80% of its HDI trajectory since 1913 was achieved then. If we compare the growth differences in terms of HDI points, an interesting pattern emerges: Mexico has experienced most human growth, and India's trajectory is comparable to that of the United States. Of course, part of the reason why this is the case has to do with the comparatively high initial level of HDI in the United States (0.506 as compared to 0.143 and 0.27 in India and Mexico, respectively). However, these trends are very different if one considers income growth, as income levels in the United States have grown much more than in other parts of the world.

How do well-being growth rates compare to those of income? As explained above, it is conceptually difficult to compare GDP per capita growth in constant prices with HDI scores. Utility indicators, however, are better suited for this, because well-being or utils are measured in monetary terms. For the United States, Nordhaus (2003) provides this during the twentieth century and finds that the value of life expectancy improvements is somewhat larger than the total consumption of goods and services. He argues that while consumption per capita grew at an average of 2.1% per year, improved health added between 2.2% and 3% annually (Nordhaus 2003, p. 29). If we consider specific periods during the twentieth century, his calculations are in line with the patterns shown by the HDI because most welfare gains from rising life expectancy took place before 1950.

This pattern does not support the usual narrative of disappointing economic performance during the first half of the twentieth century. In fact, this makes this period particularly relevant for examining how income and welfare metrics compare due to the contrasting trends in key aspects of people's lives. On one side, income per capita exhibited historically low growth rates due to the negative economic impact of armed conflicts, macroeconomic policies, and rising protectionism. On the other hand, European citizens experienced unprecedented improvements in health, leisure, and inequality, due to the diffusion of the germ theory of disease, the discovery of antibiotics, and the implementation of the 8-hour day. Gallardo-Albarrán (2019) provides comparable GDP per capita and welfare growth rates for some European countries and the United States drawing on the framework by Jones and Klenow (2016), which is similar to that of Nordhaus (2003) but more comprehensive because it allows for including working time and inequality. According to his estimations, human welfare in Western Europe grew by 2.4 percentage points annually between 1913 and 1950, while the growth rate of GDP per capita barely reached one percentage point. The main contributor to human welfare was life expectancy, which contributed to overall well-being by 1.4 percentage points yearly. Similar, and even higher, growth rates have been found for other countries during the twentieth century, such as Japan, Taiwan, or Chile (Usher 1973, p. 224). The contribution of rising leisure and equality was much more modest, about 0.2

percentage points (Gallardo-Albarrán 2019, p. 69). In sum, although these frameworks do not consider the role of education, we can conclude from the evidence presented here that health is a major contributor to human welfare during the twentieth century (if not the most important), as it tends to double the growth of living standards as conventionally measured with GDP per capita.

I now turn to cross-sectional comparisons to establish how between-country inequality of well-being has evolved since 1870. As before, a good starting point is the work of Crafts (2002) using the HDI. The trends shown by this index tend to be more optimistic in terms of cross-country convergence than those using GDP per capita. Developing countries in the late twentieth century exhibit comparatively high levels of well-being relative to industrialized countries in 1870. Indeed, countries such as Venezuela, Sri Lanka, or Paraguay scored higher than the United States already in 1950. By 1999, most countries in his sample had a higher level of human development than North America in 1913 (Crafts 2002, pp. 396–398). As a result of this, the notion of "divergence, big time" associated with income inequality during the twentieth century is not supported by the HDI. In fact, the ratio between the top and bottom 20% of the HDI (country) distribution has declined from 4.3 to 1.8 during the period 1913–1999, while that for income rose from 4 to 13.2 (Crafts 2002, p. 403). Rapid increases in educational attainment and life expectancy after 1950 explain this pattern (Neumayer 2003).

The idea that convergence has marked the international evolution of well-being is not uncontested. Prados de la Escosura (2022) draws on the HDI framework but argues that the upper bounds of the health and education variables creates a natural tendency to converge when analyzing long periods of time. He tries to overcome this by applying a nonlinear transformation to the nonincome dimensions of the HDI, arguing that marginal increases in these dimensions at high levels represent larger marginal contributions to human development than at low levels. In addition, he adds a fourth dimension to his augmented HDI, liberal democracy, to incorporate the notion that individual freedom is necessary to choose between different bundles of functionings. He finds that while the relative gap between the *West* and the *Rest* has declined, especially after 1930, the absolute gap between these two groups of countries has risen substantially (Prados de la Escosura 2022, pp. 104–109). Therefore the degree of human development convergence is much lower than that implied by HDI scores. Indeed, despite substantial gains in life expectancy and educational attainment pushing up well-being in lower-income countries, progress has slowed down in the last decades.

Another study that allows us to assess the international distribution of well-being across countries since 1870 is that of Rijpma (2014). He takes a data-driven approach using principal component analysis to aggregate information from nine indicators that capture changes in dimensions that are not included in the HDI, such as environmental diversity, income inequality, or violence. Although his results are not directly comparable with the studies mentioned previously, the comprehensive nature of his index makes it interesting to examine whether overall trends are in line with those of the HDI. He finds increasing between-country inequality between 1850 and 1900 as a group of forerunners exhibit much higher levels of well-being than the

rest of the world. Inequality kept increasing until 1950 and then it declined by 2000 (Rijpma 2014, pp. 264–265). Therefore, and in line with HDI-based results, it seems that the international distribution of human welfare has become more equal during the second half of the twentieth century.

What do utility-based indicators show about welfare convergence? Sadly, similar worldwide estimates as those presented earlier are not available since 1870. This is unfortunate given that this approach provides an alternative way to assign relative weights that can yield substantially different results, as discussed earlier with the case of the English industrial revolution. However, some studies allow us to get a grasp of the likely trends we might expect. Becker et al. (2005) constructed a welfare index combining income and health for a large sample of countries between 1960 and 2000, and they found that welfare convergence was the norm, due to mortality reductions from infectious, respiratory, and digestive diseases. One important limitation of this study, though, is that it uses GDP per capita to proxy for cross-country welfare differences in 1960 because their methodology does not allow them to create welfare levels then. This is problematic given that the post-1950 period is precisely the one in which the vast majority of low-income (and low welfare) countries have experienced substantial life expectancy gains. d'Albis and Bonnet (2018) revise these calculations, using a methodology that allows for computing welfare levels in 1960 and 2000, and argue that welfare inequality did not decline during this period. Similarly, although for a much smaller sample, Gallardo-Albarrán (2019, p. 68) indicates that welfare dispersion across countries in a sample of ten European countries and the United States rose between 1913 and 1950, while that of the HDI decreased significantly. All in all, these results suggest that indicators based on capabilities or principal component analysis tend to yield more optimistic cross-sectional trends than utility-based measures. At present we have no tools to assess which one of these views is more likely to be true. However, I think these different results should be taken as a sign of how little we know about the world distribution of well-being and that we might be wary of overly optimistic assessments of between-country inequality.

Moving Forward

This section discusses three points that might be important in future studies of well-being. The first relates to potential concerns about double counting. The correlation between the underlying components of, say, the HDI or a utility measure can be partly causal. For instance, consider how income leads to better nutrition, health care, and sanitary infrastructures. In this case, changes in purchasing power contribute to well-being via a direct channel (e.g., consumption) and an indirect channel (e.g., as an input in the health production function). Therefore, by including health outcomes in a composite index, we are already taking into account the indirect channel, which would result in double counting. This point applies to other dimensions, such as education or leisure, as well as reverse relationships between these and income, since health and education are causal determinants of income as well. Earlier studies in the field took this issue into account by adjusting the contribution of health improvements to well-being according to the share of those improvements

that are due to income (Williamson 1984). However, recent work has largely neglected this issue, which has an important implication: metrics of (supposedly) actual or experienced well-being include inputs, intermediaries, and outputs of their underlying dimensions. Clearly, more research is needed to understand and model the dependencies between the underlying components of composite indices.

The second point that needs further investigation concerns fixed weighting, a common approach in the literature, which can be problematic when making comparisons over long periods of time. For instance, historical evidence points at the rising relative importance of health for both nation-states and individuals (Costa and Kahn 2004; Mokyr and Stein 1996). Therefore, if current-day values about health are applied to historical settings, then this dimension is likely to determine a disproportionally large part of the indicator (Gallardo-Albarrán 2019). Further research is needed into the implications of projecting current beliefs about well-being when it comes to other elements such as education, both at a conceptual and practical level.

The third point, and perhaps the most difficult, is about choosing a specific framework. The results presented in the previous section show clearly that well-being indices display different patterns than purely economic indicators, both in the cross-sectional and in the temporal dimension. GDP per capita is a good index of average material living standards, but it should *not* be used to proxy for well-being in history. Which indicator should we use for this purpose? This question is hard to answer given that there are several approaches to measure well-being and, in turn, many versions of each one of these. In addition, some frameworks ignore sources of well-being stemming from gender equality or the preservation of nature. Therefore, looking for a single overarching or unifying metric might be an impossible task. Before accepting this somewhat disappointing conclusion, however, it is worthwhile exploring whether there is an underlying component that unifies them. Gallardo-Albarrán (2019, Appendix E) looked at country rankings obtained with his method and the HDI, and he concluded that the two of them are highly correlated. However, this exercise refers to a small sample and might not be representative of earlier and later periods.

We can extend this idea and repeat the exercise for a larger number of countries over a longer time span drawing on the literature presented earlier. For this purpose, I created an unbalanced sample with information on country-specific levels of HDI-, AHDI-, and PCA-based well-being for 1870, 1913, 1950, 1975, and 2000; for utility welfare, the data refers to 2007.[2] Then, for each year considered I ranked each

[2] The HDI scores are taken from Crafts (2002, pp. 396–398) for the years 1870, 1913, 1950 1975, and 1999; AHDI figures by Prados de la Escosura (2022) for the years 1870, 1913, 1950, 1975, and 2000 were downloaded from https://frdelpino.es/investigacion/category/01_ciencias-sociales/02_economia-mundial/03_desarrollo-humano-economia-mundial/ [accessed on 25-11-2022]; the PCA-based indices by Rijpma (2014) were generously shared by the author via email, and they refer to 1870, 1910, 1950, 1970, and 2000; utility-based welfare calculations by Jones and Klenow (2016) for 2007 were downloaded from https://web.stanford.edu/~chadj/papers.html#rawls [accessed on 25-11-2022].

country according to its level of human progress according to the four indicators. The focus on rankings is appropriate here due to their straightforward interpretation; the same does not apply to relative levels or growth rates due to different measurement units. Table 1 presents their pairwise correlation. As we can see, the correlation coefficients are extremely high, generally above 0.9 and as high as 0.96, which means that in terms of ranking countries by their level of well-being, the four measures result in very similar results on average. This result is encouraging, since it implies that the overall picture portrayed by each one of them is rather similar. However, when considering country-specific experiences, there can be some deviations, given that the correlations between indices are not perfect (see Fig. 1).

Table 1 Country ranking correlation using different measures of human welfare

	HDI	AHDI	PCA index	Utility index
HDI	1			
AHDI	0.93	1		
PCA index	0.96	0.94	1	
Utility index	0.96	0.90	0.91	1

Note: the underlying data refer to country rankings (1 is the highest score, 2 the second, etcetera). The correlations between the HDI, AHDI, and the PCA index are based on 259 observations unequally distributed over 5 years: 1870, 1913, 1950, 1975, and 2000. Those referring to the utility index are based on 69 observations. See footnote 4 for the sources

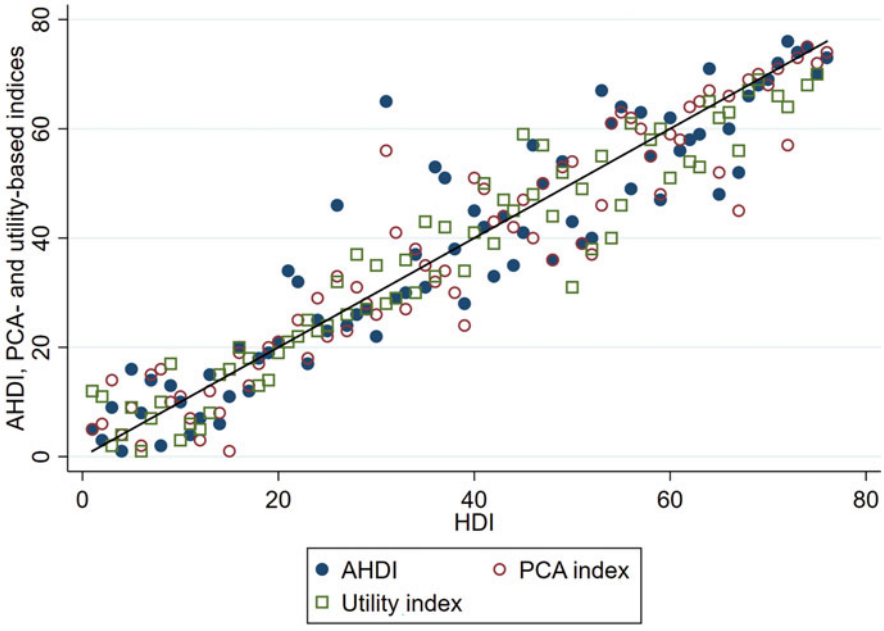

Fig. 1 Country ranking by its level of well-being in ca. 2000 according to four indices. Note: See Table 1 for the sources. The 45-degree line illustrates what a perfect correlation with the HDI would look like

In sum, the viewpoint that this chapter tries to convey is that each approach has positive and negative aspects, and that researchers should be aware of the tradeoffs involved. A theoretically grounded framework, however, does present important benefits as compared to data-driven exercises, given that theory guides which dimensions should be included in the final indicator, and how these contribute to well-being. In addition, the evidence presented above suggests that the theoretical underpinning of the HDI is less consistent than that of welfare economics when it comes to taking into account citizens' individual preferences and their implied trade-offs (Fleurbaey and Blanchet 2013), or including additional variables that are not part of its traditional formulation.

Conclusion

Economic historians have put together a vast amount of evidence in the last two decades that sheds new light on the proximate determinants of human welfare. Capital and productivity have been significant contributors to the rise in material living standards since 1870 as well as the ensuing cross-country inequality that followed. Income gaps between poor and rich economies soared to unprecedented levels during the twentieth century, mostly due to widening productivity differentials. The opposite was true in the late nineteenth century, when productive and human capital accounted for a large share of income differences between countries. Future research can underpin the trends highlighted in this chapter, as new datasets are built with more complete information covering a larger number of world regions, and new methodologies are applied to these.

When it comes to understanding the proximate determinants of human welfare, it seems that the main bottleneck of the field is not data driven, but rather theoretical. A plethora of methods exist to measure historical well-being as consensus on which one to use is still lacking. I identified three broad approaches (i.e., capability, data-driven, and utility frameworks) which seem to have enjoyed some popularity among practitioners. After discussing some of their relative strengths and weaknesses, I highlighted the need to further develop existing theoretical frameworks to create well-being indices that capture the most important aspects of people's lives throughout history. In this respect, utility-based indicators can be particularly useful given that they provide a solution to relevant issues mentioned in the literature, such as fixed weighting schemes, individuals' trade-offs, or easiness of interpretation.

Acknowledgments I acknowledge financial support from NWO (grant no. VI.Veni.201H.048).

References

Alexopoulos M, Tombe T (2012) Management matters. J Monet Econ 59(3):269–285

Allen RC (2012) Technology and the great divergence: global economic development since 1820. Explor Econ Hist 49(1):1–16

Allen RC (2019) Class structure and inequality during the industrial revolution: lessons from England's social tables, 1688–1867. Econ Hist Rev 72(1):88–125
Amendola N, Gabbuti G, Vecchi G (2023) On some problems of using the human development index in economic history. Eur Rev Econ Hist. https://academic.oup.com/ereh/advance-articles
Anand S, Sen A (2000) The income component of the human development index. J Hum Dev 1(1): 83–106
Arora S (2001) Health, human productivity, and long-term economic growth. J Econ Hist 61(3): 699–749
Beach B, Hanlon WW (2018) Coal smoke and mortality in an early industrial economy. Econ J 128(615):2652–2675
Becker GS, Philipson TJ, Soares RR (2005) The quantity and quality of life and the evolution of world inequality. Am Econ Rev 95(1):277–291
Bergeaud A, Cette G, Lecat R (2016) Productivity trends in advanced countries between 1890 and 2012. Rev Income Wealth 62(3):420–444
Broadberry S, Gardner L (2022) Economic growth in sub-Saharan Africa, 1885–2008: evidence from eight countries. Explor Econ Hist 83:Article 101424
Caselli F (2005) Accounting for cross-country income differences. In: Aghion P, Durlauf SN (eds) Handbook of economic growth, vol 1A. Elsevier, Amsterdam, pp 679–741
Clark G, Feenstra RC (2003) Technology in the great divergence. In: Bordo MD, Taylor AM, Williamson JG (eds) Globalization in historical perspective. University of Chicago Press, Chicago, pp 277–321
Comin D, Mestieri M (2018) If technology has arrived everywhere, why has income diverged? Am Econ J Macroecon 10(3):137–178
Costa DL, Kahn ME (2004) Changes in the value of life, 1940–1980. J Risk Uncertainty 29(2): 159–180
Costa DL, Steckel RH (1997) Long-term trends in health, welfare, and economic growth in the United States. In: Steckel RH, Floud R (eds) Health and welfare during industrialization. University of Chicago Press, Chicago, pp 47–90
Crafts NFR (1997) The human development index and changes in standards of living: some historical comparisons. Eur Rev Econ Hist 1(3):299–322
Crafts NFR (2002) The human development index, 1870–1999: some revised estimates. Eur Rev Econ Hist 6(3):395–405
Crafts NFR, O'Rourke KH (2014) Twentieth century growth. In: Aghion P, Durlauf SN (eds) Handbook of economic growth, vol 2A. Elsevier, Amsterdam, pp 263–346
Crafts NFR, Woltjer PJ (2021) Growth accounting in economic history: findings, lessons and new directions. J Econ Surv 35(3):670–696. https://doi.org/10.1111/joes.12348
d'Albis H, Bonnet F (2018) Inequalities in life expectancy and the global welfare convergence. Econ Lett 168:49–51
Fleurbaey M (2009) Beyond GDP: the quest for a measure of social welfare. J Econ Lit 47(4): 1029–1075
Fleurbaey M, Blanchet D (2013) Beyond GDP: measuring welfare and assessing sustainability. Oxford University Press, Oxford
Floud R, Fogel RW, Harris B, Hong SC (2011) The changing body: health, nutrition, and human development in the Western world since 1700. Cambridge University Press, Cambridge
Fogel RW (1994) Economic growth, population theory and physiology: the bearing of long-term process on the making of economic policy. Am Econ Rev 84(3):369–395
Gallardo-Albarrán D (2018a) Health and economic development since 1900. Econ Human Biol 31: 228–237
Gallardo-Albarrán D (2018b) Health, well-being and inequality over the long term. Doctoral dissertation, University of Groningen (SOM research school), Groningen
Gallardo-Albarrán D (2019) Missed opportunities? Human welfare in Western Europe and the United States, 1913–1950. Explor Econ Hist 72:57–73

Gallardo-Albarrán D, de Jong H (2021) Optimism or pessimism? A composite view on English living standards during the Industrial Revolution. Eur Rev Econ Hist 25(1):1–19
Gallardo-Albarrán D, García-Gómez JJ (2022) Growth or stagnation? Well-being during the Spanish industrialization in Alcoy (1860–1910). Investig Hist Econ 18(1):26–37
Gallardo-Albarrán D, Inklaar R (2021) The role of capital and productivity in accounting for income differences since 1913. J Econ Surv 35(3):952–974
Goldin C, Katz LF (2000) Education and income in the early twentieth century: evidence from the prairies. J Econ Hist 60(3):782–818
Gordon RJ (2016) The rise and fall of American growth: the U.S. standard of living since the civil war. Princeton University Press, Princeton
Harris B (2021) Anthropometric history and the measurement of wellbeing. Vienna Yearb Popul Res 19:1–33
Hills TT, Proto E, Sgroi D, Seresinhe CI (2019) Historical analysis of national subjective wellbeing using millions of digitized books. Nat Hum Behav 3:1271–1275
Jerzmanowsky M (2007) Total factor productivity differences: appropriate technology vs. efficiency. Eur Econ Rev 51(8):2080–2110
Jones CI (2016) The facts of economic growth. In: Taylor JB, Uhlig H (eds) Handbook of macroeconomics, vol 2A. Elsevier, Amsterdam, pp 3–69
Jones CI, Klenow PJ (2016) Beyond GDP? Welfare across countries and time. Am Econ Rev 106(9):2426–2457
Kaldor N (1961) Capital accumulation and economic growth. In: Hague DC (ed) The theory of capital. Palgrave Macmillan, London, pp 177–222
Klenow PJ, Rodríguez-Clare A (1997) The neoclassical revival in growth economics: has it gone too far? In: Bernanke BS, Rotemberg J (eds) NBER macroeconomics annual 1997, vol 12. MIT Press, Cambridge, pp 73–114
Kukić L (2021) The nature of technological failure: patterns of biased technical change in socialist Europe. J Econ Surv 35(3):895–925
Lack P (2021) Using word analysis to track the evolution of emotional well-being in nineteenth-century industrializing Britain. Hist Methods J Quant Interdiscip Hist 54(4):228–247
Maddison A (1995) Monitoring the world economy, 1820–1992. Organization for Economic Cooperation and Development, Paris
Mokyr J, Stein R (1996) Science, health, and household technology: the effect of the Pasteur revolution on consumer demand. In: Bresnahan TF, Gordon RJ (eds) The economics of new goods. University of Chicago Press, Chicago, pp 143–206
Neumayer E (2003) Beyond income: convergence in living standards, big time. Struct Chang Econ Dyn 14(3):275–296
Nordhaus WD (2003) The health of nations: the contribution of improved health to living standards. In: Murphy KM, Topel RH (eds) Measuring the gains from medical research: an economic approach. University of Chicago Press, Chicago, pp 9–40
Piketty T, Zucman G (2014) Capital is back: wealth-income ratios in rich countries 1700–2010. Q J Econ 129(3):1255–1310
Prados de la Escosura L (2022) Human development and the path to freedom. 1870 to the present. Cambridge University Press, Cambridge
Pritchett L (1997) Divergence, big time. J Econ Perspect 11(3):3–17
Ravallion M (2012a) Mashup indices of development. World Bank Res Obs 27(1):1–32
Ravallion M (2012b) Troubling tradeoffs in the human development index. J Dev Econ 99(2): 201–209
Rijpma A (2014) A composite view of well-being since 1820. In: van Zanden JL, Baten J, d'Ercole MM, Rijpma A, Timmer MP (eds) How was life? Global well-being since 1820. OECD Publishing, Paris, pp 249–269
Sen A (1999) Development as freedom. Alfred Knopf, New York
Slottje DJ (1991) Measuring the quality of life across countries. Rev Econ Stat 73(4):684–693

Steckel RH, Floud R (eds) (1997) Health and welfare during industrialisation. University of Chicago Press, Chicago
Tafunell X, Ducoing C (2016) Non-residential capital stock in Latin America, 1875–2008: new estimates and international comparisons. Aust Econ Hist Rev 56:46–69. https://doi.org/10.1111/aehr.12076
Taylor AJ (1975) The standard of living in Britain in the industrial revolution. Debates in economic history. Methuen, London
Timmer MP, Veenstra J, Woltjer PJ (2016) The Yankees of Europe? A new view on technology and productivity in German manufacturing in the early twentieth century. J Econ Hist 76(3):874–908
UNDP (1990) Human development report 1990. UNDP, New York
UNDP (2022) Human development report 2021–22: uncertain times, unsettled lives: shaping our future in a transforming. UNDP, New York
Usher D (1973) An imputation to the measure of economic growth for changes in life expectancy. In: Moss M (ed) The measurement of economic and social performance. Columbia University Press for National Bureau of Economic Research, New York, pp 193–223
Voth H-J (2001) The longest years – new estimates of labor input in England, 1760–1830. J Econ Hist 61(4):1065–1082
Voth H-J (2004) Living standards and the urban environment. In: Floud R, Johnson P (eds) The Cambridge economic history of modern Britain. Volume 1: Industrialisation, 1700–1860. Cambridge University Press, Cambridge, pp 268–294
Ward M, Devereux J (2021) New income comparisons for the late nineteenth and early twentieth century. Rev Income Wealth 67(1):222–247
Williamson JG (1984) British mortality and the value of life, 1781–1931. Popul Stud 38(1):157–172
Williamson JG (1990) Coping with city growth during the British industrial revolution. Cambridge University Press, Cambridge, MA
Wolcott S, Clark G (1999) Why nations fail: managerial decisions and performance in Indian Cotton Textiles, 1890–1938. J Econ Hist 59(2):397–423
Woltjer P (2013) The roaring thirties: productivity growth and technological change in Great Britain and the United States during the early twentieth century. Doctoral degree, University of Groningen, Groningen
Young A (1995) The tyranny of numbers: confronting the statistical realities of the East Asian growth experience. Q J Econ 110(3):641–680
Ziebarth NL (2013) Are China and India backward? Evidence from the 19th century U.S. census of manufactures. Rev Econ Dyn 16(1):86–99