INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

On Some Problems of Using the Human Development Index in Economic History

Nicola Amendola [a]
Giacomo Gabbuti [b]
Giovanni Vecchi [a]

[a] University of Rome "Tor Vergata", Italy.
[b] St. Antony's College, University of Oxford, UK and Scuola Superiore Sant'Anna, Pisa, Italy.

2021/42 **November 2021**
ISSN(ONLINE) 2284-0400

On Some Problems of Using the Human Development Index in Economic History

Nicola Amendola

(University of Rome "Tor Vergata")

nicola.amendola@uniroma2.it

Giacomo Gabbuti

(St. Antony's College, University of Oxford,

and Scuola Superiore Sant'Anna, Pisa)

giacomo.gabbuti@santannapisa.it

Giovanni Vecchi

(University of Rome "Tor Vergata")

giovanni.vecchi@uniroma2.it

Acknowledgments. We are grateful to Brian A'Hearn, Andrea Brandolini, Stefano Fenoaltea, Ernesto Savaglio, Gianni Toniolo, and to the participants at the vice presidential session "The Development of Wellbeing in History" at the XVIII World Economic History Congress for their helpful comments.

On Some Problems of Using the Human Development Index in Economic History

Abstract

We argue against the use of composite indices, such as the Human Development Index (HDI), in economic history. We show that the HDI can be interpreted as a formal representation of the analyst's ethical system. We support our claim by introducing a new class of *paternalistic social welfare functions* (Graaff 1957, Mas-Colell, 1995) which encompasses all the HDI formulas put forth by the literature. The theoretical framework is illustrated by an empirical investigation of the long-run evolution of Italians' living standards and civic liberties. We conclude that any history based on composite indices is one where both data and history play a minor role, if any.

JEL codes: N01, N3, O15

Keywords: human development index; economic wellbeing; composite indices; living standards; CES; social welfare functions; Italy.

Introduction

Composite well-being indices, the best known of which is the Human Development Index (HDI), have been undeniably successful in the field of economic history. The first applications date back to the pioneering work of Morris (1979), Federico and Toniolo (1991), but it is mainly thanks to contributions from Crafts (1997a, 1997b, 2002) and Costa and Steckel (1997), relating to industrialisation processes in Great Britain and the United States and their impact on the living conditions of the population, that the HDI has received increasing attention from economic historians. Over the last two decades, the literature has been consolidated to explore, with a broad geographical reach, the implications of the use of the HDI in the historical analysis of economic development processes. In some cases, the use of the HDI or its variants such as the Hybrid HDI, Improved HDI, Historical HDI, Augmented HDI (Felice and Vasta 2014, and Prados de la Escosura 2010, 2015a, 2018, 2021, respectively) has made it possible to redesign long-term growth paths otherwise based on the traditional macroeconomic gross domestic product indicator (e.g. Bèrtola and Ofield, 2012; Crafts, 2004; Felice and Vasta, 2014; Devereux, 2020; Prados de la Escosura, 2013, 2015a, 2015b; Bértola and Williamson, 2017). Overall, the use of composite indices is therefore not merely an enrichment of the quantitative framework but rather an analytical and methodological innovation with substantial consequences for historical economic analysis.

In this paper we do not want to question the importance of the "marriage" between composite indices and economic history, nor do we want to discuss the substance of

the historiographic debate from which this relationship was derived. Instead, we will try to identify and discuss the interpretative limitations of the long-term estimation exercises of composite indices, in the light, above all, of the unclear or undefined link between these indices and economic theory. In our view, this is a very reasonable objective, since those who use this tool to carry out quantitative historical economic analysis clearly require it to be fully consistent with the principles underlying the economic theory of reference, be it unorthodox or mainstream.

The main conclusion of our analysis is that the use of the HDI for historical and economic analysis is subject to some important conceptual limitations resulting from the lack of a well-defined relationship between individual well-being and social preferences. Fully combining the HDI with modern economic theory is possible by assimilating it to a paternalistic social welfare function (de Graaff 1967, Mas-Colell 1995, Yang 2018). In this case, however, the HDI index will only reflect the analyst's preferences as regards the aggregate dimensions of wellbeing and not the preferences of the individuals for whom an aggregate measure of wellbeing is to be provided. The "decoupling" of individual preferences clearly renders the HDI index sensitive to the discretionary choices of the analyst, in this case the economic historian, not only in relation to the relative weight of the dimensions of well-being included in the index itself, but more generally in relation to individual preferences over possible alternative social arrangements.

The HDI was explicitly created as an attempt to respond to the deep dissatisfaction with traditional macroeconomic welfare indicators (Land and Michalos, 2018). In

this sense, there are two possible interpretations of the HDI index: *i)* an interpretation consistent with the "capability approach" proposed by Amartya Sen (Sen, 1985), whereby aggregating health, education and income dimensions is a way of empirically capturing the complex relationships between functionings and capabilities; *ii)* a welfarist interpretation, where health and education represent an extension of the dimensions of individual well-being otherwise encapsulated by the dimension of monetary income alone.

The first interpretation is certainly the most coherent in analytical historical terms. As noted by Prados de la Escosura (2018), the HDI originates precisely from the attempt to operationalise Sen's capability approach.[1] However, there are a number of criticisms of the ability of the HDI index to adequately capture the full scope of the methodology proposed by the Indian economist and philosopher (Sagar, 1997). According to Fleurbaey and Blanchet (2013), the HDI index is only "a pale reflection of the general and ambitious methodology proposed by the capability perspective". Offer (2003, 2006) notes that although Sen has influenced the development of HDI, he has never embodied its approach in any metrics, precisely because "in keeping with Liberal values, he has not privileged any particular good. Even under indigence it was necessary to respect individual priorities." In the same theoretical framework as Sen, Offer notes the components "do not lend themselves to aggregation, indexation, or a focal-point summary figure". Even leaving aside these criticisms,

[1] See also UNDP (1999) and Haq (2003).

the index would be a measurement tool based on a theory of non-traditional social choices, characterized by an ambiguous relationship with methodological individualism (Robeyns, 2017) and utilitarian consequentialism (Scanlon 2001, Qizilbash 2021), the two pillars of modern economic theory. The possibility of a consistent methodological comparison with the main traditional welfare indicators generally used to describe the trajectory of modern economic development would therefore not be possible. An additional complication then arises from the specific application of the HDI tool to long-term historical data, which tacitly assumes that there is a stable relationship between functionings, capabilities and the dimensions of well-being embodied in the index itself, which is by no means easy to justify from a historical point of view.

The second, welfarist, interpretation, although partially at odds with the intellectual origins of the HDI, is the most widespread, if not predominant, interpretation. This is illustrated by the fact that all economic historians, when interpreting their HDI estimation exercises, (a) make use of the classic tools for the standard value theory, such as marginal substitution tests between index components, marginal tests for changes in the index itself and the elasticities in relation to the components, and (b) assess their consistency with individual preference schemes considered to be more or less plausible.[2] The underlying ambiguity is, of course, to embrace the capability

[2] This point, discussed further below, was raised by Ravallion (2010), and then taken up again by Klugman et al (2011) and Ravallion (2011).

approach and then adopt the welfarist approach at the interpretative level. In this paper we will focus exclusively on this second interpretation and keep out of the picture all the problems related to the ability of the HDI index to capture the world of functionings and capabilities.

In the light of this discussion, the fundamental issue is whether the HDI is consistent with the standard value theory. Economic historians are well aware of the subjectivity of the weighting system embedded in any composite indicator (Brandolini and D'Alessio 2009, Decancq and Lugo 2013), but this is an old critique and has nothing to do, as we will show, with our main concern, namely the consequences of the paternalistic nature of the HDI. In the rest of the paper we follow Graaff (1957) and Yang (2018), and interpret the HDI as a *paternalistic* social welfare function. Paternalism is not a synonym of arbitrary trade-offs between the HDI dimensions: it means the absence of any analytical connection, in terms of the traditional value theory, between individual wellbeing and the HDI, that leads us to challenge the increasing hegemony of this and other, related composite indices as a way to carry out economic history analysis. The consequence is that the ethical system of the analyst is imposed on the reader and, more importantly, on the individuals of the society under scrutiny.

The second contribution of this paper, its *pars construens*, is to take seriously the implicit preferences of economic historians who rely on the HDI and its variations.

We bring those preferences to light by analysing marginal rates of substitution (MRS) for most of the HDI formulas adopted in economic history (Ravallion 2011). This is a tedious exercise, but MRSs reveal the implicit value judgments imposed by the researcher, and the conclusion is that those applied by economic historians so far differ greatly, and are sometimes difficult to accept. To generalize this point, we propose a specification of the HDI based on the constant elasticity of substitution (CES) utility (or production) function[3]. Our CES-HDI turns out to be a convenient tool for revealing ethical judgments hidden in composite indices formulas. In a CES-based HDI, the elasticity of substitution parameter determines the degree of substitutability among the socio-economic indicators that enter the definition of the index: by varying this parameter, we can a) reproduce all the HDI versions adopted by economic historians, and b) identify their ethical systems. The CES-HDI reveals what HDI hides.

If we interpret the HDI as a *cardinal* measure – as most economic historians seem to have done in the past – then we find that the choice of the substitution parameter fully determines the results: however subjective is the choice of the former, equally subjective are the latter. If, instead, we interpret the HDI as an *ordinal* measure – as

[3] Maasoumi (1986) argued, on the basis of economic and statistical criteria, in favour of the CES function as a flexible formula to aggregate multiple dimensions of wellbeing, but he referred explicitly to individual welfare and not to the aggregation of social indicators. Yang (2018) introduced the CES function referring explicitly to the HDI.

envisaged by its original proponents but embraced by few, if any, economic historians – then we find that the rank of the index is robustly identified only when its constituent elementary indicators are positively correlated (as is the case, typically, with income, education and health outcomes). In practice, economic historians have often been interested in using composite indices to resolve puzzling contradictions in the evidence, cases in which elementary indicators show diverging trends – think, for instance, of income and height during industrialisation (Engerman 1997, Crafts 1997a). We find that in the presence of negatively correlated indicators, composite indices fail to deliver robust results, even when interpreted as ordinal indices.

After setting out the theoretical framework for composite indices, we illustrate our arguments using post-unification Italy (1861 to the present day) as a case study. Between 1922 and 1943, Italians lived "under the axe" of the first fascist regime: indices of civil and political rights are negatively correlated with the HDI's components: income, longevity, and schooling (Amendola et al. 2017). This circumstance, combined with the abundance of high-quality, long-run estimates of several socio-economic indicators (Toniolo 2013, Vecchi 2011, 2017), provides us with all the necessary ingredients to test the inter-temporal properties of composite indices. Our conclusion is clear: the history of Italians' living standards, as identified by any composite index, is *necessarily* and *entirely* subjective. Whether the HDI is interpreted as an ordinal or cardinal index is irrelevant: results depend entirely on the choice of the substitution parameter, that is, on the analyst's preferences.

The HDI and economic history

Looking back at the literature of the 1960s and 1970s, one realizes that considerable effort has gone into replacing or supplementing GDP as an indicator of socio-economic development (Hicks and Streeten, 1979: 572). Some authors proposed to *adjust* the GDP, to account for the monetary value of aspects of human development it neglects, such as pollution or longevity. Usher (1973, 1980), for instance, proposed a method to adjust GDP per capita for longevity, while Williamson (1984: 162) proposed a revised version of this last index to correct for the endogeneity of secular improvements in mortality and income.

Other scholars chose an alternative route in devising composite indices that, instead of adjusting GDP, aggregated several indicators of development into a single number. Morris (1977) introduced the Physical Quality of Life Index (PQLI), an arithmetic mean of literacy, infant mortality and life expectancy at age one. He argued in favour of a 'historical PQLI', and estimated it back to the 19th century, for countries as diverse as Sri Lanka and the US (Morris 1979: 74). The PQLI attracted the interest of economic historians: to mention one case, Federico and Toniolo (1991) estimated decadal series from 1870 to 1910, comparing Italy to England, France and Belgium.

The PQLI was soon buried by the HDI, launched by the United Nations Development Program (UNDP) in 1990. In its initial formulation, the HDI was defined as a simple, arithmetic mean:

(1) $$HDI = \frac{1}{3} I_E + \frac{1}{3} I_S + \frac{1}{3} L_Y$$

where the terms on the right-hand side stand for life expectancy at birth (I_E), literacy (I_S), and income (L_Y). Contrary to the components of the PQLI, which can be easily expressed on a common 0 to 100 basis, the ingredients of the HDI must be made comparable. Each component x is normalized between 0 and 1 as follows: $I_x = (x_t - \underline{x})/(\overline{x} - \underline{x})$, where $\overline{x}$ and $\underline{x}$ stand for the maximum and minimum value of variable x, respectively. The notation L_Y signals that per capita GDP (y_t) is transformed by applying the natural logarithm to $y_t, \overline{y}$ and $\underline{y}$, to convey the idea of the diminishing returns of income for wellbeing[4].

The HDI found simultaneous applications on both shores of the Atlantic, in parallel debates on living standards and industrialisation – the so-called 'antebellum puzzle' in the U.S. (Komlos 2012), and the 'quality of life' in Industrial Revolution Britain (Voth 2003). In 1997, Richard Steckel and Roderick Floud edited an NBER volume, *Health and Welfare during Industrialization*, where a wide array of contributors expressed confidence in the fact that composite indicators would strengthen our

[4] Initially UNDP (1990, 12-13) adopted an 'average official poverty line income' for nine industrial countries as $\overline{y}$, but later reports increased this threshold up to $75,000 (UNDP 2016, 198). The concave transformation has also been expressed in different forms (see UNDP 1994: 108). As noted by Kelley (1991: 317) the choice of any specific transformation is 'necessarily arbitrary'.

understanding and interpretation of history[5]. Engerman (1997: 33) emphasized the advantages offered by these new (composite) indices of welfare, including the PQLI, the HDI, and the Dasgupta-Weale Index (DWI), an index including the dimensions of political and civil rights. Other authors, while joining the enterprise, expressed awareness of the analytical limitations of these tools. Costa and Steckel (1997: 71), for example, stressed that 'of particular concern in economic history is the choice of indicators and the selection of maximum and minimum values'; similarly, in the monograph's epilogue Steckel and Floud (1997: 437) shared their concern about the use of composite indices - a 'debatable method' in their words.

In the same year, 1997, in London, Nicholas Crafts led the way to the use of the HDI, in his investigation on the British 'quality of life' debate. Crafts (1997a) estimated both the HDI and 'its most ambitious cousin', the DWI, for six benchmark years between 1760 and 1850.[6] He showed how the steady increase in composite indices

[5] Retrospectively, it is of great significance that the adoption of composite indices in economic history was pioneered by the same authors that had established anthropometric measures in the discipline. As noted by Komlos (2009: 342), both "the Human Development Index and the Biological Standard of Living attempt to quantify the quality of life as a complex, multidimensional entity." Compared to anthropometric indicators, composite indices have the advantage of a clear, specified formula, and of a well-defined list of components; on the other hand, they have the disadvantage of lacking an objective bedrock comparable to the biological relations on which the adoption of height and other anthropometric indicators is based.

[6] Crafts used a slight variation of equation 1, a formula first proposed in UNDP (1994).

was at odds with the gloomy evidence based on height. The comparison between Britain's HDI- and DWI-based achievements at 1860 with those of eleven other countries, called into question Britain's 'leadership' based on GDP: inter-country rankings differed markedly depending on the measure used. Crafts concluded that composite indices could 'be even more important for economic historians than for contemporary development economists' (p. 634). In a second paper, Crafts (1997b) produced estimates of the HDI for 16 industrialized countries, arguing that HDIs offered 'a new angle on comparisons of economic progress in different economic eras', new with respect to those based on real GDP per person (p. 301). [7]

Two decades later, while it failed to replace GDP, one can probably claim that the HDI has become a standard welfare measure in economic history. It has earned a place in the historiography of every continent – Astorga et al. (2005) constructed a series for Latin America, Prados de la Escosura (2013) for Africa – and in authoritative textbooks (Broadberry and O'Rourke 2010, Persson 2010).

[7] The index was proposed with two alternatives: together with the one adopted in Crafts (1997a), a second HDI was computed normalising income, I_Y, along the same lines as the non-monetary components, with $\overline{y}$ set at the US per capita GDP in 1992.

Since its first appearance, the HDI has undergone several modifications, often in response to criticism from academia.[8] The main reason for dissatisfaction with the original formula in equation (1) is that it implicitly assumes *perfect substitutability* between arguments. In equation (1), one year less of life expectancy is perfectly compensated by an increase of equal magnitude in the schooling index (Desai, 1991)[9]. Paradoxically, the human development of a modern industrialized economy may be made equivalent to the degree of human development of a population with a zero life expectancy, as long as its citizens – who have not even had the time to go beyond the cradle – are sufficiently educated or wealthy. Perfect substitutability is incompatible with the idea that the components of the index are *essential* dimensions of wellbeing: by definition, that which is essential cannot be replaced (Sagar and Najam, 1998). The latest and most important revision of the HDI was carried out for its twentieth anniversary (UNDP 2010). On that occasion, it was decided to change equation 1 by introducing a geometric mean instead of the simple arithmetic mean:

$$HDI = I_E^{\frac{1}{3}} \cdot I_S^{\frac{1}{3}} \cdot L_Y^{\frac{1}{3}} \quad (2)$$

[8] See, among others, Srinivasan (1994), Ravallion (1997, 2011, 2012a), and Klugman et al. (2011). Kovacevic (2010) reviews the first twenty years of the debate, instrumental to the 2010 revision. See also Hirai (2017).

[9] This is clearly not true, due to the log transformation, for the degree of substitutability between GDP and the other components of the HDI.

At the same time, GDP was replaced by Gross National Income (GNI), and literacy by average years of schooling. The change in the functional form of the HDI was promptly accepted by economic historians, while the change in the indicators was not: the 'old' indicators (literacy, GDP, and life expectancy) have been maintained in preference to the new. Following Gidwitz et al. (2010), equation (2) is probably the most popular version currently in use among economic historians.

A notable exception is Leandro Prados de la Escosura. He took up an idea of the Indian development economist Nanak Kakwani, according to which development indicators should reflect the greater difficulty of improvement 'as the standard of living reaches progressively higher limits' (Kakwani 1993: 308). To implement this idea, Kakwani proposed a convex transformation of social indicators: $F_x = \left[ln\left(\overline{x} - \underline{x}\right) - ln(\overline{x} - x_t)\right] / ln\left(\overline{x} - \underline{x}\right)$[10]. Applying Kakwani's transformation F_x to the non-monetary components of HDI, Prados de la Escosura (2015a) proposed a new HDI:

$$HIHD = {F_E}^{\frac{1}{3}} \cdot {F_S}^{\frac{1}{3}} \cdot {L_Y}^{\frac{1}{3}} \quad (3)$$

[10] More precisely, Kakwani (1993) introduced the following parametric "achievement function": $\left[\left(\overline{x} - \underline{x}\right)^{\varepsilon} - (\overline{x} - x_t)^{\varepsilon}\right] / \left(\overline{x} - \underline{x}\right)^{\varepsilon}$. The limit of a linear transformation of this function, as ε approaches 0, coincides with the expression in the text and is fully consistent with the axiomatic structure proposed by the author.

Equation (3) attributes increasing value to marginal increases of life expectancy and education, often costlier to achieve, but maintains the traditional decreasing marginal benefits of income, guaranteed by L_Y.

Concerned by the arbitrariness of both convex and concave transformations, Amendola et al. (2017) proposed a new version of the HDI that aggregates the three components using a geometric mean, but normalizes all of them symmetrically with the linear transformation I_x, as in Crafts (1997b):

$$HDI_{AGV} = I_E^{\frac{1}{3}} \cdot I_S^{\frac{1}{3}} \cdot I_Y^{\frac{1}{3}} \quad (4)$$

The index HDI_{AGV} does not escape the interpretative limitations of all other HDIs, but does have the virtue of greater simplicity (due to the absence of any non-linear transformation in the original variables), in addition to another hidden virtue that will be brought to light in Table 1, when discussing the marginal rates of substitution built into any HDI. The different formulas of the HDI reviewed in this section are assessed in the next, in the context of the history of Italian living standards.

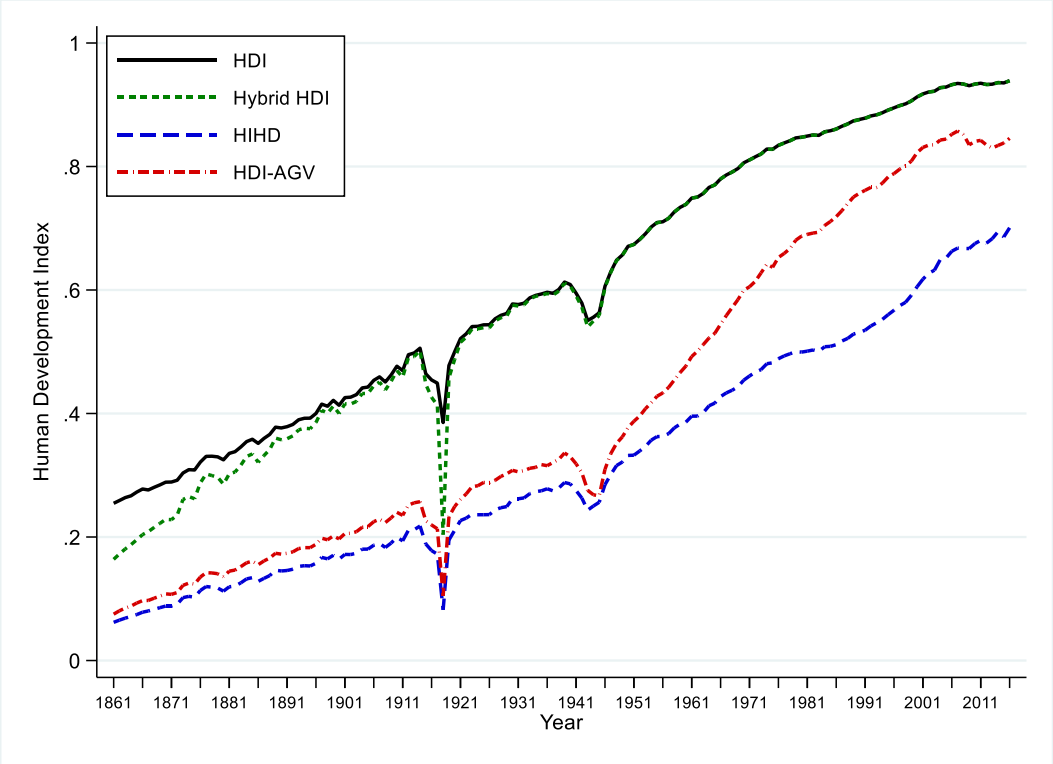
Italian living standards through the lenses of the HDIs

Italian economic historians did not miss the opportunity to use the HDI, with the aim of achieving a better understanding of the long-run dynamics of welfare in the country. Alternative estimates of the HDI – both at the national and the regional level – appeared in the 2000s: Conte, Della Torre, and Vasta (2001), Felice (2007), and then Brandolini and Vecchi (2013) produced new time series of the HDI from Italy's

unification (1861) to the present day. More recently, Felice and Vasta (2014) constructed new regional estimates in support of their investigation of the roots of the *Southern Question*, while Amendola et al. (2017) focused on the difficulties of adopting the HDI in interpreting Italian development. The fact that different authors used different versions of the HDI suggests two questions. First, do different aggregation rules lead to different results? Second, did the HDI succeed in identifying the welfare trend of the Italian population during the last 150 years? Our answer to the first question is positive, while the answer to the second question is, by and large, negative. We conclude arguing that economic historians need a unified conceptual framework before using the HDI and other composite measures – a task that we take on in the next section.

In Figure 1, adopting the new evidence on literacy, longevity and per capita GDP (Vecchi 2017), we calculate four alternative HDIs for Italy: the 'old' HDI adopted by Crafts (1997) (equation 1), the 'hybrid' HDI as in Felice and Vasta (2014) (equation 2), the HIHD devised by Prados de la Escosura (2015a) (Equation 3), and the new-born HDI of Amendola et al. (2017) (equation 4). Even a quick glance reveals that the series differ greatly – not only in levels, but also in trends. How should we interpret these differences? Which index is the most 'correct'?

Figure 1. Human Development Indices in Italy, 1861-2016

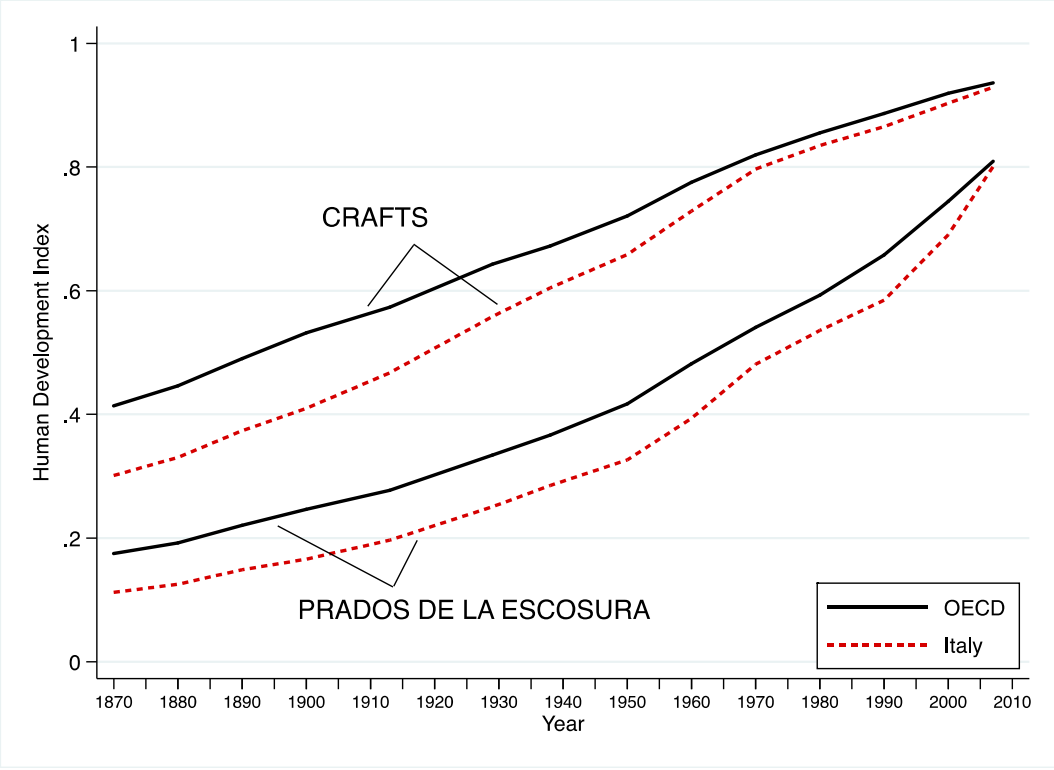


Source: Authors' calculations based on Amendola et al. (2017).

Consider the difference between the HIHD (blue, dashed line) and Amendola et al.'s HDI (red, broken line). They proceed hand in hand until World War II, when they diverge in both trend and level. The reason is the different transformations of the components: all possible transformations are at work here, linear, concave, and convex transformations (I_x, L_Y and F_x, respectively). Within the Prados de la Escosura index, in fact, monetary and non-monetary dimensions are treated in opposite ways. While an increase of average income is 'discounted' by the logarithm, as in the 'old' HDI, improvements in education and longevity are magnified by the

convex transformation F_x. Both assumptions might be reasonable in historical terms, but they have to be recognized as *discretionary* choices. The analyst's personal judgments are embedded into all the formulas in Figure 1, and therefore hidden from view, so that they cannot be assessed by the reader. No unambiguously best index exists; subjective judgment is the only criterion available in deciding which story one should send to the printer.

Figure 2. Italy and the OECD compared, 1870 - 2010



Source: Authors' calculations based on Amendola et al. (2017: 465).

Figure 2 elaborates on the consequences of this last argument, by comparing Italy's performance (dashed lines) to the OECD average (solid lines). The idea is to show the implications at the interpretative level that can result from using one formula rather than another. As we can see, the choice of formula changes not just the level of the index (Prados de la Escosura's HIHD geometric mean systematically shifts index values down compared to the arithmetic mean of Craft's HDI), but also changes its dynamics. Regarding *levels*, Crafts (1997b) noted that the 1870 values of the OECD countries were comparable with those of developing countries today (around 0.4). In contrast, using a geometric formula, the OECD countries of 1870 would fare worse than the Central African Republic (0.352), the country with the lowest HDI in UNDP (2016). Turning to the *dynamics*, according to the old HDI, Italy converged with the OECD throughout most of the last century, with an acceleration during the economic miracle of the 1950s, achieving near parity by the 1970s. If we use the HIHD, instead, the lesson to be drawn is the exact opposite: Italy failed to converge for almost 130 years, and managed to recover and align with OECD standards only in the last twenty years.[11] Alternative HDI formulas do not reflect more or less innovative 'technologies', that is amendments to the original formulation; they correspond to different sets preferences or ethical systems. The choice between the old HDI and the HIHD is a choice between different

[11] Similarly, with reference to of European living standards during the interwar period, Gallardo Albarrán (2018: 2-3) noted that HIHD overturns the optimistic interpretations of Millward and Baten (2010) based on the old HDI.

understandings of living standards – those of Crafts, Amendola and coauthors, Felice and Vasta, or Prados de la Escosura.

Figures 1 and 2 offer an opportunity to discuss a second issue of great importance. Graphs comparing the HDI series for various countries over time have been created, typically, in order to study growth rates and convergence patterns between the various areas.[12] This interpretation is clearly based on the *numerical value* of the index. According to the creators of the HDI, however, this is inappropriate. The HDI was created explicitly to *rank* the *relative* performances of the various countries *at a given moment* in time (UNDP 1993: 110). Anand and Sen (1994: 8) did not overlook the possibility of constructing a historical series of the index but came to the conclusion that 'no special significance is attached to the absolute value of the index, the entire analysis being conducted in terms of the ranking of countries relative to one another'.

[12] Crafts (1997a: 310) used the HDI to measure "the speed of development in different eras"; Boyer (2007) in a discussion on 'convergence of living standards'; Prados de la Escosura (2010: 850) asked "whether the human development gap between the 'core' and the 'periphery' deepened over time" and, eventually, commented on the "absolute gap" and the "rate of variations" (Prados de la Escosura 2014: 12); for Astorga et al. (2005: 775) the wellbeing of Latin Americans "almost doubled between 1900 and 1939, and more than doubled to 1980"; according to Millward and Baten (2010: 253) "HDI showed signs of convergence within Europe during the interwar period"; finally, for Baines et al. (2010: 399) "the average HDI score for Europe rose by almost 30 percent between 1950 and 2003".

To be consistent with the purposes of the original creators of the HDI, the index should thus be interpreted as a purely *ordinal* index, which can be used to create a ranking.[13] This consistency requirement poses, however, serious limits to its use in economic history. When looking at the progress made by a country over time, as in Figures 1 and 2, the HDI – in any of its guises – tells us that the 1950s were better than the interwar years, which were mostly better than the 19th century; but we cannot tell by how much, nor compare the different speeds at which the HDI increased; the very definition of a growth rate is based on a cardinal interpretation of the HDI. For most of their history as a unified country, Italians have improved their standards of living. Although not completely mundane as a historical insight, this is unfortunately

[13] In the economic history literature, we found explicit acknowledgment of this fact only in Baines et al. (2010, 399), according to whom the HDI is a 'relative', or 'comparative' measure of development. As mentioned in the previous footnote, however, they also consider its absolute variations. Herrero et al. (2012, 250) note that the use of the geometric mean gives the index "a true cardinal dimension and then allows performing comparisons on how much the human development has changed". In fact, they show that an HDI based on the geometric mean satisfies the "ratio consistency axiom" which means that the relative values of the indices of two social states does not depend on the value of a common component. They do not exclude, however, that other aggregation rules may preserve the same ordering of social states implied by the geometric mean-based HDI, and this is what really matters for our argument.

all that we can say.[14] Turning to international comparisons, according to Figure 2, we can simply conclude that Italians have always experienced a lower degree of wellbeing, on average, than their contemporaries in northern Europe or in North America. We cannot say by *how much*: when indices are ordinal ones, the vertical distance between the two curves in Figures 1 and 2 has no meaningful interpretation. The cardinal interpretation of the HDI (see Footnote 12), is by far the dominant one in economic history – while not intrinsically incorrect, it testifies to the distance between its practical application in economic historical analysis, and its intellectual origins in Sen's capability approach.

Historical HDI as a paternalistic social welfare function

In this section, we put composite indices, and in particular the HDI, into a more general conceptual framework. We know, from Arrow's impossibility theorem (Arrow, 1951), that there are no democratic decision rules that can be used to aggregate individual preferences and that satisfy a minimal set of consistency criteria. As observed by Sen (1999), dictatorship would avoid inconsistencies in social ordering systems. In this regard, the HDI is not an exception. Roughly speaking, Arrow rules out the possibility that economic historians could come up

[14] More interesting, from the interpretative side, is the exercise of decomposing the HDI into contributions made by each of the three components – life expectancy, education and income, see Amendola et al. (2017).

with an HDI-based ordering of two societies by aggregating individual preferences, unless they use subjective judgment, that is, unless they play dictator. A simple method to produce social rankings is via Bergson-Samuelson social welfare functions (BSWF), which allow the analyst to rank social alternatives, according to a dictatorial criterion and starting from individual orderings. We suggest following this method, as BSFWs allow us to appreciate the details underlying a history of welfare based on the HDI.

Assume that individual preferences are defined over the three variables – longevity, schooling, and income – and that they can be described by a standard utility function:

$$u_i = u_i(e_i, s_i, y_i)$$

where the index i refers to individuals, and e_i, s_i and y_i denote life expectancy, educational attainment and income, respectively. A BSWF can be defined as follows:

$$W = W\big(u_1(e_1, s_1, y_1), u_2(e_2, s_2, y_2), \dots\dots, u_n(e_n, s_n, y_n)\big) \quad (5)$$

where W (·) is a real valued function that maps individual utilities into real numbers. The shape of the function W (·) reflects the ethical system of the dictator, and is defined *independently* of individual preferences.

One might be tempted to interpret the HDI as a BSWF. In this perspective, the choice of a specific functional form of the HDI would reflect the ethical system of the economic historian. This is not the case. The degree of arbitrariness imposed by the HDI is stronger than that implied by a proper BSWF. To see this, let us write the HDI as follows:

$$HDI = HDI(E, S, Y)$$

where *E* stands for life expectancy, *S* for schooling, and *Y* for income (GDP). This shows that the index does not depend on the wellbeing of individuals $u_i(e_i, s_i, y_i)$, but rather on aggregate indicators (*E, S, Y*), which in turn can be assumed to depend on individual-level variables: $E = E(e_1, e_2, \dots, e_n)$, $S = S(s_1, s_2, \dots, s_n)$, and $Y = Y(y_1, y_2, \dots, y_n)$. The HDI can therefore be written as:

$$HDI = HDI\big(E(e_1, e_2, \dots, e_n), S(s_1, s_2, \dots, s_n), Y(y_1, y_2, \dots, y_n)\big) \quad (6)$$

Equation (6) clarifies that the HDI does *not* qualify as a SWF, the former being defined on social indicators, the latter on individual preferences. Both equations (5) and (6) ultimately depend on individual levels of longevity, but the definition of an SWF requires that each and every individual in the society comes up with an ordering of possible outcomes based on *e*, *s* and *y*, and then individual orderings are aggregated by the SWF.[15] In contrast, equation (6) does not require any individual

[15] One way to make (5) and (6) equivalent is to assume linear individual preferences and an additive $W(\cdot)$. This is, basically, a special case of a Benthamite (or cardinal) social welfare function. On the difference between aggregation at the individual vs. country level, see Atkinson et al. (2002) and Brandolini (2009). Another option is to adopt a "symmetric geometric aggregation rule", for individual preference and over aggregated variables, to exploit the Foster and Shneyerov (2000) path independence property. In this case the HDI in eq (6) would not depend on the fact that that aggregation is first carried out over individuals or over variables. However, this "symmetry" restriction on the aggregation rule is equivalent to impose to individuals the analyst's own preferences, a more sophisticated way to say that individual preferences do not matter.

ordering. This difference explains why, following Graaff (1957), we can interpret the HDI as a *paternalistic* social welfare function (PSWF) – an argument recently also advanced by Yang (2018).[16] Here the role of the *pater* is played by the economic historian, who arbitrarily chooses not only the shape of the social welfare function, but also the implicit system of individual preferences[17].

The fact that the HDI can be interpreted as a PSWF is crucial to our argument, as it formally establishes that the economic historian's HDI-based welfare ordering is entirely dictated by their own preferences. Once we have proven that HDI is a PSWF, then to say that the HDI is higher in A than in B is no more to say that the economic historian would choose A rather than B, if allowed to make the choice (Graaff, 1957: 5). Irrespective of its specific formulation, the HDI is *not* an objective measure of human development.

Even so, an interesting question remains to be answered: what are the defining characteristics of these ethical systems? In what ways does Prados de la Escosura's ethical system differ from Crafts', or from Felice and Vasta's? The marginal rates of substitution (MRS) provide a possible answer (Brandolini 2009). In the HDI context,

[16] Technically, the HDI is a social welfare function that does not satisfy the *nonpaternalism* property, which prescribes that "(…) in the expression of social preferences only the individual preferences matter. The planner does not have direct preferences on the final alternatives" (Mas-Colell et al. 1995: 825).

[17] In principle, we do not even know if such an implicit preference system exists.

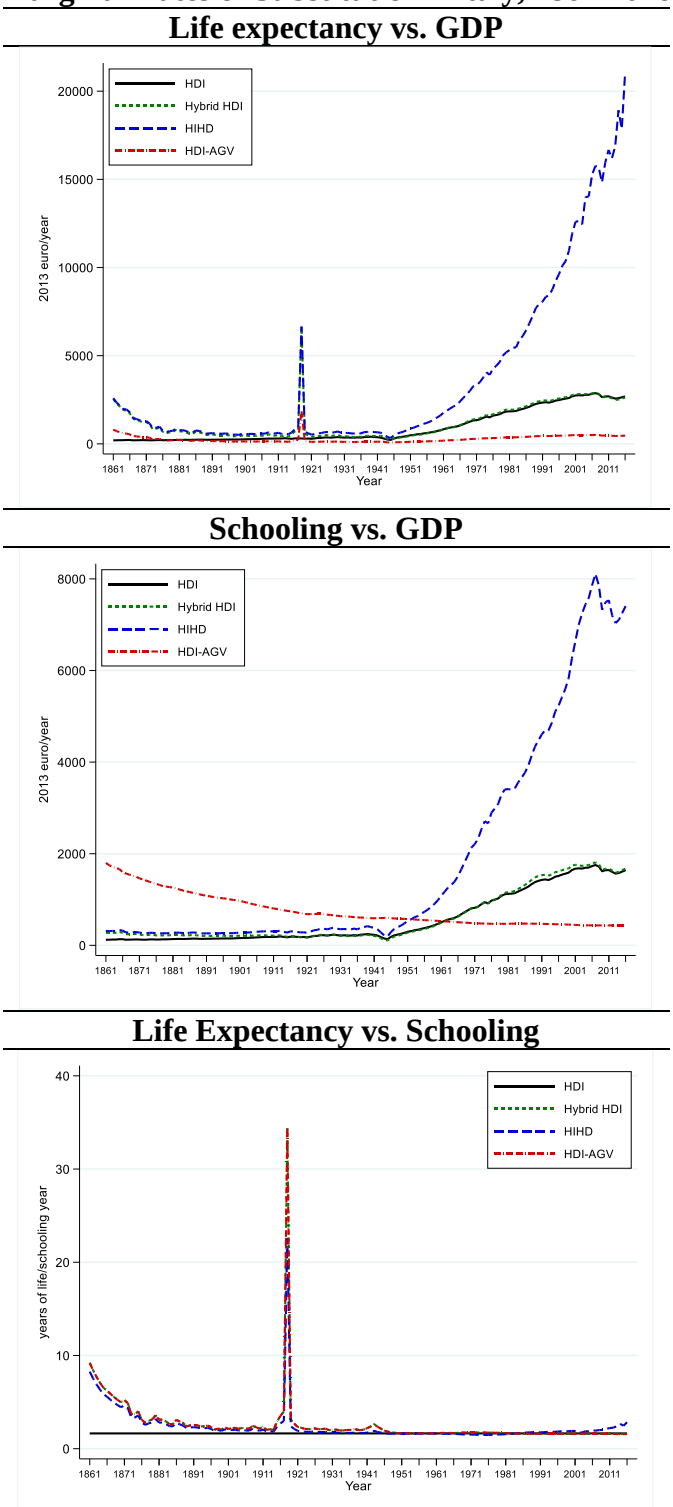
the MRS of life expectancy E_t with respect to per capita GDP Y_t is the amount of dollars that one (the economic historian) has to give up when increasing life expectancy by one year, in order to keep the HDI unchanged. To all intents and purposes, the MRS is the 'exchange rate' or the relative importance of the population's average life expectancy compared to average income. Table 1 shows the MRS implied in each of the formulas proposed for the HDI (equations 1 - 4).

Table 1. Marginal rates of substitution (MRS) for different HDI specifications

	GDP vs. Life expectancy $\lvert MRS_{e_t/y_t} \rvert$	**GDP vs. Schooling** $\lvert MRS_{s_t/y_t} \rvert$	**Life expectancy vs. Schooling** $\lvert MRS_{e_t/I_t} \rvert$
HDI (eq. 1)	$\frac{y_t(lny - lny_0)}{e - e_0}$	$\frac{y_t(lny - lny_0)}{s - s_0}$	$\frac{s - s_0}{e - e_0}$
Hybrid HDI (eq. 2)	$\frac{y_t(lny_t - lny_0)}{e_t - e_0}$	$\frac{y_t(lny_t - lny_0)}{s_t - s_0}$	$\frac{(s - s_0)}{(e_t - e_0)}$
HIHD (eq. 3)	$\frac{y_t(lny_t - lny_0)}{(e - e_t)(\ln(e - e_0) - \ln(e - e_t))}$	$\frac{y_t(lny_t - lny_0)}{(s - s_t)(\ln(s - s_0) - \ln(s - s_t))}$	$\frac{(s - s_t)[\ln(s - s_0) - \ln(s - s_0)]}{(e - e_t)[\ln(e - e_0) - \ln(e - e_t)]}$
$\boldsymbol{HDI_{AGV}}$ **(eq. 4)**	$\frac{(y_t - y_0)}{(e_t - e_0)}$	$\frac{(y_t - y_0)}{(s_t - s_0)}$	$\frac{(s_t - s_0)}{(e_t - e_0)}$

Faced with different HDI formulations, the MRSs facilitate the understanding of which is more consistent with our preferences. Moreover, since the MRS is a relative magnitude, a sort of relative price, it means that it has a *cardinal* interpretation even if one does not share a cardinal interpretation of the HDI: it makes perfect sense to carry out comparisons of different MRSs both in levels and in changes over time. Figure 3 shows the development over time of the MRS for life expectancy with respect to per capita GDP (top panel), for schooling and per capita GDP (panel in the middle), and for life expectancy versus schooling (bottom panel) in Italy (1861-2016).

Figure 3. HDI Marginal Rates of Substitution – Italy, 1861-2016



Source: Authors' calculations.

To begin with, let us consider the index proposed by Amendola et al. (2017) (red dashed line, equation 4). According to this specific HDI, in 1861 one extra year of life would be worth around $400 (in 2011 Geary-Khamis dollars). The top panel in Figure 3 shows that for as much as a century and a half after unification, the MRS remains close to this value. The indices used by other authors show a very different trend. The index calculated by Crafts starts from a lower MRS (circa $200), while the indices put forward by Felice and Vasta and by Prados de la Escosura start from a much higher $1,800 and shoot up in the aftermath of World War II. Crafts' formulation assumes that in the 2010s a year of life is worth around $2,600 – a value not too far removed from the one assumed by Felice and Vasta (about $2,500). The index created by Prados de la Escosura exceeds all of these: from 1861 the value of one additional year of life rises from $1,894 to $17,641. These discrepancies stem from using a geometric mean. The point was well grasped by Ravallion (2012a), on the occasion of the introduction of the new HDI formulation: the HDI 'puts a higher value to an extra year of life for people in rich countries than poor ones', with the 'unacceptable implication that rich people, or residents of rich nations, are worth more than the poor' (p. 206). Ravallion's observation naturally applies to inter-period comparisons or to economic history. As the above calculations show, different HDIs attribute very different weights to life expectancy in different periods of Italian history: this evaluation may be legitimate, but it should be made transparent.

A ‘quasi folk theorem’ for historical HDIs

Comparison of the marginal rates of substitution between the components of the HDI illuminates differences in the ethical systems implicit in the various formulations of the HDI proposed in the literature. We can make a more systematic comparison by allowing for smooth variation in the degree of substitutability between the index components. We start by reformulating the HDI as a constant elasticity of substitution function[18] (CES). This allows us to vary the degree of substitutability between the HDI components through a single parameter: the elasticity of substitution σ.[19] Maintaining the same notation as the previous sections, the generalized HDI can be written as follows:

$$HDI_{CES} = \left[\alpha_1 (E_t)^{\frac{\sigma-1}{\sigma}} + \alpha_2 (S_t)^{\frac{\sigma-1}{\sigma}} + \alpha_3 (Y_t)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \quad (7)$$

[18] Mas-Colell et al. (1995) introduced the CES function as an instance of a *generalized utilitarian social welfare function* (see example 22.c.4 p. 828-29). However, in the present context, equation 7 generalizes a SWF that violates the *non-paternalism* property, i.e. a PSWF.

[19] One might be tempted to further generalize the HDI formula by allowing for different elasticities of substitution among the arguments of the HDI. Unfortunately, Uzawa (1962) proved that if the elasticities of substitution between every pair of arguments are to be held constant, then the elasticities must be identical.

where the parameter α_i represents the weight attributed to component *i* of the HDI, $\alpha_i \geq 0$ for every *i* and $\sum \alpha_i = 1$. In what follows we assume that $\alpha_i = \alpha_j \; \forall \; i, j$, that is, we treat symmetrically all the arguments of the HDI.

In equation (7), the parameter σ plays a crucial role. As σ approaches infinity, the components become perfect substitutes and we obtain the Crafts (1997) formulation of the HDI (equation 1). As σ approaches 1, HDI_{CES} converges to the *hybrid* HDI proposed by Felice and Vasta (2014) (equation 2) and, with slight modifications, to the HIHD by Prados de la Escosura (2015a) (equation 3). When σ approaches 0, then $HDI_{CES} = min\{E_t, S_t, Y_t\}$, that is, we obtain a new Leontief-like HDI specification, never considered in the literature to the best of our knowledge. In this case the components of the HDI are perfect complements, a characteristic that fits well with the idea that they capture the essential dimensions of wellbeing. A number of intermediate cases can be obtained by varying σ between 0 and infinity[20].

[20] Zambrano (2014, 2017) follows an axiomatic approach to characterize different specifications of the HDI. He finds that, i) in contrast to multiplicative forms of the HDI, the additive HDI does not satisfy either the "subsistence axiom", or the "capabilities growth independence axiom", a stronger form of independence with respect to the "independence axiom" satisfied by the multiplicative HDI. All these axioms can be associated to different degrees of substitutability and are consistent with the role played by the parameter σ in equation (8). An advantage of our approach is that the parameter σ is a continuous variable with an intuitive interpretation, a drawback is that we need to impose the CES functional form for the HDI.

Unfortunately, equation 7, even if simple and of immediate interpretation, is not sufficiently flexible to encompass *exactly* all the historical HDIs proposed in the literature. This is because these HDIs introduce specific 'achievement functions', i.e. transformations of the original, elementary indicators. A general CES formulation of the HDI, able to encompass all the cases explored in the literature, must include additional parameters that model the achievement functions. The following general HDI_{CES} does this:

$$HDI_{CES}(\sigma,\varepsilon,\eta)=\left[\sum_{i=1}^{K}\alpha_i\left[\frac{\left(\bar{x}^i-\underline{x}^i\right)^{\varepsilon}-\left(\bar{x}^i-x_t^i\right)^{\varepsilon}}{\left(\bar{x}^i-\underline{x}^i\right)^{\varepsilon}-1}\cdot a_i\right]^{\frac{\sigma-1}{\sigma}}\right. \\ \left.+\left(1-\sum_{i=1}^{K}\alpha_i\right)\left[\frac{(y_t)^{\eta}-\left(\underline{y}\right)^{\eta}}{(\bar{y})^{\eta}-\left(\underline{y}\right)^{\eta}}\right]^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{8}$$

Notation in equation 8 is as follows: x_t^i is the value of the *i*-th indicator (*e.g.*, life expectancy at birth) in year *t*; $\bar{x}^i$ and $\underline{x}^i$ are the theoretical, time-invariant maximum and minimum values, respectively, of the *i*-th indicator. Some authors refer to these boundaries as *goalposts*; y_t denotes real per capita GDP. The parameter $\varepsilon \in (0,1]$ is an important one; it describes the degree of convexity of the achievement functions for the *K* components other than GDP. Finally, the parameter $\eta \in (0,1]$ regulates the

degree of concavity of the achievement function for GDP[21]. In short, equation 8 identifies a class of HDIs characterized by the parameter vector $(\sigma, \varepsilon, \eta)$.

Table 2 below shows how the main historical HDIs proposed in the literature can be obtained from equation (8) by choosing the parameter set $(\sigma, \varepsilon, \eta)$ appropriately. The first four rows refer to the indices described in the previous sections. The basic HDI formula assumes perfect substitutability among the HDI components, and the achievement functions are linear with the exception of the GDP component. This implies $\sigma = \infty$, $\varepsilon = 1$ and $\eta = 0$. The hybrid HDI is based on the same assumptions about the achievement functions but allows for Cobb-Douglas type imperfect substitutability among the components, corresponding to $\sigma = 1$. The HIHD introduces a variation on the achievement functions for the non-GDP components by assuming, other things being equal, $\varepsilon = 0$. The HDI_{AGV} index assumes linear achievement functions for all the components, which means $\varepsilon = \eta = 1$. The fifth row identifies a case that, to the best of our knowledge, has not been explored in the literature but deserves particular attention. This case assumes the complete absence of substitutability, which means $\sigma = 0$, with all the possible combinations of the parameters ε and η. In this case, the HDI, as observed above, takes the shape of a Leontief type social welfare function that captures, in a utilitarian framework, the ethical system proposed by Rawls (1971). All the social indicators included in the

[21] The adjustment coefficients $\alpha_i = (\bar{x}^i - \underline{x}^i - 1)/(\bar{x}^i - \underline{x}^i)$ are necessary to guarantee the exact convergence of the $HDI_{CES}(\sigma, \varepsilon, \eta)$ with the HDIs proposed in the literature.

HDI are supposed to be essential dimensions of wellbeing for which a trade-off cannot be defined. As a consequence, the pattern of the HDI is entirely driven by the most deficient dimension of wellbeing. The last rows capture the residual possibilities characterized by imperfect substitutability and different degrees of concavity/convexity of the achievement functions.

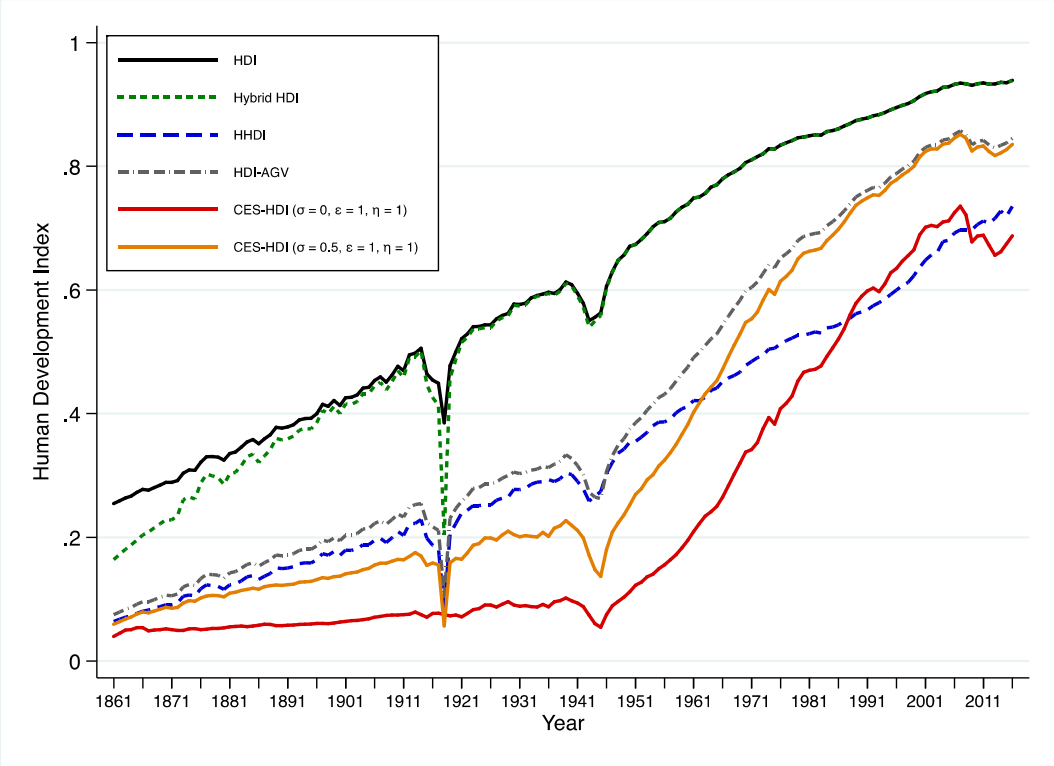
Table 2. A taxonomy of indices within the generalized historical HDI

Denomination	σ	ε	η	Preferences	Notes
HDI (eq. 1)	∞	1	0	almost perfect substitutes	(a)
Hyrid HDI(eq. 2)	1	1	0		(b)
HIHD (eq. 3)	1	0	0		(c)
HDI_{AGV} **(eq. 4)**	1	1	1	Cobb Douglas	(d)
Rawlsian HDI	0	[0,1]	[0,1]	perfect complements	
Imperfect substitute	$(0,\infty) \setminus 1$	[0,1]	[0,1]		
Perfect substitute	∞	1	1	linear preferences	

Notes: (a) eq. 1, United Nations (1990), (b) eq. 2, Felice and Vasta (2014), (c) eq. 3, Prados de la Escosura (2015a), (d) eq. 4, Amendola et al (2017).

Figure 4 below plots the HDI_{CES} series that we obtain by varying the elasticity of substitution σ and the parameters η and ε. The figure contains all the HDIs previously illustrated in Figure 1, but also a subset of other indices based on the 'Rawlsian' or 'imperfect substitute' parameterization. It is evident that a fairly wide range of patterns results from this exercise.

Figure 4. The many facets of the CES-HDI in Italy, 1861-2017



Source: Authors' calculations.

The results in Figure 4 suggests an analogy with game theory's 'folk theorem'. Within the family of human development indices encompassed in equation (8), it is not difficult to find an ethical system, summarized by the parameters $(\sigma, \varepsilon, \eta)$, which can generate any (arbitrary) trajectory for the HDI, for the same underlying elemental indicators. This is not true for *all* possible trajectories, but it is possible for a consistent subset of all the possible trajectories. We are not reasoning in a repeated game analytical setting, of course, but the message is similar: under fairly general conditions, 'anything goes'. That is, given a suitable choice of the parameters $(\sigma, \varepsilon, \eta)$ almost any HDI-based story can be told. Note that this result does *not* depend on the weights (α_i) used for the components of the HDI, but on their

substitutability: the parameter σ regulates, symmetrically for all components, the degree of substitutability that, in turn, identifies the analyst's subjective *exchange ratios* among them.

Thus, the parameter σ reveals the analyst's preferences. The fact that σ can take on, at least theoretically, any value greater than zero implies a fundamental 'indeterminacy' in the HDI, a feature that deserves a comment. The negative side of HDI indeterminacy is that the HDI seems to be unable to deliver a history of human development based on transparent and non-discretionary criteria. There are many, virtually infinite, stories that can be told according to the ethical judgment of those who build the HDI index. The positive side is that all these stories can be supported on the basis of a vector in the $(\sigma, \varepsilon, \eta)$ parameter space.

We used the expression '*quasi* folk theorem' in the title of this section, where the qualifier 'quasi' reminds us that the validity of the analogy is limited by the fact that the standard components of the HDI are, for most countries, highly positively correlated. This is a well-documented empirical regularity. Many authors have pointed out the redundant nature of the additional information incorporated in the human development index with respect to the single GDP component[22]. An

[22] See Stewart (1985), McGillivray (1991), Srinivasan (1994), Felipe and Resende (1996: 187-190), and Diener and Suh (1997: 192-200), Brandolini (2009). Ogwang (1994) and Ogwang and Abdou (2003), adopting a principal components analysis, showed that using a single indicator (life expectancy) would result in almost the same country ranking, 'without loss of too much information'.

immediate consequence of the fact that life expectancy at birth and education co-move with per capita GDP is that, irrespective of the elasticity of substitution, most of the time, and most of all, in the long run we are *not* able to reverse the trend of the historical HDI. Surely, the different degree of substitution embodied in the formula brings down the HDI in years characterized by sharp GDP contractions – the 1930s, and most notably, the decade of stagnation inaugurated by the 2008 recession, the most severe in the history of Italy (Baffigi, 2015), but also the shorter-term falls after the oil shock of 1973 and the currency crisis of 1992. Excluding these exceptions, and considering the long-run correlation between the dimensions of the HDI, interpreting it exclusively as an ordinal index amounts to saying that the 'quasi-folk theorem' does not apply, as the ranking by years is unequivocally determined by the index and is independent of σ.

A possible rehabilitation of our quasi folk theorem relies on the introduction of new indicators that are not so closely correlated with per capita GDP. This would not only enrich the concept of wellbeing inherent in the HDI but would also enlarge, as we will see in the next section, the number of conceivable trajectories of the HDI that can be sustained by an appropriate ethical system. A strong candidate to extend the dimensions of wellbeing considered in the historical HDI is political and civil rights, a strategy already explored by Dasgupta and Weale (1992), Crafts (1997) and others. In the next section we explore this line of research, which seems particularly suitable

for the case of Italy, a country where, over extended periods, political and civil rights followed an eccentric pattern with respect to the other dimensions of wellbeing included in the HDI.

Beyond the HDI

It comes as no surprise that economic historians consider political and civil rights crucial in the intertemporal assessment of living standards. The creators of the HDI argued that 'human development is incomplete without human freedom' (UNDP 1990: 15), but they ended up limiting the aggregation to education. The choice was not an accident: it was put forward by Paul Streeten, a distinguished development economist, but also a refugee from Austria who had fought in Sicily against Nazi-Fascist forces (Haq 1995: 61). Streeten (1994: 236) argued in favour of measuring and including political rights: they are 'so important' – he wrote – 'that no trade-off should be possible'. On the other hand, the volatility of such indices and the inevitable subjectivity in quantifying such indicators represent a weakness. Crafts (1997a: 621-622) acknowledged the last argument, but also noted that Dasgupta and Weale's idea of including this dimension echoed radical stances in the standard of living debate, such as Thompson (1963). Prados de la Escosura (2021) presented an augmented HDI that integrates the standard index with a measure of political freedom and applied it to a worldwide dataset over a very long period of time. In what follows we carry out a similar exercise, limited to the case of Italy.

The case of Italy makes clear how the inclusion or exclusion of political and civic liberties significantly alters the historical evolution of wellbeing. It is hard to reconcile the evolution of Figure 1 in the interwar years (1918-1939) with the definition of human development as 'enlarging people's choice' (UNDP 1997, p. 15). After a wave of violence against political opponents, in October 1922 Benito Mussolini was appointed Prime Minister following the so-called 'March on Rome'. Within a few years, his coalition government evolved into a dictatorship that restricted in many ways the scope of citizens' free will. Aside from the killings of opposition leaders, and the illiberal laws abolishing the freedom of the press, banning workers' strikes, and installing a one-party system by 1926, the Fascist regime restricted many civic liberties. The infamous Race Laws (1938), that expelled Jewish citizens from public education and employment and prohibited mixed marriages, were only the most extreme examples because Fascist rule restricted many aspects of individual freedom, for instance by banning internal migration in the 1930s.[23] Still, all versions of the HDI increase in this period, driven by the slow but steady increases in education and life expectancy.[24]

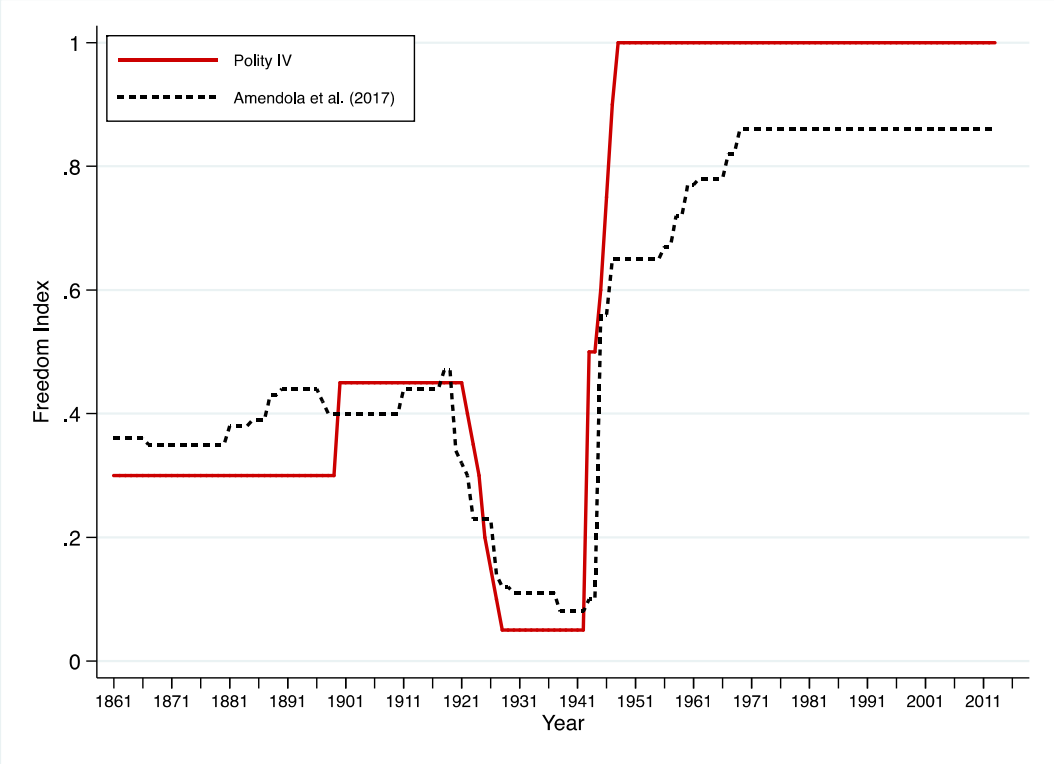
While no indicator of political rights can be considered as better than the others, two sources have gained growing consensus in the literature (Hogstrom: 2013). The first

[23] For a brief review of the history of civil and political rights in Italy, see Amendola et al. (2017: 475-479) and the literature mentioned there.

[24] Ivanov and Peleah (2010) discuss the case of the Soviet Union, whose last decades are covered by the UNDP data.

is Freedom House, a US non-governmental organization that provides annual estimates of the level of political and civil rights enjoyed in almost all polities in the world. Their annual report, *Freedom in the World*, assesses the Political Rights and Civil Liberties indicators, proposed by Taylor and Jodice (1983: 58-68) and available from 1972. For each year, a score between 0 and 4 is assigned to 25 questions on various dimensions of political life and civil rights. The results can be expressed as a score between 0 and 100, or as two separate indices, ranging between 1 (most free) to 7 (least free). Amendola et al. (2017: 475-479) used these two indicators to estimate the DWI for Italy from unification to the present. A second alternative, the 'Polity2 Indicator', is provided by the Polity IV Project (Marshall et al., 2017) and is arguably the most widely used in historical application. The index evaluates political institutions (such as the openness of elections and checks on the executive branch) and assigns a score between -10 and +10 (-6 and +6 being the thresholds for dictatorship and democracy, respectively). Figure 5 illustrates the evolution of political and civil rights in Italy, after rescaling the *Polity2* score to between 0-1. The series show similar trends, and agree on the timing of the major reversal experienced by Italians in the interwar years.

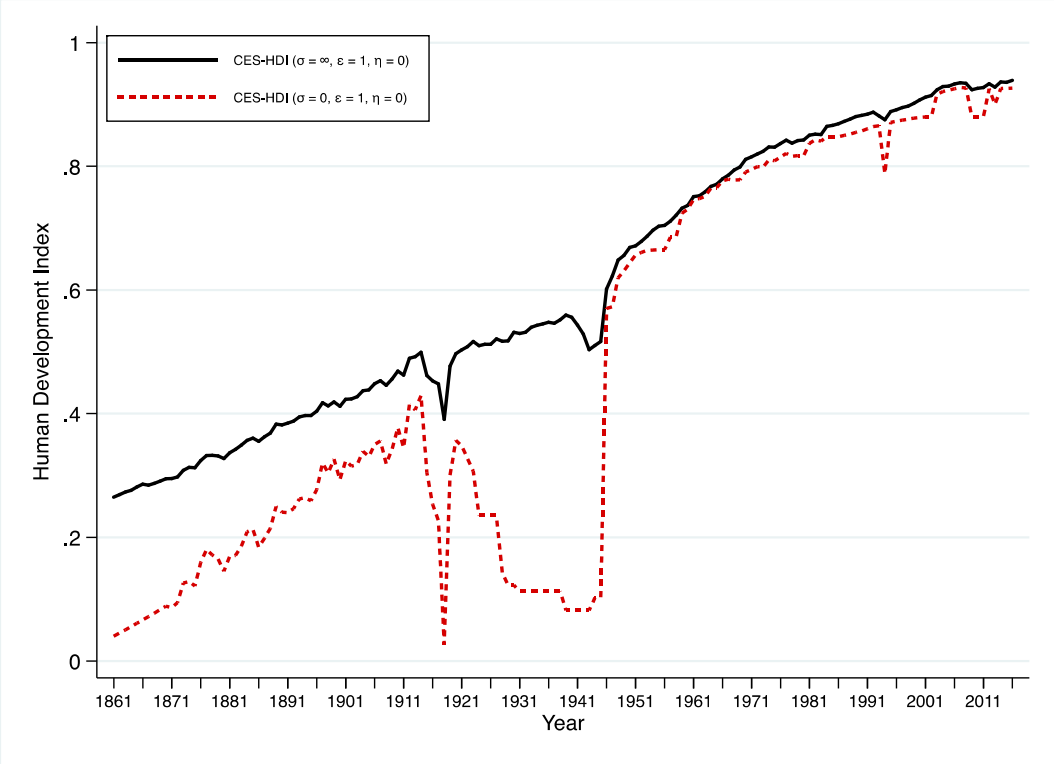
Figure 5 – Freedom indices for Italy, 1861-2016



Source: Authors' calculations.

By taking civil and political rights into consideration, the CES-HDI index better reflects Sen's idea of 'development as freedom', embodied in the concept of *capabilities*. Amendola et al. (2017) showed that including freedom among the components of a human development index defined as in equation 4 alters the long-run evaluation of wellbeing in Italy. Borrowing their freedom index (the dashed line in Figure 5, but the result is robust to the adoption of the Polity IV), Figure 6 shows alternative versions of the CES-HDI augmented with our own freedom index.

Figure 6 –Human development index, preferences and trend reversal



Source: Authors' calculations.

Figure 6 shows the 'folk theorem' at work, once again. Though the two curves are based on the same dataset, adopt the same dimensions in defining human development, and use the same weighting scheme[25], they tell opposite stories. The

[25] The weights underlying Figure 6 are 0.3 for each of the three dimensions of the HDI, and 0.1 for the fourth dimension (freedom index); this scheme is the same for both curves. Furthermore, the result is robust to the choice of $\sigma = 0$ and $\sigma = \infty$; other, less extreme values for σ , would deliver the same result as stylized in Figure 6.

CES-HDI variant depicted by a solid line assumes a high elasticity of substitution, and delivers the message that the interwar years saw a steady improvement in aggregate welfare, despite the fascist dictatorship; the dashed line variant, which assumes a low elasticity of substitution, instead shows a dramatic deterioration – historically unparalleled outside of wartime – over the same period. Depending on the parameter σ – the ethical system of the analyst – 'anything goes', any history can be told.

Conclusion

Faced with the challenge of going beyond the traditional, one-dimensional analysis based on GDP, development economists of the 1990s produced a composite index that was easy to calculate and communicate. Economic historians were not slow to take this idea on board, and began to use the HDI and other composite indices with growing enthusiasm. The original sin, we have argued in this paper, is that the creators of the HDI paid no attention to the analytical foundations of the index (Fleurbaey et al. 2009, Fleurbaey and Blanchet 2013). True, they warned HDI-users about the many choices involved in moving from a single indicator to a composite index, but they focused their recommendations either on the selection of the relative weights for the different dimensions or on the choice of the dimensions themselves.

In this paper, we have made two arguments, one theoretical, and one empirical. The former hinges on standard theoretical tools: following Graaff (1957), Mas-Colell et

al. (1995) and, more recently, Yang (2018), we showed any HDI can be interpreted as a paternalistic social welfare function. Paternalism, it bears reiterating, is not used here as a synonym of arbitrary trade-offs between the HDI dimensions, that is, it should not be confused with the discretionary choice of the weights: instead, paternalism is the feature makes the HDI unsuitable for producing social rankings that are transparently derived from individual rankings. No matter how trivial the proof of this result might look, its implication is powerful, as it rules out any interpretation of the HDI as a transparent, value-free technique. It also rules out any HDI-based conclusion qualifying as evidence-based; historical data feeds the HDI machinery, but the product is a portrait of the owner of the machinery, not of the reality. If the person who uses the HDI is a historian, then the HDI will entirely reflect his/her own judgment of history. The HDI does not tell us about the data, nor about history, but it tells us about the historian.

If the creators of the HDI aimed to vanquish complexity by constructing an objective aggregate measure of social wellbeing, then the battle can be considered lost. The HDI only gives a ranking of wellbeing consistent with the value system associated with the particular specification of the HDI. The plurality of value systems reviewed in this article implies that any specific HDI is just one among equals – as long as preferences and ethical systems cannot be ranked, so it is for HDIs. This conclusion critically depends on the assumption that the interpretation of the HDI must be consistent with the standard value theory. If, instead, economic historians opt to embrace heterodox foundations like, for instance, the functioning/capability

approach, then the HDI remains a viable choice. The price that comes with it is that the resulting analysis would lack standard microeconomic analytical foundations. A big step backward, paradoxically, into the era when statistics, economic theory and economic history were worlds apart.

A second argument against the use of composite indices in economic history comes from our extended empirical investigation of Italian living standards during the last 150 years. Our results have been summarized by borrowing the expression ‘quasi folk theorem' from repeated game theory: we find that (almost) any ranking between two societies can be established by choosing an appropriate specification of the HDI. To demonstrate this result, we introduce a new CES-based HDI family of composite indices. This tool, despite its conceptual simplicity, is a powerful one. We have shown that almost any conclusion can be reached by choosing the CES-HDI parameters appropriately. We can conclude, for instance, that welfare either increased or decreased during the interwar years, when the Fascists were in power, depending on our choice of the elasticity of the substitution parameter. This has nothing to do with the well-known (and very important) arbitrariness associated with the choice of the weighting scheme.

All this said, to many scholars “going beyond GDP” represents a conceptual improvement over any GDP-based analysis, and one that may well be worth paying an empirical price for: the use of proxy variables, the adoption of some arbitrary choices, inconsistency with the standard theory of value, are all lesser evils. In this perspective, the fact that the HDI can be interpreted as a paternalistic social welfare

function, suggests that we should consider the strategy of testing the robustness of any HDI-based analysis with respect to the choice of the analyst's preferences – in our context, with respect to the choice of a different set of values for the parameters σ, η and ε in the CES-HDI index family. When results are consistent across the different specifications, then one may claim that the empirical strategy is a robust one (at least conditional on the restriction of CES-like preferences). This is a solution in line with the original use of the HDI. When they pioneered its use to appraise living standards in antebellum USA, Costa and Steckel (1995) explicitly mentioned social welfare functions, and were clearly aware of the impossibility of identifying one that can be claimed to be superior to the others, "Economic historians must understand that the HDI is a retrospective index of welfare; which asks how and when modern levels of welfare were attained" (p. 36). Costa and Steckel were concerned by the fact that the preferences embedded by the HDI are those of the modern analyst, as opposed to those that analysts in the past would have had; we raise a similar concern, but we focus on the multiplicity of modern analysts' preferences. Sensitivity analysis might be a pragmatic compromise.

References

Amendola, G., G. Gabbuti and G. Vecchi, "Human Development", in G. Vecchi (2017), *Measuring Wellbeing. A History of Italian Living Standards.* New York: Oxford University Press, pp. 454-491.

Arrow, K.J. (1951), *Social Choice and Individual Values*. New York: Wiley.

Astorga, P., A.R. Bergés, and V. Fitzgerald, (2005). "The standard of living in Latin America during the twentieth century", *The Economic History Review*, 58(4): 765-796.

Atkinson, T., Cantillon, B., Marlier, E. and B. Nolan (2002) *Social Indicators: The EU and Social Inclusion*, Oxford University Press.

Baines, D. and N. Cummins, and M.-P.- Schulze (2010), "Population and living standards, 1945-2000", in S. Broadberry and K. O'Rourke, (eds.), *The Cambridge economic history of Modern Europe*. Volume 2: 1870 to the Present. Cambridge University Press, Cambridge, UK: 391-420.

Bértola, L. and J. A. Ocampo (2012), "The Economic Development of Latin America since Independence," Oxford University Press.

Bértola, L. and J. Williamson (2017), "Has Latin American Inequality Changed Direction? Looking Over the Long Run," Springer.

Brandolini, A and G. D'alessio (2009) "Measuring Well-Being in the Functioning Space", in E. Chiappero Martinetti (eds), *A Multi-Voiced Dialog on Global Society*, Milano: Giangiacomo Feltrinelli Foundation.

Brandolini, A., and G. Vecchi (2013), "Standards of Living", in G. Toniolo (ed.), *The Oxford Handbook of the Italian Economy Since Unification*, Oxford University Press, pp. 227-48.

Broadberry, S. and K.H. O'Rourke (2010), "The Cambridge Economic History of Modern Europe," Cambridge Books, Cambridge University Press.

Conte, L., Della Torre, G., and Vasta, M. (2007), "The Human Development Index in historical perspective: Italy from political Unification to the present day", Working paper no. 491, Department of Economics, University of Siena.

Costa, D. L., and Steckel R. H. (1997), "Long-term Trends in Health, Welfare, and Economic Growth in the United States", in R. H. Steckel and R. Floud (eds.), *Health and Welfare during Industrialization.* Chicago: University of Chicago Press , pp. 47-89.

Crafts, N.F.R. (1997a), "The Human Development Index and Changes in the Standards of Living: Some Historical Comparisons", *European Review of Economic History*, 1(3): 299-322.

Crafts, N.F.R. (1997b), "Some dimensions of the 'quality of life' during the British Industrial Revolution", *Economic History Review*, 4: 617-639.

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

Dasgupta, P., and Weale, W. (1992), "On Measuring the Quality of Life", *World Development*, 20(1): 119-131.

Decancq, K. and M. Lugo (2013), "Weights in Multidimensional Indices of Wellbeing: An Overview", *Econometric Reviews*, 32(1): 7-34.

Desai, M. (1991), "Human Development: Concept and Measurement", *European Economic Review*, 35(2-3): 350-357.

Devereux, J. (2020), "The Absolution of History: Cuban Living Standards After 60 Years of Revolutionary Rule", *Revista de Historia Economica – Journal of Iberian and Latin American History*, 39(1): 5-36.

Engerman, S. L. (1997), "The Standard of Living Debate in International Perspective: Measure and Indicators", in R.H. Steckel and R. Floud (eds.), *Health and Welfare during Industrialization, Chicago*. University of Chicago Press/NBER: 17-45.

Federico, G., and G. Toniolo (1991), "Italy", in R. Sylla and G. Toniolo (eds.), *Patterns of European Industrialization*. London: Routledge, 197-217.

Felice, E. (2007), "I divari regionali in Italia sulla base degli indicatori sociali (1871-2001)", *Rivista di Politica Economica*, 67(3-4): 359-405.

Felice, E. and M. Vasta (2015), "Passive Modernization? Social Indicators and Human Development in Italy's Regions (1871-2009)", *European Review of Economic History*, 19(1): 44-66.

Fleurbaey, M. (2009). "Beyond GDP: The quest for a measure of social welfare", *Journal of Economic Literature*, 47(4): 1029-1075.

Fleurbaey, M., and D. Blanchet (2013), *Beyond GDP: Measuring Welfare and Assessing Sustainability*. New York: Oxford University Press.

Fleurbaey, M., E. Schokkaert, and K. Decancq (2009), "What good is happiness?", CORE Discussion Papers 2009017, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Floud, R., Harris, B. (1997), "Health, Height, and Welfare: Britain, 1700-1980", in R.H. Steckel and R. Floud (eds), *Health and Welfare During Industrialization*. Chicago, IL: University of Chicago Press, 91-126.

Foster, J. E., and A. A. Shneyerov. 2000. "Path Independent Inequality Measures." *Journal of Economic Theory*, 91(2): 199–222.

Freedom House (2013), *Freedom in the World: The Annual Survey of Political Rights and Civil Liberties*, Lanham: Rowman & Littlefield.

Gastil, R.D. (1986), *Freedom in the world*. New York: Freedom House.

Gidwitz, Z., M.P. Heger, J. Pineda, F. Rodríguez, (2010), Understanding Performance in Human Development: A Cross-National Study, UNDP Human Development Research Papers 2010/42.

Gormely, P.J. (1995), "The Human Development Index in 1994: Impact of Income on Country Rank", *Journal of Economic and Social Measurement*, 21: 253-67.

Graaff, J. de V. (1957), *Theoretical Welfare Economics*. Cambridge University Press.

Haq, M. (1995), *Reflections on Human Development.* New York: Oxford University Press.

Haq, M. (2003), "The Birth of the Human Development Index", in S. Fukuda-Parr and A. K. Shiva Kumar (eds.), *Readings in Human Development*. Oxford: Oxford University Press , pp. 127-137.

Herrero, C., R. Martínez, and A. Villar (2012), "A Newer Human Development index", *Journal of Human Development and Capabilities*, 13: 247-268.

Hicks, N., and P. Streeten (1979), "Indicators of Development: The Search for a Basic Needs Yardstick", *World Development*, 7: 567-480.

Hirai, T. (2017), "The Creation of the Human Development Approach", Palgrave Mcmillan.

Hogstrom, J. (2013) "Does the Choice of Democracy Measure Matter? Comparisons between the Two Leading Democracy Indices, Freedom House and Polity IV". Government and Opposition, Vol. 48, No. 2, pp. 201–221.

Ivanov, A., and M. Peleah (2010). "From centrally planned development to human development", Human Development Research Paper, 38.

Kakwani, N. (1993), "Performance in Living Standards. An International Comparison", *Journal of Development Economics*, 41, 307-336.

Kelley, C.A. (1991), "The Human Development Index: 'Handle with Care'", *Population and Development Review*, 17(2): 315-24.

Klugman, J., Rodrìguez, F., Choi, C. (2011), "The HDI 2010: New Controversies, Old Critiques", *Journal of Economic Inequality* 9(2): 249-88.

Komlos, J. (2012), "A three-decade history of the antebellum puzzle: explaining the shirking of the US population at the onset of modern economic growth", *The Journal of the Historical Society*, 12: 395-445.

Kovacevic, M. (2010), "Review of HDI Critiques and Potential Improvements", Human Development Research Paper 33, UNDP-HDRO, New York.

Land, K, C. and A. C. Michalos, (2018), "Fifty years after the social indicator movement: has the promise been fulfilled? An assessment and an agenda for the future, *Social Indicator Research*, 3, 835-868.

Maasoumi, E. (1986), "The Measurement and Decomposition of Multi-dimensional Inequality", *Econometrica*, 54: 991-997.

Mas-Colell, A., M.D. Whinston, and J.R. Green (1995), *Microeconomic Theory.* Oxford University Press.

Millward, R., and J. Baten (2010) "Population and living standards 1914–1945", in S. Broadberry and K. O'Rourke, (eds.), *The Cambridge economic history of Modern Europe*. Volume 2: 1870 to the Present. Cambridge University Press, Cambridge, UK.

Marshall, M.G., T.R. Gurr, and K. Kaggers (2017), "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2016. Dataset users' manual." Center for Systemic Peace.

Morris, D. M. (1979), *Measuring the Conditions of the World's Poor: The Physical Quality of Life*. New York: Pergamon.

Offer, A. (2006), *The Challenge of Affluence*, Oxford University Press.

Offer, A. (2003), "Economic Welfare Measurements and Human Well-being", in David, P.A. and M. Thomas (eds.), *The Economic Future in Historical Perspective*, Oxford University Press.

Ogwang, T. (1994), "The choice of the Principal Variable for computing the human Development Index", *World Development*, 22: 2011-14.

Ogwang, T. and A. Abdou (2003), "The Choice of Principal Variables for Computing Some Measures of Human Well-Being", *Social Indicators Research*, 64 (1), 139-152.

Persson, K. G. (2010). *An economic history of Europe: knowledge, institutions and growth, 600 to the present*. Cambridge University Press.

Prados de la Escosura, L. (2010), "Improving Human Development: a Long-run View", *Journal of Economic Surveys*, 24(5), 841-894.

Prados de la Escosura, L. (2013), "Human Development in Africa: A Long-run Perspective", *Explorations in Economic History*, 50(2): 179-204.

Prados de la Escosura, L. (2015a), "World Human Development: 1870–2007", *Review of Income and Wealth*, 61(2): 220-247.

Prados de la Escosura, L. (2015b), "Human Development as Positive Freedom: Latin America in Historical Perspective", *Journal of Human Development and Capabilities*, 16(3): 342-373.

Prados de la Escosura, L. (2018), Well-Being Inequality in the Long Run, CEPR DP 12920

Prados de la Escosura, L. (2021), "Augmented Human Development in the Age of Globalization", *Economic History Review*, 74(4): 946-975.

Mozaffar Qizilbash (2021), "On "Consequentialism" and the Capability Approach", *Journal of Human Development and Capabilities*, 1-21.

Robeyns, I. (2005), "The Capability Approach: a Theoretical Survey", *Journal of Human Development* 6(1), 93-114.

Ravallion, M. (1997) "Good and Bad Growth: the Human Development Reports.", *World Development*, 25(5): 631-638.

Ravallion, M. (2011), "The Human Development Index: A response to Klugman, Rodriguez and Choi", *Journal of Economic Inequality*, 9: 475.

Ravallion, M. (2012a), "Troubling Tradeoffs in the Human Development Index." *Journal of Development Economics* 99L: 201-209.

Ravallion, M. (2012b), "Mashup Indices of Development.", *The World Bank Research Observer*, 27(1): 1–32.

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

Robeyns, I. (2017), *Wellbeing, Freedom and Social Justice. The Capability Approach Re-examined*, OpenBook Publisher.

Sagar A.D., and A. Najam (1998), "The Human Development Index: A Critical Review", *Ecological Economics*, 25: 249-264.

Scanlon, T. M. (2001), Symposium on Amartya Sen's philosophy: 3 Sen and consequentialism. *Economics and Philosophy* 17(1): 39-50.

Sen, A. K. (1985), *Commodities and Capabilities*. Amsterdam: North Holland

Srinivasan T.N. (1994), "Human Development: A New Paradigm or Reinvention of the Wheel?", *American Economic Review, Papers and Proceedings*, 84(2): 238-243.

Streeten, P. (1994), "Human Development: Means and Ends", *American Economic Review, Papers and Proceedings*, 84(2): 232-237.

Taylor, C. L., and D. A. Jodice (1983), *World Handbook of Political and Social Indicators*, Vol. 1. New Haven, CT: Yale University Press.

Thompson, E.P. (1963), *The making of the English working class*. London: Victor Gollancz

Toniolo, G., (2013) (ed.), *The Oxford Handbook of the Italian Economy Since Unification*. Oxford: Oxford University Press.

United Nations Development Programme (UNDP) (1990), *Human Development Report 1990*. New York: Oxford University Press.

United Nations Development Programme (UNDP) (1994), *Human Development Report 1994*. New York: Oxford University Press.

United Nations Development Programme (UNDP) (2013), *Human Development Report 2013*. New York: Oxford University Press.

Usher, D. (1973), "An Imputation to the Measure of Economic Growth for Changes in Life Expectancy", in M. Moss (ed.), *The Measurement of Economic and Social Performance, Studies in Income and Wealth*, vol. 38. New York: National Bureau of Economic Research: 193-226.

Uzawa, H. (1962), "Production Functions with Constant Elasticity of Substitution", *Review of Economic Studies*, 29(4): 291-299

Vecchi, G. (2011), *In ricchezza e in povertà*. *Il benessere degli italiani dall'Unità a oggi.* Bologna: il Mulino.

Vecchi, G. (2017), *Measuring Wellbeing. A History of Italian Living Standards.* New York: Oxford University Press.

Voth, H.-J. (2003), "Living Standards during the Industrial Revolution: An Economist's Guide", *The American Economic Review*, 93(2): 221-226.

Yang, L. (2018) "Measuring Well-being: A Multidimensional Index Integrating Subjective Well-being and Preferences", *Journal of Human Development and Capabilities*, 19(4): 456-476.

Zambrano, E. (2014), “An axiomatization of the Human Development Index”, *Social Choice and Welfare*, 42: 853-872.

Zambrano, E. (2017), “The ‘Troubling Tradeoffs’ Paradox and a Resolution”, *Review of Income and Wealth*, 63(3): 520-541.