How Was Life?
Global Well-being Since 1820

Chapter 13

A composite view of well-being since 1820

by
Auke Rijpma, Utrecht University

This chapter provides a parsimonious overview of the trends in various well-being dimensions covered in the previous chapters by constructing a composite index of well-being. It discusses the crucial problem of choosing a set of weights to calculate such a composite index. Related problems include normalisation of individual indices and dealing with missing observations. The chapter discusses the advantages of various options, and their implications for the final results. It finds that empirically a wide range of aggregation methods generate comparable results. They all indicate that progress in well-being was commonplace since the early 20th century, with the possible exception of Sub-Saharan Africa. It is also found that since the 1970s between-country inequality in composite well-being is lower than in GDP per capita, while being more pronounced in the period before.

Introduction

Economic historians have made great strides in expanding the long-term view of economic development in the decades since Angus Maddison (1982, 1995, 2001) began his historical reconstruction of the world economy. By measuring economic performance through GDP and real wages, the deep origins of global inequality are clearer than ever.

However, what ultimately matters for people is not economic performance *per se*, but the extent to which it improves their well-being. In turn, it is widely believed that the total production of goods and services in an economy, GDP, is an insufficient measure of well-being (Sen 1987, 1999, 2001). Focusing exclusively on GDP implies ignoring distributional issues, as well as the contribution to well-being of non-market goods and services such as health, education, security, governance and the environment. Attention to other aspects of well-being is therefore warranted (Stiglitz et al., 2009).

This book has explored the many aspects of well-being in a long-term perspective. After sketching the contours of world population from 1820 to 2010, the authors have looked at classic income measures such as GDP, real wages and income inequality. They have also analysed long-term changes in health and the biological standard of living by discussing life expectancy and population heights. Furthermore, safety, the quality of the environment, and institutions guaranteeing political freedoms have been analysed. Finally, there has been attention to gender equality in a wide range of well-being indicators.

Population developments matter not only because it is important to know how many people lived in the countries covered in this study, but also because population is a factor in other well-being indicators such as economic growth, health and pressure on the environment. The overall pattern is one of accelerating growth in the 19th century, with European populations showing strong growth at the beginning of the period. As population growth in the developed regions slowed down in recent decades, the population share of fast-growing Asia and Africa has increased.

Although the point of this book is to look beyond GDP per capita as a measure of well-being, the total productive capacity of an economy is still an important measure. Income allows people to consume and gives governments the opportunity to provide public goods and services relating to health and education. Historical national accounting has shown that Europe and its Offshoots already had higher levels of GDP per capita at the start of the 19th century, while much of the rest of the world was still at near-subsistence incomes. World income has increased eighteen-fold since then. However, the gap between rich and poor countries actually widened until the middle of the 20th century, after which a catch-up process began.

Real wages are an important complementary measure of income. Because GDP per capita is an average for the economy as a whole, it does not necessarily capture the income of most people. Income may be measured more satisfactorily by looking at the purchasing power of wage income. Throughout the 19th and 20th centuries, real wages were highest

in the Western Offshoots, with Europe catching up only well into the 20th century. Because progress in real wages was limited outside Europe, its Offshoots and the Middle East and North Africa, inequality between countries has increased strongly over the past two centuries.

Education is important for well-being because improved access to information is of intrinsic importance, but also because there are indirect effects through the impact of education on other well-being indicators, such as income, health and political stability. In respect to basic education, the world has progressed from low to near-universal literacy attainment. Whereas literacy was limited to Europe and its Offshoots in the late 19th century, reading and writing is now widespread throughout the globe. However, the gap in advanced educational attainment has decreased much less, an important fact given the higher educational requirements of today's technological world.

The ability to lead a long and healthy live is a prerequisite for achievements in all other aspects of well-being. Looking at life expectancy to see whether people can enjoy long lives shows that there was a strong, worldwide increase from the 1950s onwards. However, there are substantial inequalities due to the fact that the upwards trend in life expectancy in Europe and its Offshoots was already well underway in the second half of the 19th century. Asia and Latin America began to catch up from the early 20th century and Africa after that, but continued increases in life expectancy in the developed world mean that convergence is not complete.

Human height is determined by nutrition as well as the disease environment, thus capturing two important aspects of well-being. People's average height has generally tended to rise throughout the world. However, there was a temporary stagnation and even decline in height in Europe in the early 19th century and in the Western Offshoots and Asia in the second half of the 19th century.

Personal security, measured here by homicide rates, is also of great importance for well-being. The data coverage for this indicator is problematic, but what it does show is that homicide rates in Western Europe continued the decline started prior to the 19th century, to reach very low levels. One striking aspect of homicide rates is that the high United States rates did not decline to the lower European levels. Sub-Saharan Africa and Latin America stand out for still being very unsafe regions in recent years.

Political institutions matter for well-being because they determine people's control over government decisions, which affect their own lives. Moreover, decisions influencing many aspects of well-being are made in politics, making institutions instrumentally important as well. The tradition of statehood goes back furthest in Asia, Europe and its Offshoots. These regions also have the highest political stability. Democratic institutions and elections were found to have increased throughout the 19th and 20th centuries, though there were important intermezzos, notably in socialist countries. East Asia, the Middle East and North Africa, and Sub-Saharan Africa were late to develop democratic institutions.

The quality of the environment directly influences well-being, because people obtain enjoyment from their environment and often care about the planet they live on. Moreover, the environment also influences other well-being indicators such as health and the sustainability of other indicators. Looking at biodiversity, Europe was already at relatively low levels in 1820, and over time biodiversity declined further until the 1980s. In the 19th and 20th century, biodiversity deteriorated in North America too. Although biodiversity is deteriorating globally, the rest of the world is still at higher levels.

Income inequality is important in order to assess whether total income has benefited large shares of the population. It matters indirectly as well, because high inequality has an adverse impact on other well-being indicators such as crime rates, health and education. Income inequality increased throughout the world over the course of the 19th and 20th centuries, though there was an egalitarian moment in the middle of the 20th century. Asia, Sub-Saharan Africa, and especially Latin America stand out as particularly unequal regions. Globally speaking however, the country in which a person was born became more important for his expected income.

Each chapter has investigated the correlation between per capita GDP and the well-being indicators considered in that chapter. Although many of them have a fairly high correlation with GDP, each indicator also differs sufficiently to warrant separate consideration from GDP per capita in a comprehensive view of well-being. The correlation coefficients are usually between 0.5 and 0.8.

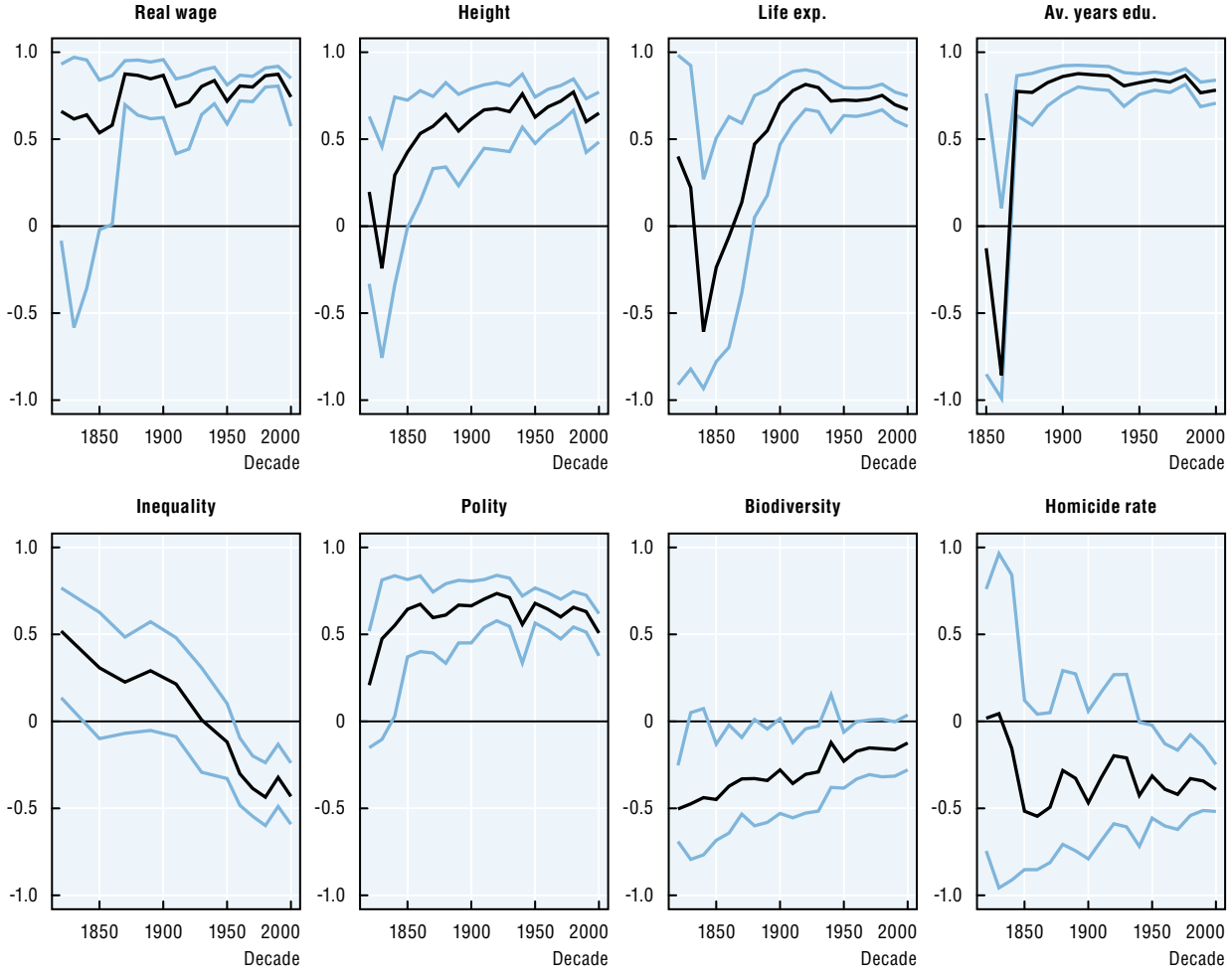
The correlation of the indicators with GDP changes over time (Figure 13.1). For a number of well-being indicators (real wages, height, life expectancy, education, political institutions), the correlation coefficients are closer to zero in the first half of the 19th century than they are in the 20th century, though their estimated values are less precise in this period as there are fewer observations and less variance in per capita GDP and the other well-being indicators. When trying to explain these changing correlations, it is important to keep in mind that these are unconditional correlations. A host of confounding variables could be driving changes in the relationship between per capita GDP and the other well-being indicators, so causal interpretations are not warranted. This change in correlation is not due to new countries entering the data as time moves on: doing the calculations for a smaller set of countries that enjoy good data coverage from an early moment onwards gives qualitatively similar, if more noisy developments.[1] That said, the weaker relationship between per capita GDP and a number of important well-being indicators suggests that economic gains before roughly the 1870s were less likely to be associated with broader improvements in well-being. Only by the turn of the century had a consistent positive correlation of most well-being indicators with GDP come into being. However, income inequality is seen to move from a positive correlation with GDP in the 19th century to a negative correlation by the end of the 20th century.

One prominent candidate for explaining these trends in the relation between per capita GDP and well-being indicators is technology. Preston (1975) famously observed that as medical technology improved and spread across the globe, the same income resulted in higher life expectancy than thirty years earlier. However, Preston observed an upward shift, not changes in the strength or direction of the relationship between per capita GDP and life expectancy. This is reflected in the correlation coefficients presented here, which hardly changed over the course of the 20th century. Preston's observation on the role of technology can, however, be extended backwards. Arguably, it took until the end of the 19th century before medical technology had advanced sufficiently and spread far enough throughout the globe so as to give rise to a relationship between health and per capita GDP (Cutler et al., 2006). This would explain the lower correlation during the 19th century.

A technological argument may also matter for the relation between education and per capita GDP. Early 19th century industrial technology did not require a highly skilled workforce and could thus limit educational attainment (Nicholas and Nicholas, 1992; Goldin and Katz, 1998). More generally, the finding that per capita GDP did not have a consistent positive relation with the other well-being indicators in the 19th century is not

Figure 13.1. **Correlation between well-being indicators and GDP per capita, 1820s-2000s**

Pearson correlation coefficient and upper/lower bounds of 95% confidence interval per decade



Source: Clio Infra, *www.clio-infra.eu*.

StatLink http://dx.doi.org/10.1787/888933096483

entirely unexpected. It has been claimed that the gains from economic growth among early industrialisers did not end up with the common workers, who made up the bulk of the population. For example, nutrition stagnated or even declined during the early decades of the Industrial Revolution (Komlos, 1998), and the disease environment worsened until around 1850 (Szreter, 2003). The data gathered here shows this to have been a pervasive pattern in the 19th century.

Although the value of considering multiple indicators of well-being is clear, the composite picture of well-being has not yet been explored. What does well-being look like when all the indicators are considered together? Do countries compensate a bad performance in one aspect of well-being with a good performance in another? A good example is a socialist country like Cuba, which had achieved only a modest level of income and lacked political freedoms but nonetheless managed to achieve impressive health outcomes (Cooper et al., 2006). However, it might also be the case that poor performance

in one aspect of well-being is accompanied by low scores on other measures as well. The numerous problems befalling many Sub-Saharan Africa countries are prime examples.

These questions are pertinent not only to understanding well-being in the world today, but also for the long-term view of development. What happens to increasing between-country inequality since the Industrial Revolution – the Great Divergence – once the many dimensions of well-being are taken into consideration?

A common approach to questions like these is to aggregate multiple indicators into one composite indicator. Though very popular, such composite indicators are not without problems, and these will be discussed below. Each way of constructing a composite indicator has its own distinct advantages and disadvantages, but so far no fully satisfactory method exists. After discussing methodological issues, this chapter moves on to show the distribution, trends and regional differences in two composite indicators that have been constructed in different ways. Doing this shows a fairly positive picture. Progress is common, and the world is more equal than is indicated when per capita GDP alone is taken as the well-being indicator. However, this lower inequality between countries is also a fairly recent phenomenon, with signs of convergence beginning only in the second half of the 20th century. During the second half of the 20th century, divergence in terms of the composite indicators could well have been stronger than in terms of GDP alone. A number of robustness checks show that the main findings are fairly robust in relation to the way the composite indicator is constructed. The large number of indicators entering the index and their strong correlation mean the overall trends are unlikely to change much due to the weighting scheme. One important exception is the early 19th century, when the weaker correlations between the indicators and the need for imputations add considerable uncertainty. A discussion of the future research agenda into historical composite indicators concludes the chapter.

Description of the concepts used

A composite indicator can be constructed in many ways, and there is no clear consensus on how it should be done. The theoretical problems inherent in combining indicators based on different metrics have made this a topic of substantial debate in academic circles. Despite this, composite indicators have become quite popular. While recognising their theoretical difficulties, those in favour of them hold that they can make a valuable contribution in summarising the messages of the many variables by providing a parsimonious view (Nardo et al., 2005). However, the theoretical difficulties of a composite indicator indicate a need to closely examine how robust they are.

The academic debate on composite indicators usually focuses on the twin issues of weights and trade-offs. The indicators entering a composite indicator usually have widely different ranges and different units of measurement, and this needs to be addressed. For example, if average years of education are combined with life expectancy without any transformations, the range of 19 to 82 years for life expectancy makes a larger contribution to the composite indicator than does the 0-13 years for education. For this reason, the indicators are often normalised or standardised (e.g. Morris, 1980; UNDP, 2010; Boarini and Mira d'Ercole, 2013; Nardo et al., 2005). However, a normalisation procedure does not fully resolve the aggregation and weighting issue. Trade-offs exist in the composite indicator (Ravallion, 2011a, 2012a, 2012b). The core of the problem is that combining two or more indicators into one introduces the possibility of exchanging some amount of one indicator for some amount of another while keeping the overall score equal. This means

that aggregation amounts to a statement on the relative importance of the indicators. For instance, four years of education might be worth as much as one additional year of life expectancy in the composite indicator.

One solution is to gather subjective information on the determinants of well-being. The study of subjective well-being (life satisfaction or "happiness") and its determinants is one way to obtain such insights (Fleurbaey et al., 2009; Schokkaert, 2007), though measurement of subjective well-being is not without its own challenges, particularly when making comparisons across cultures. These methods hold great promise, but historical public opinion is not available, thus limiting its use for the historical data discussed in this book. A final problem is that these surveys can give counterintuitive results. For example, broad agreement exists about the importance of longevity for well-being: one should be alive to experience any well-being. Though there is a relationship between subjective well-being and health status at the individual level (Dolan, Peasgood and White, 2008), Deaton (2008) finds that conditional on income, life expectancy has no significant effect on life satisfaction at the country level. The data gathered for this book give similar results when they are regressed on country-level life satisfaction scores from the World Value Surveys (Rijpma, 2014). Moreover, conditional on income, none of the other well-being indicators has a significant association with life satisfaction. Only average years of education and income inequality are statistically significant, but they have an unexpected sign (negative and positive, respectively). Regressing each well-being indicator separately on life satisfaction gives results that are more in line with expectations, though inequality still has a positive rather than a negative effect on well-being. Using individual-level survey data would probably allow for more accurate estimation of the contribution of the various indicators of well-being (e.g. Boarini et al., 2012), but the indicators gathered by these surveys do not match the (country-level) data in the present volume. Surveys among experts or the public can also provide information on the relative importance of indicators of well-being (Boarini and Mira d'Ercole, 2013).

Research in welfare economics has also made contributions by constructing composite indicators of well-being using economic theory (Becker et al., 2005; Dowrick et al., 2003). To achieve this goal, it is necessary to express other well-being indicators in monetary terms (Fleurbaey and Gaulier, 2009). The consistency with economic theory and the expression of the resulting indicator as a correction to GDP are useful properties of these approaches. However, setting prices on other dimensions of well-being requires a great deal of data, and results in complexity that hampers their interpretation.

Another line of research into composite indicators considers poverty measures from a multidimensional perspective. Poverty measures take a poverty line and count the number of people who fall below it and the extent of their poverty. The point of a multidimensional poverty line is to consider multiple indicators of poverty (Alkire and Foster, 2011). Since such measures are designed to count the number of poor, they require individual-level data and are not well suited to country-level data. Moreover, these poverty measures still require a decision about weights and, additionally, the poverty cut-off for each indicator.

One way to circumvent the issue of weighting altogether is to be content with only ranking the countries, rather than assigning an exact number to each of them. The most famous proponent of this approach is probably Amartya Sen (1999; 1987, pp. 32-3). This approach has been worked out in some detail (Atkinson and Bourguignon, 1982; Duclos et al., 2006). The basic idea is to rank a country higher than another country only if it scores as good or better on all indicators. In that case the country will score higher than the other

regardless of the weighting procedure. The problem with this approach is that rankings can be incomplete when a country scores better on one indicator but worse on another. As the number of indicators and observations increases, an unambiguous ranking becomes more difficult to achieve. Since this book has considered over nine indicators and historical data provide many observations, this is especially problematic. Only 728 unambiguous rank relations could be established among 40 200 possible complete-case comparisons (there are over 22 million comparisons if including incomplete cases). A further issue with using ranking techniques is that it makes comparisons over time difficult. If an improvement of rank between two decades is found, it is not known whether this is an actual improvement or a relative improvement.

A final set of methods is data-driven. Principal components analysis (PCA) is one popular example of these techniques (e.g. Slottje, 1991; Chakravarty, 2003). Other latent variable models such as factor models are very similar and are also widely used in the construction of composite indicators (Høyland et al., 2012; Kaufmann et al., 2004). These models can be used to construct a composite indicator that reflects as much as possible the information shared by the indicators (Nardo et al., 2005). One advantage of these methods is that a statistical model provides a coherent way of thinking about the issue of aggregation. Given assumptions about the relations between the indicators, the weights are chosen so that they differentiate between countries as best as possible. Moreover, statistical models make it possible to account for sources of uncertainty, such as measurement error in a composite indicator (Høyland et al., 2012; Kaufmann et al., 2004; Treier and Jackman, 2009). Allowing for the possibility of measurement error is very important in the context of historical data (Feinstein and Thomas, 2002).

The main disadvantage of a latent variable approach is that there is no reason to suppose that a statistical property, for instance, the correlation between variables, captures the correct trade-offs between the indicators. Moreover, the statistical justification behind a latent variable model is not without issues. Having each indicator capture only part of well-being seems to fit well with a multidimensional concept of well-being. Such concepts are frequently used, and examples include healthy condition, cognitive ability, ideological disposition and democracy (Treier and Jackman, 2008; Lee, 2007). However, this does require that the structure of the latent variable – well-being – is reflected in the covariance structure of the country-level indicators.

Two approaches to constructing a composite indicator will be used here. This will give a first assessment of the sensitivity of the composite indicator to the way it is constructed. The first approach is to calculate an arithmetic average of the standardised indicators with equal weights for each dimension. This is probably the most well-known way of constructing a composite indicator. The second approach is to estimate a composite indicator through a latent variable (factor) model. There is no guarantee that the latent variable captured by this procedure is well-being, especially not when the indicators are measured at the country-level (well-being is primarily an individual-level concept). Nonetheless, there are good reasons to consider such a model. First of all, the results turn out to be qualitatively similar to the equal-weighting approach. The weights this procedure gives turn out to be fairly similar to the equal-weighting case, thus giving some assurance about the robustness of the approach.

The second and most important reason to consider the factor analysis model separately is to explore the important issue of imputations. Computing a composite indicator requires observations of all indicators for a given country at a given point in time, which is not

readily achieved with historical data. Only 159 such complete cases exist in the data. Imputations are therefore inevitable, and the way this is done can influence the results. The latent variable model allows for imputation of missing data using the covariance between the indicators. This is a useful property, since, if lacking data on one variable, say, height, it is desirable to estimate this missing value based on other highly correlated variables (and less so for variables with a lower correlation, such as political institutions). Furthermore, the underlying data structure of the latent-variable model can be specified. This allows the imputations to reflect the fact that countries in a region or period are expected to have similar scores on indicators (a multilevel structure) or that indicators within a country are correlated over time (a time-series cross-sectional structure). Finally, the latent variable model also allows the uncertainty caused by imputing to be reflected in the imputations and in the composite indicator itself.

Main highlights of trends in composite indicators of human well-being

The indicators used to construct the composite indicator are roughly equal to those considered in the preceding chapters. Some of the indicators capture a similar aspect of well-being and are thus grouped. First is income, which is captured through per capita GDP and real wages. Incorporating inequality is important for a good understanding of whether income benefits a large share of the population. However, income inequality enters the model separately, because it has a low and negative correlation with GDP per capita. Life expectancy serves as an indicator for the population's health. Height captures both income and health. Education is measured by average years of education. Next, political freedoms are incorporated through data on democratic institutions (Polity2). Finally, the environment and personal security are measured through mean species abundance (MSA) and homicide rates, respectively. Table 13.1 summarises some key statistical features of the well-being indicators used in this chapter.

The first method used here to construct a composite indicator is to calculate an equally-weighted average over all the indicators. The indicators are standardised so that the mean and standard deviations for the entire 1820–2000 period and all countries are zero and one respectively. No further transformations were performed on the data (following advice in Ravallion, 2012b, and Chakravarty, 2003). Multiple indicators for the same group (per capita GDP and real wages in the case of income; height and life expectancy in the case of health) are given half the weights of the other indicators so that each group of indicators has an equal

Table 13.1. **Well-being indicators: coverage and summary statistics**

	Mean	SD	Min	Max	N	Period
GDP per capita	3 702.59	4 763.11	225.06	34 440.92	1 500	1820-2000
Real wages	24.50	32.29	0.48	349.38	1 023	1820-2000
Height	167.75	4.57	152.36	183.20	1 357	1820-2000
Life expectancy	56.09	14.57	18.90	82.19	1 618	1820-2000
Average years of education	3.72	3.32	0.01	13.07	1 638	1820-2000
Income Inequality	43.18	8.71	16.14	73.70	895	1820-2000
Polity2	7.74	7.15	1.00	21.00	1 896	1820-2000
Mean species abundance	0.86	0.15	0.05	1.00	3 857	1820-2000
Homicide rate	5.93	8.61	0.00	71.75	812	1820-2000

Source: Clio Infra, *www.clio-infra.eu*.

StatLink http://dx.doi.org/10.1787/888933097585

weight. Inequality and homicides are thought to contribute negatively to well-being, and they are given a negative weight. Composite indicator scores were calculated for each decade and country, which were then used to calculate population-weighted regional averages.

Because the composite indicator can be calculated only when all indicators are present, it is necessary to impute missing data. Here, the same procedure was used as elsewhere in this book to calculate regional averages. First, the growth rate of a similar country from the same region for which data on a given indicator was available was calculated. This growth rate was then applied to the country for which observations were missing so as to make projections. Although this is a transparent way of addressing missing observations, assuming growth for the imputations will make finding progress more likely. To prevent imputations driving the results too strongly, cases where more than half the composite indicator would consist of imputed values have been dropped. The influence of imputations will be tested extensively below by using other imputation techniques.

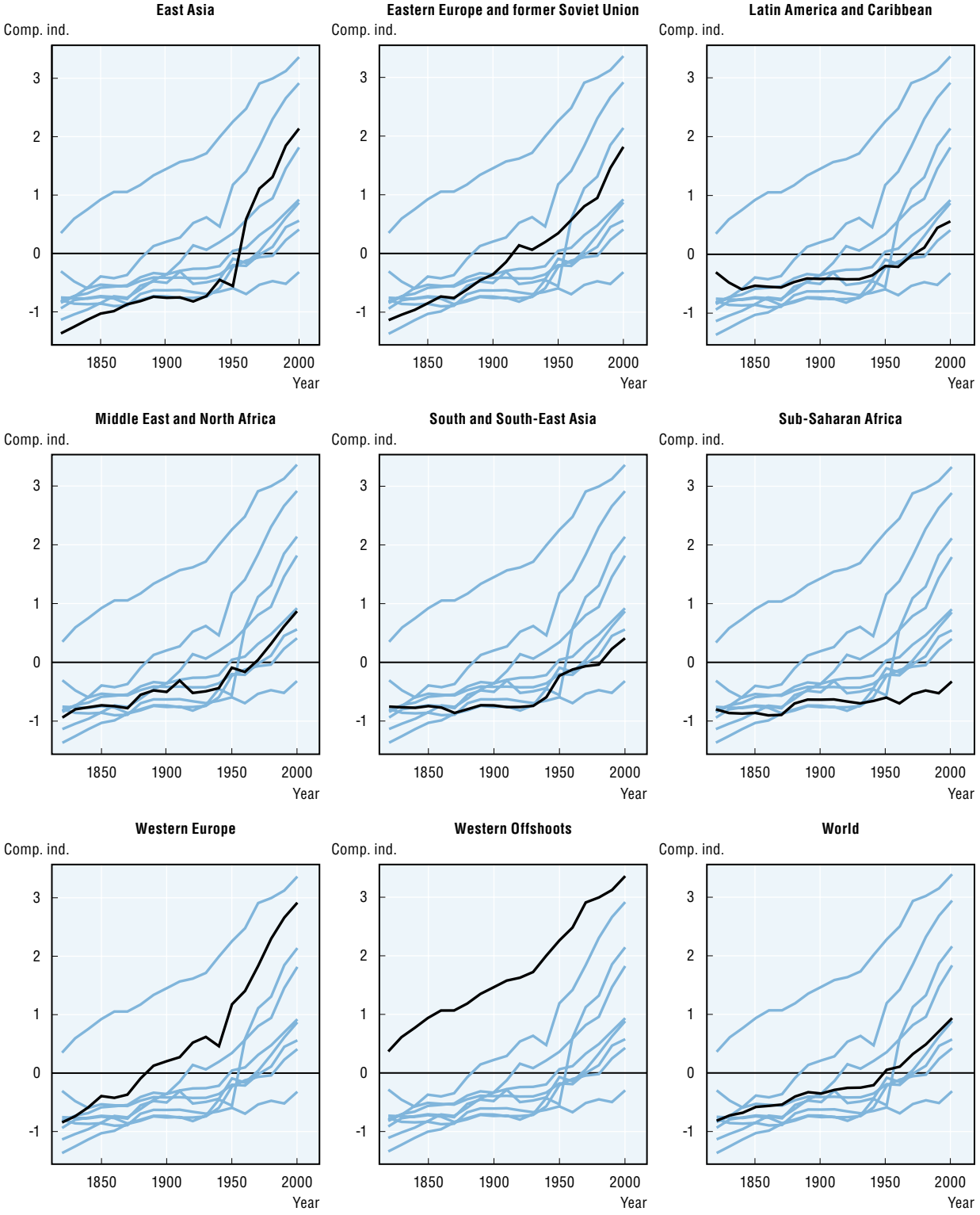
Regional population-weighted averages for this composite indicator are presented in Figure 13.2. Throughout the period, the composite indicator was highest in the Western Offshoots. Already in 1820, its scores on the composite indicator were higher there than they were in the worst-performing regions today – Sub-Saharan Africa and South and Southeast Asia. These are also the two regions showing the least progress since 1950. Elsewhere, there was substantial progress (about two standard deviations of the composite indicator). Convergence with the Western Offshoots was limited, however, with only Western Europe and East Asia making up substantial ground.

Though divergence between regions still exists, the development process looks different when the composite indicator is compared with trends in per capita GDP (Figure 13.3). The divergence in terms of the composite indicator until the 1950s is stronger than it is in terms of per capita GDP alone. The very populous regions of Asia even performed worse in terms of the composite indicator in this period. This changed from the 1970s onwards as progress in Europe and its Offshoots is substantially less than it was in terms of per capita GDP.

The second way of constructing a composite indicator used here is through a latent variable model. This approach has benefits in terms of dealing with missing data and accounting for the uncertainty this causes in the composite indicator, but this comes at the expense of transparency. To understand the results, it is therefore important to make the workings of a latent variable model explicit. At its core, the model is similar to the previous method in that the composite indicator is created by standardising the variables and calculating their average. The difference is in how the weights are determined. One way to think about the latent variable model is that it gives the weights to construct a composite indicator that reflects as much shared information between the variables as possible to differentiate between countries as best as possible. To do this, highly correlated indicators get higher weights, since this is a sign of shared information and increases the robustness of the rankings (Foster et al., 2013). For these weights to make sense, however, the key assumption of a latent variable model is that the indicators are correlated with each other because of their correlation with the latent variable. For this to hold, a unitary, underlying concept of well-being linked to the observable indicators has to be plausible at the cross-country level. This is not straightforward for all the variables. The model estimated here is a Bayesian variant of a multilevel latent variable model (Jackman, 2009; Høyland et al., 2012; Gelman and Hill, 2007; Merkle, 2011; Lee, 2007). More details on the estimation procedure can be found in the background paper to this chapter (Rijpma, 2014).

Figure 13.2. **Regional averages of a composite well-being indicator, 1820s–2000s**

Standardised values (global mean equals zero), population-weighted decadal averages

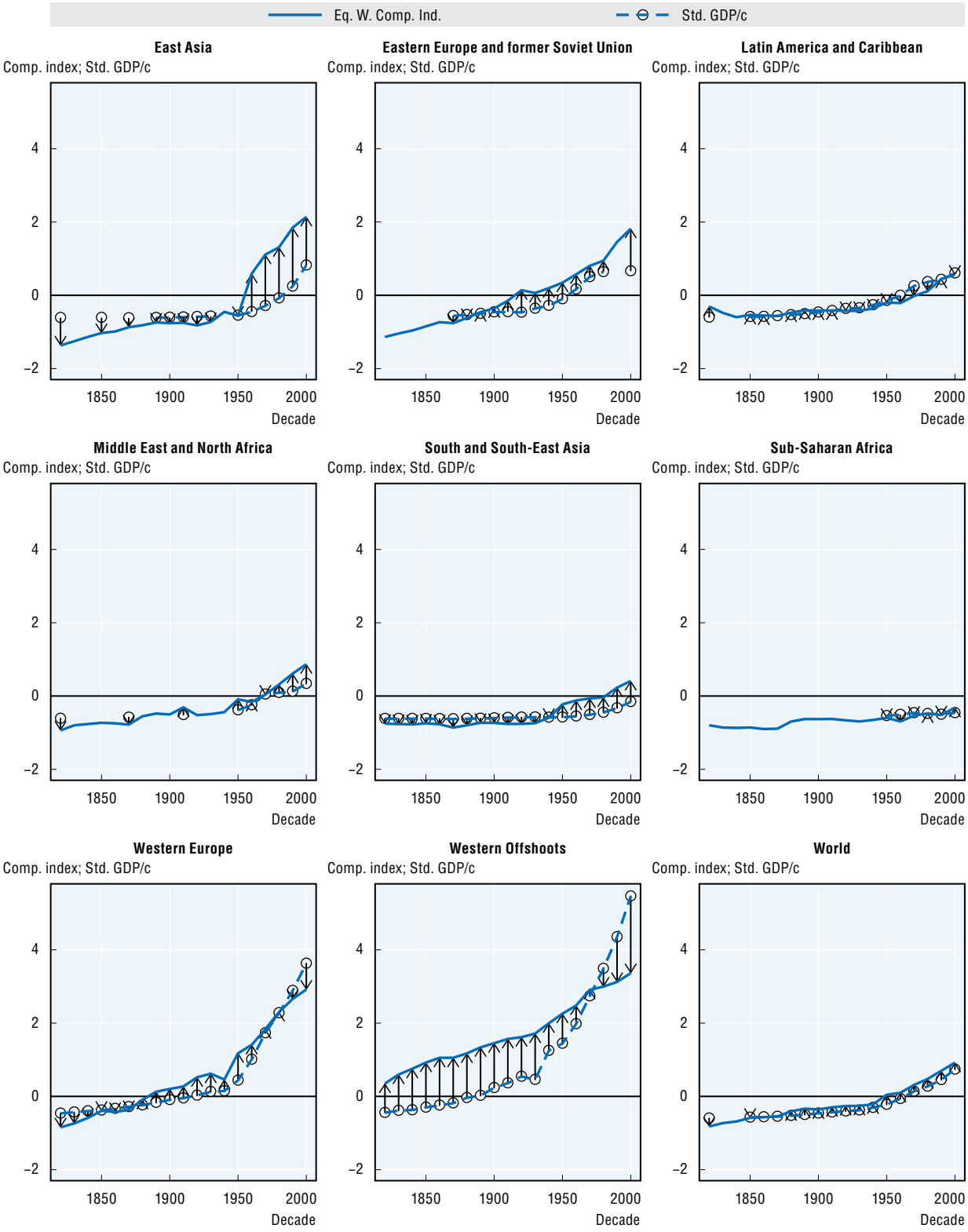


Source: Clio Infra, *www.clio-infra.eu*.

StatLink http://dx.doi.org/10.1787/888933096502

Figure 13.3. **Regional averages of a composite well-being indicator and standardised GDP per capita, 1820s–2000s**

Standardised values (global mean equals zero), population-weighted decadal averages



Source: Clio Infra, *www.clio-infra.eu*.

StatLink http://dx.doi.org/10.1787/888933096521

Table 13.2. **Factor loadings for a composite well-being indicator**
Posterior means and quantiles

	Mean	q05	q50	q95
GDP per capita	0.77	0.73	0.77	0.80
Real wages	0.70	0.66	0.70	0.74
Height	0.72	0.68	0.72	0.77
Life expectancy	0.88	0.85	0.88	0.90
Average years of education	0.92	0.90	0.92	0.94
Income inequality	-0.22	-0.27	-0.23	-0.18
Polity2	0.67	0.64	0.67	0.70
Mean species abundance	-0.36	-0.39	-0.36	-0.34
Homicide rate	-0.15	-0.21	-0.15	-0.09

Source: Clio Infra, *www.clio-infra.eu*.

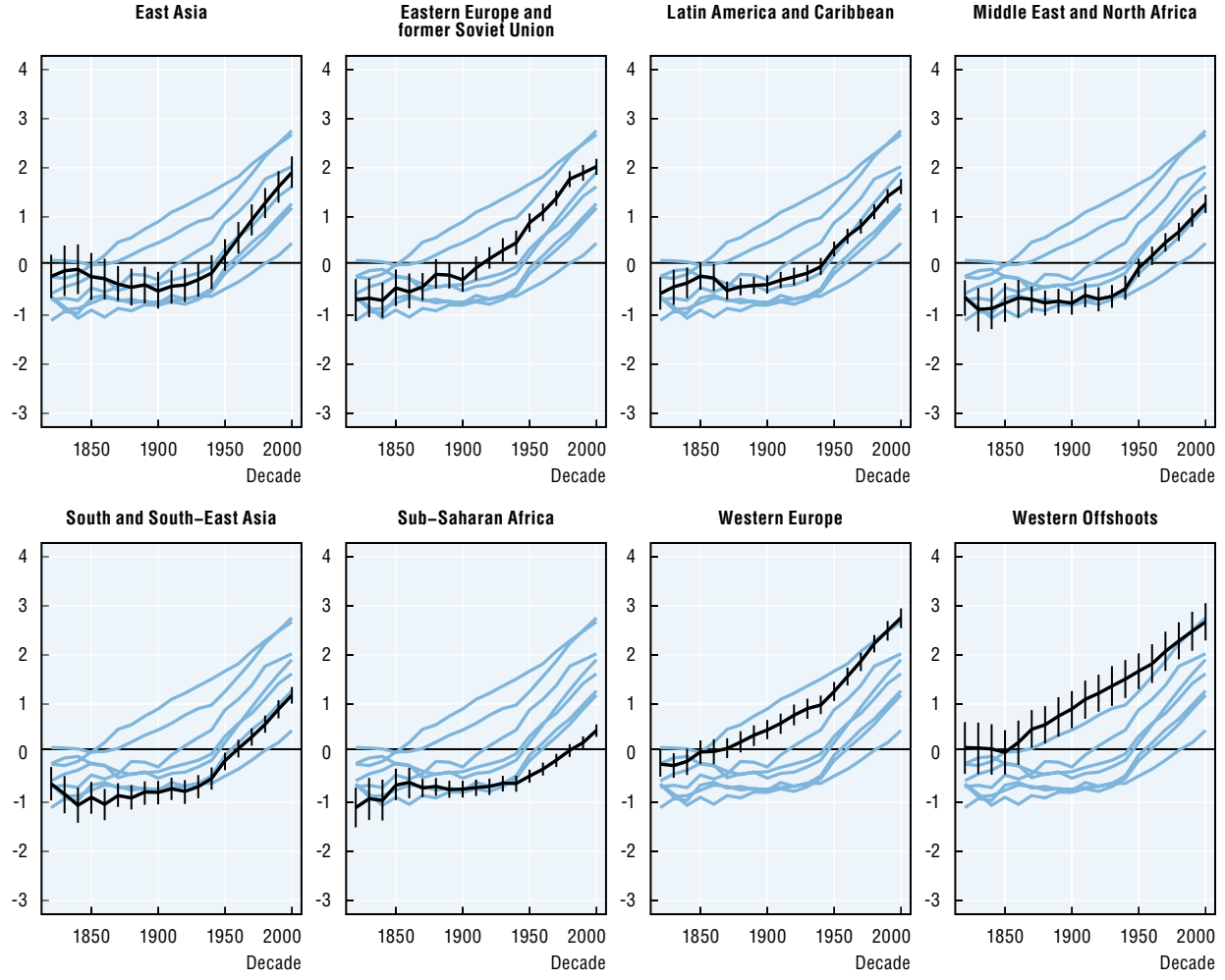
StatLink http://dx.doi.org/10.1787/888933097604

First, a look at the weights (factor loadings) is warranted. Most of the indicators contribute to the composite indicator in the expected direction. Moreover, the contributions are significantly different from zero. GDP, height, life expectancy, education, political institutions and the quality of elections all contribute positively to the composite indicator. Their weights are in the 0.7-0.9 range, meaning that the two grouped indicators (health and income) have a stronger contribution than in the equal-weighting scheme where they were at half the weights of the other indicators. As expected, income inequality and homicide rates contribute negatively to the composite indicator, though their contribution is smaller than that of the other indicators. Biodiversity turned out to be a problematic indicator. Because of its negative correlation with most of the other variables, it had a negative loading. This means higher biodiversity lowered the score on the composite indicator, which is opposite to how biodiversity is thought to affect well-being (though excluding the biodiversity indicator did not substantially change the results). This clearly shows that a statistical approach to a composite indicator can lead to counter-intuitive results. With one of the indicators working in an unexpected manner, there is some uncertainty about what concept the factor model measures: it could arguably be broader social progress rather than well-being. It is nonetheless worthwhile to explore the results of the factor model. For one, the results are generally very similar to the results obtained from other ways of constructing a composite indicator from this book's data. With this in mind, improved imputations are the main reason for estimating the latent variable model. The model estimated here is a multilevel model, which means that the imputations are driven not only by the correlations between the indicators, but also by observations from the same region and decade. However, unlike the previous imputations, there is no assumption of growth in the imputation procedure. Moreover, the number of imputations and their strength is reflected in the uncertainty of the estimates of the composite indicator.

Key results from this analysis are described below. First, consider developments at the level of the geographic regions (Figure 13.4). Wherever possible, the figures also give an indication of the precision of the estimates by providing 90% confidence intervals. Overall, the regional estimates of progress in the early 19th century are not precise, largely because of the number of observations that had to be imputed. The lack of precision makes it difficult to distinguish between regions in this period, clearly showing how important the imputations are for the earlier period. After about 1950, the precision on the regional

Figure 13.4. **Composite well-being indicator by region, latent variable model, 1820s–2000s**

Median and 90% confidence interval, standardised values (global mean equals zero), decadal averages



Source: Clio Infra, *www.clio-infra.eu*.

StatLink http://dx.doi.org/10.1787/888933096540

estimates is definitely good enough to distinguish between most regions. The influence of imputations is also reflected by the higher scores on the composite indicator in the early 19th century in East Asia, Sub-Saharan Africa and Latin America compared to the previous indicator. In these decades, there were very few observations, causing the overall mean (zero by definition) to have a strong impact on the imputations.

The global picture here is still one of progress, as all regions improve (though Sub-Saharan Africa to a much lesser extent). Yet it is also a picture of divergence, as some regions start from a higher level and also start advancing earlier. At the beginning of the 19th century, the Western Offshoots already had the highest scores on the composite indicator. They kept their lead until the last decades of the 20th century, when Western Europe caught up, though this result is sensitive to weighting. Halfway through the 19th century progress began in Europe and its Offshoots. Due to this early start, by the end of the 19th century a clear gap had arisen between Europe and its Offshoots on the one hand and the rest of the world on the other.

The rest of the world only began catching up later. In the early 20th century, Eastern Europe and especially the Soviet Union witnessed substantial progress and closed some of the gap with the leaders. Convergence with the West is stronger there in terms of the composite indicator than what is shown by per capita GDP. Despite some flaws in socialist countries (for instance, the lack of political freedom), progress was nonetheless made in many well-being-indicators. Asia, Latin America, and the Middle East and North Africa began making up lost ground in the first half of the 20th century. Sub-Saharan Africa did not see any substantial progress until the 1950s, and even after that convergence with the rest was very limited.

These results are similar to those found by Leandro Prados de la Escosura (2010, 2014). His composite indicator shows substantial progress and convergence after the First World War, though he finds that these processes slowed down after the 1940s. Nicholas Crafts (2002) has also calculated Human Development Index (HDI) scores for the 1870–1990 period using the UNDP's pre-2010 procedure, which relied on the arithmetic mean of the normalised indicators. Crafts too observes that substantial convergence took place after 1950. A comparison of the country ranks with his scores is instructive. Broadly speaking, Craft's composite indicator and the one presented here rank countries similarly (rank correlation of 0.96). However, some substantial outliers exist, mostly a reflection of the larger number of indicators used here and the logarithmic transformation on per capita GDP used by the UNDP and Crafts.

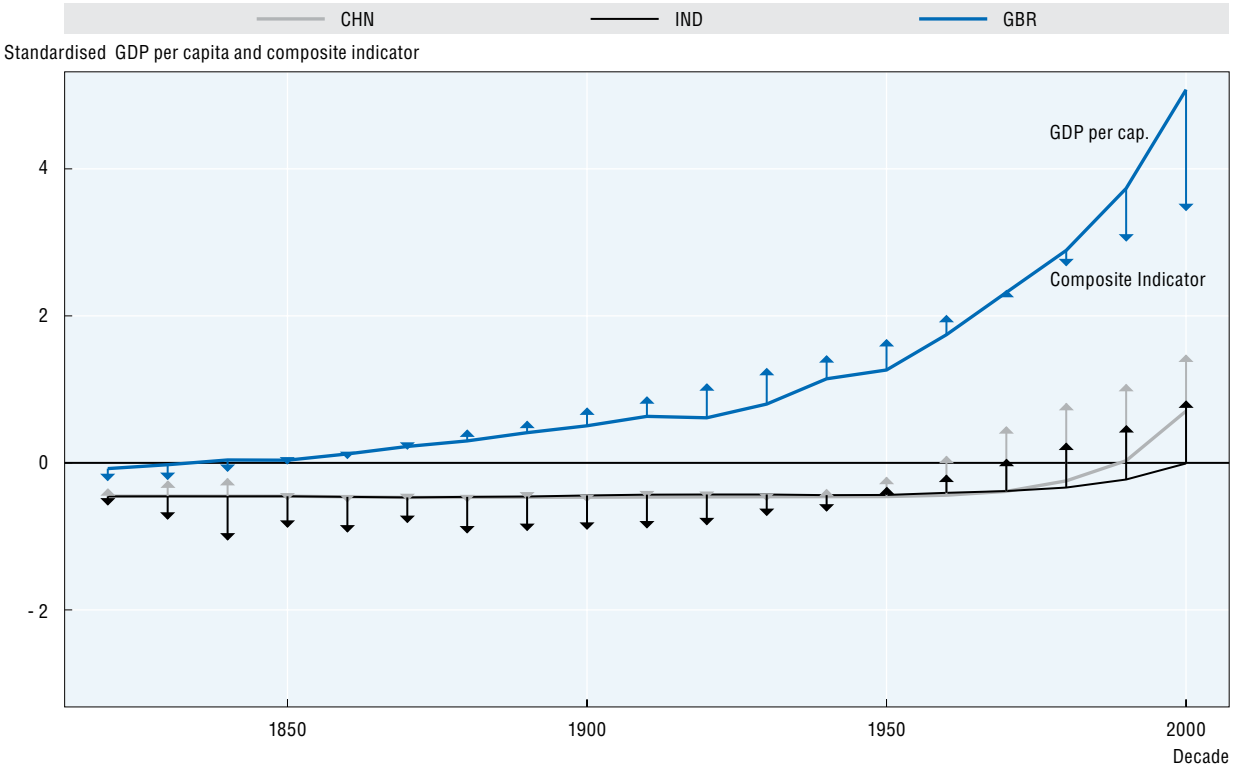
Although there is a divergence in progress measured via this indicator between *ca* 1820 and 1950, it looks different compared to GDP alone. Strikingly, the divergence in the late 19th and early 20th century between Europe and the Western Offshoots on the one hand and the rest of the world on the other is more pronounced in the case of the composite indicator. From the early 20th century onwards, however, Eastern Europe and the former Soviet Union, Latin America and Asia make up more ground than GDP alone suggests. Looking at some of the most prominent countries in the debate on the Great Divergence (the United Kingdom, China and India, see Figure 13.5) confirms this picture in more detail. Only after about 1950 does the composite indicator show more convergence than GDP.

Progress in the composite indicator over the entire period was stronger than growth in per capita GDP. In terms of GDP, some of the worst-performing countries in 2000 are no better-off than the poorest countries in 1820. The composite indicator paints a different picture. The countries with the lowest composite indicator scores in 2000 are generally better-off than the lowest-scoring countries in 1820. Many even do better than the countries with the highest scores in 1820. Indicators of well-being other than GDP are spread more evenly across the globe, and this is reflected in the composite indicator. This result is, however, sensitive to the way the composite indicator is constructed. In the equal-weighting case more countries in Sub-Saharan Africa show very little improvement. This is due to the strong weight given to inequality and security and above all to the assumption of growth underlying the imputations in the equal-weighting case (using equal weights on the imputations of the latent-variable model shows qualitatively similar results to the composite indicator created by the latent-variable model itself).

Figure 13.6 presents the distribution of the composite indicators across countries at three years (1850, 1900 and 2000). In 1850, the distribution of the composite indicator was still fairly equal. Most countries in the middle and even at the bottom of the distribution have similar scores. Moreover, substantial uncertainty surrounds many of the estimates.

Figure 13.5. **Composite well-being indicator and GDP per capita in the United Kingdom, China, and India, 1820s-2000s**

Standardised values (global mean equals zero), decadal averages



Note: The continuous line refers to GDP per capita; the arrows to the composite well-being indicator.
Source: Clio Infra, *www.clio-infra.eu*.

StatLink http://dx.doi.org/10.1787/888933096559

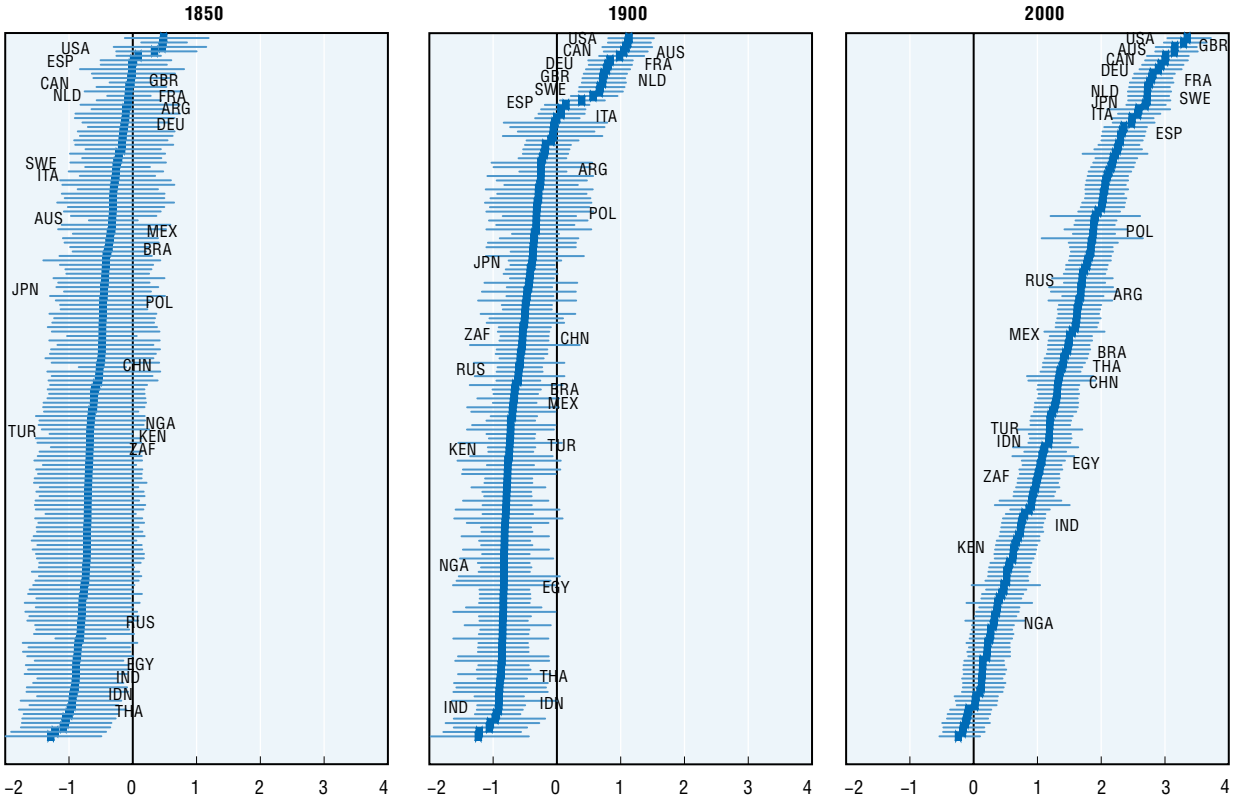
This, and the relatively equal distribution, makes the countries difficult to distinguish in terms of their score on the composite indicator. Only at the very top do substantial differences start appearing.

By 1900, between-country inequality had increased as a clear group of frontrunners had arisen. In Western Europe and especially in its Offshoots, the composite indicator was much higher than in the rest of the world. The precision in the estimates has also increased substantially, meaning it is now possible to make distinctions between many more countries. By 1950, no clear group of leaders existed anymore. The highest scores were still found in north-western Europe and the Western Offshoots, but countries in southern Europe and Latin America as well as the Soviet Union and Japan had begun closing the gap. Generally, the scores on the composite indicator have increased substantially by this time, with half the countries having a score higher than zero, the overall mean for the entire period. At the same time, however, between-country inequality was at its highest at this point in time.

By 2000, there was further progress in the composite measure. Nearly all of the countries are now above the overall 1820–2000 mean. This represents considerable progress if it is remembered that in 1850 nearly all countries were still below that point. Furthermore, many countries experienced a substantial increase between 1950 and 2000, most by more than one standard deviation of the global 1820–2000 distribution. Although there was an increase in all countries between 1950 and 2000, progress was much lower in many

Figure 13.6. **Composite well-being indicator across countries, 1850, 1900 and 2000**

Medians and 90% confidence intervals, standardised values (global mean equals zero), decadal averages



Source: Clio Infra, *www.clio-infra.eu*.

StatLink http://dx.doi.org/10.1787/888933096578

Sub-Saharan Africa countries as well as in Afghanistan and Haiti. Despite lagging development in these countries, however, between-country inequality as measured by the composite indicator in 2000 had declined relative to 1900 and 1950.

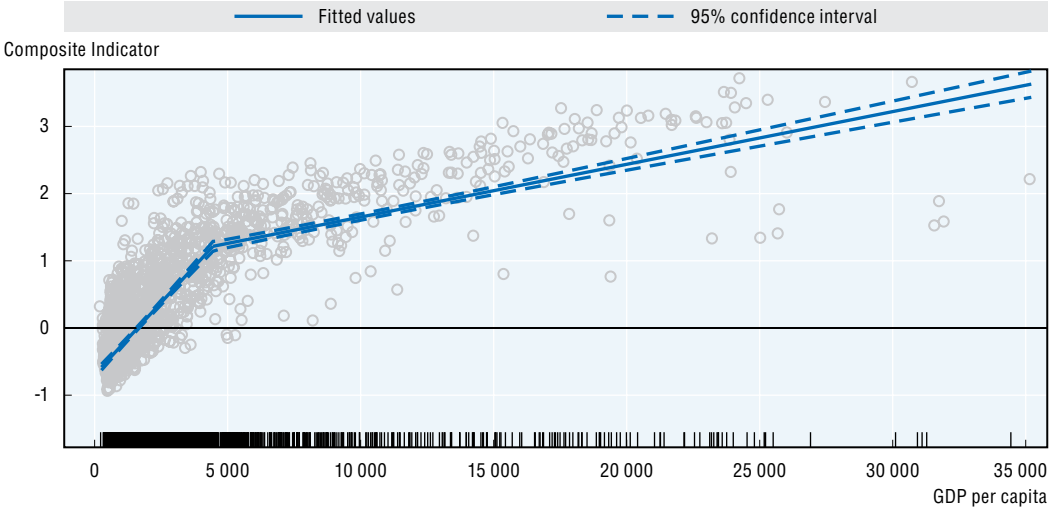
It is also useful to point out the uncertainty of the estimates. Generally speaking, it is difficult to distinguish countries that are close to each other in the rankings, especially when uncertainty is high due to missing observations. For example, the estimates suggest that the chance that two high-ranking countries in the 1850s like the United States and Norway had a different score on the composite indicator was 54%, not much better than determining this by the flip of a coin. Only when the highest countries are compared to countries below the top of the ranking does the chance of the two being different become large. For instance, the chance of the composite indicator in the United States in 1850 being higher than in Britain in the same year is about 90%. As for developments over time, only by about 1870 is it possible to state with some certainty that a country like Britain had improved its position relative to 1820. More precision and higher increases after 1850 means it becomes easier to distinguish countries and progress.

An interesting question that arises from the composite indicator is the importance of GDP for achieving progress. In developing regions, the composite indicator improved in the second half of the 20th century despite slow growth in per capita GDP. Moreover, since the 1970s, per capita GDP rose much more than the composite measure in Europe and its Offshoots. It was also observed earlier that the correlation between per capita GDP and

other well-being indicators was not very strong in the first half of the 19th century. Tests for a break in the relation between per capita GDP and the composite indicator indicate that their relation starts to change around USD 4000-5000 (Muggeo, 2008). Controlling for country and decade fixed-effects gives roughly the same results (Figure 13.7). Likewise, estimating the global trend in the composite indicator conditional on per capita GDP shows that higher scores on the composite indicator were attained each decade even if GDP had remained constant. These results can probably be ascribed to two phenomena. First is the inclusion of (weakly) negatively correlated well-being indicators (inequality, homicides). Second, the break reflects the concave relationships between per capita GDP and a few of the other indicators, an indication that well-being indicators other than per capita GDP are more equally distributed between countries.

Figure 13.7. **Segmented relation between a composite well-being indicator and GDP per capita**

Standardised values (global mean equals zero) and US dollars at 1990 PPP, decadal averages



Source: Clio Infra, *www.clio-infra.eu.*

StatLink http://dx.doi.org/10.1787/888933096597

Finally, the robustness of the results to the aggregation method should be discussed. A large number of methods of constructing the composite indicator have been tried, both here and in the background paper (Rijpma, 2014). Overall, most approaches lead to similar results. Trying a large number of weighting schemes made only small differences to the regional results. The differences between the two aggregation methods employed here insofar as they were due to weighting were also small. However, the imputation method did have a substantial impact on the results for the early 19th century, as shown by the latent variable model approach.[2] Overall, however, the results remain fairly robust, though care is warranted in making statements about differences regarding the composite indicator in the 19th century.

Discussion

A composite indicator is a tool that, while not without its share of problems, is very useful for summarising developments. Two methods for constructing such an indicator were explored here: a straightforward equal-weighting scheme and a statistical approach in the form of a latent variable model.

The correlation between the various well-being indicators was generally high, meaning that low or high well-being often went hand in hand. Nonetheless, compared to per capita GDP, many of the indicators were distributed more equally across the globe. This means that, compared to per capita GDP, progress in the composite indicator was stronger and differences between countries were less pronounced. However, this lessened between-country inequality was largely a phenomenon of the 1970s and after. Before then, between-country inequality was actually more pronounced when looking at the composite indicator. The composite indicators measured an early lead for the Western Offshoots and Western Europe. They would continue to extend their lead in the late 19th and early 20th centuries. East Asia, Latin America, and Eastern Europe and the Soviet Union began to make up grounds sometime in the first half of the 20th century. South Asia and Sub-Saharan Africa made much less progress in the 20th century, though the composite indicator suggests that this lag was less stark than suggested by income alone. In short, moving away from solely considering GDP as a measure of well-being suggests less world poverty and less between-country inequality, though substantial differences continue to exist between countries. Most ways of constructing a composite indicators will probably give similar conclusions.

An important issue is the changing relations between GDP and the composite indicator. It shows that more economic activity and income does not necessarily imply progress in other domains of well-being. GDP alone may therefore be an imperfect measure of well-being, especially for societies in the past. Although this suggests that a multi-dimensional perspective is important, it also hints at the possibility that constructing a single composite indicator to cover such a long time period can be problematic.

Priorities for future research

Three elaborations of this view of well-being in the past seem important. First, more relevant indicators could be added. For example, unemployment is an important determinant of subjective well-being (Fleurbaey et al., 2009; Fleche et al., 2011). It could therefore be relevant for an historical composite indicator as well. Second, there is the possibility of systematically incorporating valuable expert and public opinion into the aggregation procedure. Although this made little difference to the resulting composite indicators here, it is a lead worth following. Finally, there has been little attention to distributional issues beyond the inclusion of income inequality as an indicator in composite indicator. However, inequalities can exist in many more dimensions of well-being, and they can be analysed in more sophisticated ways than has been done here. Harmonised micro-data, historical as well as contemporary, can be of great value for such an effort (e.g. Ruggles et al., 2003). Doing this could also cross the gap from country-level to individual conceptualisations of well-being.

Notes

1. These and other technical matters are treated in more depth in a background paper to this chapter (Rijpma, 2014).
2. A dynamic factor model that attributes more importance to the gradual developments of the indicators over time was also estimated to this end (Zhang and Nesselroade, 2007).

References

Alkire, S. and J. Foster (2011), "Counting and Multidimensional Poverty Measurement", *Journal of Public Economics,* Vol. 95, pp. 476-487.

Atkinson, A.B. and F. Bourguignon (1982), "The Comparison of Multi-Dimensioned Distributions of Economic Status", *The Review of Economic Studies*, Vol. 49, pp. 183-201.

Becker, G.S., T.J. Philipson and R.R. Soares (2005), "The Quantity and Quality of Life and the Evolution of World Inequality", *The American Economic Review,* Vol. 95, pp. 277-291.

Boarini, R. et al. (2012), "What Makes for a Better Life?:The Determinants of Subjective Well-Being in OECD Countries – Evidence from the Gallup World Poll", OECD Statistics Working Papers, 2012/03, OECD Publishing, *http://dx.doi.org/10.1787/5k9b9ltjm937-en*.

Boarini, R. and M. Mira D'Ercole (2013), "Going Beyond GDP: An OECD Perspective", *Fiscal Studies,* Vol. 34, pp. 289-314.

Chakravarty, S.R. (2003), "A Generalized Human Development Index", *Review of Development Economics,* Vol. 7, pp. 99-114.

Cooper, R.S., J.F. Kennelly and P. Orduñez-Garcia (2006), "Health in Cuba", *International Journal of Epidemiology,* Vol. 35, pp. 817-824.

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

Cutler, D., A. Deaton and A. Lleras-Muney (2006), "The Determinants of Mortality", *Journal of Economic Perspectives,* Vol. 20, pp. 97-120.

Deaton, A. (2008), "Income, Health and Well-being Around the World: Evidence from the Gallup World Poll," *Journal of Economic Perspectives,* Vol. 22/2, pp. 53-72.

Dowrick, S., Y. Dunlop and J. Quiggin (2003), "Social Indicators and Comparisons of Living Standards", *Journal of Development Economics,* Vol. 70, pp. 501-529.

Duclos, J., D.E. Sahn and S.D. Younger (2006), "Robust Multidimensional Poverty Comparisons", *The Economic Journal,* Vol. 116, pp. 943-968.

Feinstein, C.H. and M. Thomas (2002), "A Plea for Errors", *Historical Methods: A Journal of Quantitative and Interdisciplinary History,* Vol. 35, pp. 155-165.

Fleche, S., C. Smith and P. Sorsa (2012), "Exploring Determinants of Subjective Well-being in OECD Countries: Evidence from the World Value Survey", *OECD Statistics Working Papers*, No. 2012/01, OECD Publishing, *http://dx.doi.org/10.1787/5k9ffc6p1rvb-en*.

Fleurbaey, M., K. Decancq and E. Schokkaert (2009), "What Good Is Happiness?" *www.gate.cnrs.fr/IMG/pdf/Fleurbay_Oct_2009.pdf*.

Foster, J.E., M. McGillivray and S. Seth (2013), "Composite Indices: Rank Robustness, Statistical Association, and Redundancy", *Econometric Reviews,* Vol. 32, pp. 35-56.

Gelman, A. and J. Hill (2007) *Data Analysis Using Regression and Multilevel/hierarchical Models,* Analytical Methods for Social Research, Cambridge University Press, New York.

Goldin, C. and L.F. Katz (1998), "The Origins of Technology-Skill Complementarity", *The Quarterly Journal of Economics,* Vol. 113, pp. 693-732.

Honaker, J. and G. King (2010), "What to Do About Missing Values in Time-Series Cross-Section Data", *American Journal of Political Science,* Vol. 54, pp 561-581.

Høyland, B., K. Moene and F. Willumsen (2012), "The Tyranny of International Index Rankings", *Journal of Development Economics,* Vol. 97, pp 1-14.

Jackman, S. (2009), *Bayesian Analysis for the Social Sciences,* Wiley Series in Probability and Statistics, John Wiley and Sons, Chichester.

Komlos, J. (1998), "Shrinking in a Growing Economy? The Mystery of Physical Stature During the Industrial Revolution", *The Journal of Economic History* Vol. 58 pp. 779-802.

Kaufmann, D., A. Kraay and M. Mastruzzi (2004), "Governance Matters III: Governance Indicators for 1996, 1998, 2000 and 2002", *The World Bank Economic Review* Vol. 18, pp. 253-287.

Lee, S.-Y. (2007), *Structural Equation Modeling: A Bayesian Approach,* John Wiley and Sons, Chichester.

Maddison, A. (1982), *Phases of Capitalist Development*, Oxford University Press, Oxford.

Maddison, A. (1995), *Monitoring the World Economy 1820-1992,* Development Centre Studies, OECD Publishing, Paris.

Maddison, A. (2001), *The World Economy: A Millennial Perspective*, Development Centre Studies, OECD Publishing, Paris, *http://dx.doi.org/10.1787/9789264189980-en*.

Merkle, E.C. (2011), "A Comparison of Imputation Methods for Bayesian Factor Analysis Models", *Journal of Educational and Behavioral Statistics,* Vol. 36, pp. 257-267.

Morris, M.D. (1980), "The Physical Quality of Life Index (PQLI)", *Development Digest,* Vol. 18, pp. 95-109.

Nardo, M. et al. (2005), "Handbook on Constructing Composite Indicators: Methodology and User Guide", *OECD Statistics Working Papers*, No. 2005/03, OECD Publishing, *http://dx.doi.org/10.1787/533411815016*.

Nicholas, S.J. and J.M. Nicholas (1992), "Male Literacy, 'Deskilling,' and the Industrial Revolution", *The Journal of Interdisciplinary History,* Vol. 23, pp. 1-18.

Plummer, M. (2003), "JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling," DSC (Disability Services Commissioner) 2003 Working Papers, *www.r-project.org/nosvn/conferences/DSC-2003/Drafts/Plummer.pdf*.

Prados de la Escosura, L. (2010), "Improving Human Development: A Long-Run View", *Journal of Economic Surveys,* Vol. 24, pp. 841-94.

Prados de la Escosura, L. (2014), "World Human Development: 1870-2007," *Review of Income and Wealth,* forthcoming.

Ravallion, M. (2011), "The Human Development Index: a Response to Klugman, Rodriguez and Choi", *The Journal of Economic Inequality,* Vol. 9, pp. 475-478.

Ravallion, M. (2012a), "Mashup Indices of Development", *The World Bank Research Observer,* Vol. 27, pp. 1-32.

Ravallion, M. (2012b), "Troubling Trade-offs in the Human Development Index", *Journal of Development Economics,* Vol. 99, pp. 201-209.

Rijpma, A. (2014), "A composite view of long-term well-being: construction and trends", *CGEH (Centre for Global Economic History) Working paper series, http://cgeh.nl/working-paper-series*.

Ruggles, S. et al. (2003), "Integrated Public Use Microdata Series (IPUMS): Version 3.0 Minneapolis: Historical Census Projects", University of Minnesota.

Schokkaert, E. (2007), "Capabilities and Satisfaction with Life", *Journal of Human Development,* Vol. 8, pp. 415-430.

Sen, A. (2001), *Development as Freedom,* 1st Oxford University Press.

Sen, A. (1999), *Commodities and Capabilities,* Oxford University Press, India.

Sen, A. (1987), *The Standard of Living: The Tanner Lectures,* Cambridge University Press.

Slottje, D. J. (1991), "Measuring the Quality of Life across Countries", *The Review of Economics and Statistics,* Vol. 73, pp. 684-693.

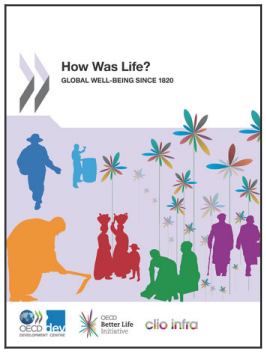
Stiglitz, J.E., A. Sen and J.P. Fitoussi (2009) "Report by the Commission on the Measurement of Economic Performance and Social Progress", *www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf*.

Szreter, S. (2003), "The Population Health Approach in Historical Perspective", *American Journal of Public Health,* Vol. 93, pp 421-431.

Treier, S. and S. Jackman (2008), "Democracy as a Latent Variable", *American Journal of Political Science,* Vol. 52, pp. 201-217.

UNDP (United Nations Development Programme) (2010), *Human Development Report 2010: The Real Wealth of Nations: Pathways to Human Development,* United Nations Development Programme.

Zhang, Z. and J.R. Nesselroade (2007), "Bayesian Estimation of Categorical Dynamic Factor Models", *Multivariate Behavioral Research,* Vol. 42, pp 729-56.



From:

How Was Life?

Global Well-being since 1820

Access the complete publication at:

http://dx.doi.org/10.1787/9789264214262-en

Please cite this chapter as:

Rijpma, Auke (2014), "A composite view of well-being since 1820", in Jan Luiten van Zanden, *et al.* (eds.), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris.

DOI: http://dx.doi.org/10.1787/9789264214262-17-en