DOI: 10.1111/joes.12618

SPECIAL ISSUE ARTICLE

JOURNAL OF ECONOMIC SURVEYS WILEY

Maddison-style estimates of the evolution of the world economy: A new 2023 update

Jutta Bolt[1,2] | **Jan Luiten van Zanden**[3]

[1]Department of Economics, Econometrics and Finance, University of Groningen, Groningen, The Netherlands

[2]Department of Economic History, Lund University, Lund, Sweden

[3]Department of Economic and Social History, Utrecht University, Utrecht, The Netherlands

Correspondence
Jutta Bolt, University of Groningen, Groningen, The Netherlands.
Email: j.bolt@rug.nl

Abstract

This paper surveys the literature on historical national accounting, discusses the importance of relative income benchmarks for, in particular, historical income estimates, and presents an update of long run global economic development with a new version of the Maddison Project Database (MPD). As benchmarks are central to methodologies for global income comparisons over time, and therefore vital to MPD, we analyze the consequences and biases of three benchmarks, the 1990 benchmark, the 2011 benchmark and the multiple benchmark method following the recent Penn World Tables (PWT) methodology, for pre-1940 income estimates. We develop a methodology to determine which benchmark in combination with time series produces the best anchor for the historical income estimates in the MPD. We conclude that the best way forward for the Maddison Project is to stick to the original 1990 benchmark, yet with two important changes. First, we integrate the 2011 benchmark for the post-1990 period, and second, we fine tune the dataset for the pre-1940 period by integrating a new historical benchmark for the US/UK comparison in 1909. By integrating more bench-

marks, the MPD moves closer to a multiple benchmark approach as developed by the PWT.

KEYWORDS
Global Economic History, historical national accounting

1 | INTRODUCTION

Having reliable and consistent data on long-term national economic performance is crucial to understanding patterns of global economic development and to developing and testing theories of economic growth. The gross domestic product (GDP) is one of the most important indicators measuring economic performance, and is a central driver of broad well-being across nations (OECD, 2021). The measurement of national incomes has its roots in attempts by William Petty to construct aggerate estimates of England's economy in 1665. Since then, the field of (historical) national accounting has developed immensely, but the pioneering research on the systematic international comparisons of national income and product only started in the 1930s with the works of Clark and Kuznets (Clark, 1940; Kuznets, 1941). Angus Maddison was one of the first economist to build a global dataset of consistent GDP (pc) estimates, allowing for comparisons of incomes across countries and over time, focusing especially on the per-1950 period (Maddison, 1995, 2001, 2007). After his passing, the Maddison Project was established as an ongoing research project aimed at standardizing and updating academic work in the field of historical national accounting, in the tradition of Maddison's syntheses of long-term economic growth.

Central to global comparisons of long-run GDP per capita estimates are the conversions of *national* income estimates from a national currency basis to a common currency using purchasing power parities (PPPs). PPPs measure price differences between countries and represent what a dollar of income buys in one country relative to another. Because relative price levels are lower in less advanced economies, exchange rates—the alternative to transform national currencies into a common currency—typically understate the real (or cross-country comparable) GDP level of such economies relative to the richer ones. But while the principle of using PPP-converted income levels for cross-country comparisons is clear, the choice of method in estimating PPPs is more difficult to make; yet it is of first-order importance to debates on historical living standards.

The aim of this paper is three-fold. First, we give an overview of the literature on historical national accounting in the next section. Second, because PPP benchmarks are central to the methodologies for global income comparisons over time, we discuss the consequences of the various PPP benchmarks available for, in particular, historical income estimates, review the different benchmark options, and discuss choice criteria for these options in section three and four. Third, we present all historical income series that have recently become available and which we integrated in the 2023 Maddison Project Database and give an updated overview of global long run economic development in section five. The final section concludes our analysis.

2 | HISTORICAL NATIONAL ACCOUNTS, SURVEY OF THE FIELD

GDP measures the level of a nation's economic activity; when calculated per person, it measures their level of access to goods and services in a society. It is one of the most consistently measured

indicators of material wellbeing over time. William Petty made the earliest known attempts to measure a nation's (England's) total economic activity in 1665 and 1676; his findings were published in 1890. Only six years later, Gregory King improved England's national income estimates by providing more detailed calculations. Many subsequent estimates were based on Petty's and King's data until Arthur Young presented the first independent estimate in 1770. The frequency of new estimates of national income increased—not only for England, but for other countries as well—especially during the 19th century, as concepts were developed alongside advancements in economic theory. These developments made it possible for the first international comparisons to be published in the latter half of the 19th century (Krantz, 1983).

However, research on the systematic measurement and international comparison of national income and production capacity started only in the 1930s, with the seminal works of Colin Clark and Simon Kuznets (Clark, 1940; Kuznets, 1941). Kuznets was the first to create a standardized framework for the measurement of GDP and its components to facilitate the construction of long historical time series of GDP. This formed the basis for his studies on economic growth in the 1960s. Kuznets's work also lay the foundations for the internationally accepted System of National Accounts (SNA), which was developed with the contributions of—in particular—John Maynard Keynes, Richard Stone, and James Meade during the 1940s. It was published by the United Nations in 1953. Since then, national statistical agencies make official income estimates available for most countries in the world using the SNA.

Historical national accounting research—on periods before official statistics were collected—has been expanding rapidly over the past decades and has resulted in a substantial increase in long-run series of per capita GDP. For countries that have sufficient historical data available, reconstructions of national incomes or products are based to a large extent on the SNA framework. The most robust estimates are often produced using the output-based approach, which involves summing up the net value added of all goods and services produced within a country in a given year. This method is based on the production of all sectors in the economy and is used for countries such as Britain, the Netherlands, and Sweden (Broadberry et al., 2015; Schön & Krantz, 2015; van Zanden & van Leeuwen, 2012). These countries have a strong tradition of quantitative historical research, which makes it possible to calculate the annual components of output.

In the context of data scarcity, national income estimates often rely on more indirect methods and stronger assumptions to obtain accurate estimates. The consumption-based method is mostly used for countries in which the necessary data for output-based estimates is not present. This method has been used to estimate GDP for countries such as France, Spain, Italy, Portugal, and Poland, dating back to the 16th century or earlier (Álvarez-Nogal & De La Escosura, 2013; Malanima, 2011; Malinowski & van Zanden, 2017; Palma & Reis, 2019; Ridolfi, 2017). The consumption-based approach begins by estimating agricultural production, which is calculated consistently across these studies using estimates of real wages and the level of urbanization during the period in question, assuming a demand function for agricultural produce. Other sectors of the economy are subsequently estimated using various methods, ultimately leading to an estimate of GDP.

While these long-run national income estimates are important and insightful, they do not allow for a direct comparison across countries, as they are generally expressed in national currency units. Angus Maddison was one of the first economists to construct a database that allowed for a comparison of incomes for a global set of countries over time (Maddison, 1995, 2001, 2007). In particular, when it comes to the pre-1950 period, the Maddison database Historical Statistics of the World Economy and its successor, the Maddison Project Database, have the widest coverage of data on GDP per capita across countries and over time currently available, including

information for over 160 countries and covering the period from Roman times to the present. In comparing income over time and across space, the crucial factor is that national income estimates are converted from a national currency basis to a common currency using PPPs. PPPs measure price differences between countries; they represent what a dollar of income buys in one country relative to another.

2.1 | The importance of benchmarks

Although the principle of using PPP-converted income levels for cross-country comparisons is clear, the choice for a particular method of estimating PPPs is more difficult to make, yet of first-order importance in many debates on historical living standards. For scholars interested in studying changing income levels over a given time period, the challenge is to estimate PPPs and the implied relative levels of GDP per capita for each year (excluding the ICP benchmark year). There are three general methods that can be used for this purpose:

1. The extrapolation (or projection) method
2. The multiple/historical benchmark method
3. The econometric method

Maddison's original work relied on the extrapolation method, under which GDP or GDP per capita levels from the benchmark year 1990 are extrapolated to earlier and later years using growth rates of GDP per capita in constant prices from each country's National Accounts (or other growth estimates). However, because it keeps prices constant, this approach disregards changes in relative prices over time and related shifts in economic structures, which may lead to biased results.

The second and third methods have been developed as alternatives to the extrapolation method. The econometric method estimates relative price levels for historical periods using the contemporaneous relationship between price levels and other variables that are observable over long time periods (Prados de la Escosura, 2000).[1] The multiple benchmark method uses more than one benchmark to anchor the long-run income series. In this paper, we use the method that was introduced in the Penn World Tables 8.0 (Feenstra et al., 2015): it covers the period since 1950, and makes use of all available ICP benchmarks.[2] The benchmarks are used at face value, and are made consistent by changing the GDP growth rate in the years between both benchmarks (for example, for the year 1980, it takes the "best" estimate of the PPP value based on the ICP round for 1980, while for 2005, the authors rely on PPPs from ICP 2005). The historical benchmark method takes this logic to the pre-1950 period estimating PPPs between a pair of countries based on available price and quantity information.

In addition to the availability of different methods for constructing PPP converted incomes, scholarship continues to produce new series of estimates of historical national accounts, new estimates of benchmarks comparing GDP levels of different nations in years past, and new ICP rounds that give updated estimates of PPPs. This raises a question: what is the best way forward for the Maddison Project Database? The database currently relies on Maddison's 1990 benchmark; this was state-of-the-art at the time of its creation, and its results have served as a point of reference in many studies since. Yet the quality of PPP measurement has improved considerably, due to the increased number of countries included during consecutive ICP rounds and the development of more sophisticated methodologies over the years.

Both most recent ICP results—for 2011 and 2017—are considered very good benchmarks (Deaton & Schreyer, 2022). However, the Maddison Project uses the extrapolation method for creating comparable historical income series. This method assumes that the underlying price structure of every nation's economy remains fixed over time. Consequently, it is no easy task to simply change to a newer, even if potentially better, benchmark. This is because the further one moves from the benchmark year, and/or the more difference evolves between national price structures, the larger the bias becomes in keeping the price structure constant. As a result, the comparisons of relative incomes generally become less reliable. An alternative to using a more recent PPP benchmark would therefore be to extrapolate income series backwards from the earliest available benchmark for each country, using the multiple benchmark approach as developed by PWT. Unfortunately, these earlier benchmarks are often considered less reliable and include only a smaller sample of countries, generating a bias in historical comparisons (Ryten, 1998). Therefore, the key problem is that the choice for the "best" benchmark for international comparisons of national incomes over long periods of time has to strike a balance between the potential biases of a too-recent benchmark and those of using earlier, lower-quality PPPs. We'll return to this in section four.

2.2 | Limitations of GDP

Gross domestic product (GDP) per capita is a widely used measure of a country's economic well-being. However, the concept has several limitations in terms of how it is measured and exactly *what* it measures. One of the main limitations is that it only measures market production, not non-market output (such as the services provided by a stay-at-home spouse or parent). In societies in the past, a larger part of household consumption was typically produced in the domestic sphere, and the commercialization of certain domestic activities in more recent time periods may have artificially inflated the measured growth of GDP (e.g., people eat out more today, instead of preparing their own meals at home, compared to 50 years ago). Furthermore, there have been substantial improvements in how market activities are recorded since the first introduction of the SNA, which could lead to an upward bias in GDP growth as well. However, there are also biases that lead to an underestimation of growth, such as products whose quality and utility have increased over time without a corresponding shift in their prices. The price of information, for example, has dropped significantly in the internet era, but this is not captured by typical measurements of GDP growth, as many internet services are free of charge and therefore, by definition, do not contribute to the GDP.

Additionally, GDP estimates for years past are subject to certain margins of error, as they are often based on partial data and assumptions about the links between data and economic activities represented. These limitations often become more relevant for GDP estimates of years in the distant past, although there may be exceptions for historical periods about which a great deal of information is available. We probably know more about Medieval England, then pre-Columbian Latin America or 19th century Sub-Saharan Africa.

Although GDP per capita is the most consistently measured indicator of *average* material well-being over time, it does not provide information on how that material wellbeing is distributed among people. There is still a large gap between national accounts—focusing on macro indicators and growth—and inequality studies, which base distributions on survey and tax data (often not fully consistent with macro totals). A way forward could be the creation of distributional national accounts, from which we can derive inequality statistics that are consistent with

national accounts, as recent work for the US has shown (Piketty et al., 2018). Finally, the measurement of GDP does not include social and environmental costs associated with higher production, including pollution and the depletion of natural resources. And national accounts do not include information about important nonmaterial dimensions of wellbeing, such as health, education, and social connections. Recent developments in the "beyond GDP" debate point to many improvements in the measurement of these other dimensions of wellbeing, which are beyond the scope of this study; however, GDP per capita continues to be a key determinant of wellbeing, (OECD, 2014, 2021).

3 | THE BENCHMARK OPTIONS

When using official or directly constructed historical benchmarks in combination with time series to construct international long-run comparable income series, there are, in principle, three alternatives to the Maddison Project for combining recent PPPs with time series of GDP growth. These are: one, making use of recent, finalized PPPs (the 2011 PPPs in this case)[3]; two, sticking to the original 1990 PPPs; or three, using all available ICP benchmarks (the multiple benchmark approach). To compare the implications of each scenario for reconstructing global long-run development, we have defined two criteria to test the accuracy (or the impact of the cumulative biases) of each option.

First, we establish the extent to which the combinations of benchmarks and time series result in plausible historical estimates, such that, for example, countries do not end up with incomes (far) below subsistence level for prolonged periods of time. To assess this, we find the number of below-subsistence estimates using our three scenarios and analyze the "plausibility" of those below-subsistence incomes relative to the quality of the time series on which the observations are based.

Second, we compare the relative performances of countries during the 19th and early 20th century as given by the alternative scenarios described above, with the relative performance of countries indicated by independent, directly measured historical benchmarks. For this, we have collected a dataset of all independent historical inter-country benchmarks of relative GDP per capita levels available in the literature. This comparison can tell us which benchmark scenario most closely aligns with these independent estimates of relative output (GDP per capita) levels, and to what extent these different scenarios result in different stories of global development. Additionally, we examine how the three benchmark scenarios compare with the results of the econometric (indirect) method as proposed by Prados de la Escosura (2000), which can be seen as an alternative source of estimates of relative income levels in the 19th and early 20th century.

3.1 | The "original" Maddison 1990 benchmark

Maddison created a benchmark for the reference year 1990, which became the "interspatial-intertemporal anchor" for his famous global estimates of GDP. In practice, Maddison could not always use PPPs based on 1990 relative prices, because many countries were not covered by the PPP exercise in 1990; for those countries, he had to link to earlier ICP rounds (Maddison, 2006, p. 610). In his 1995 study, Maddison made use of the 1990 ICP 6 round for the 22 member countries of the OECD and presented the results together with other ICP rounds. Many non-OECD countries were covered in one of the other ICP rounds—for example, in 1975, 1985, or 1993. In

the final comparison, data for 43 countries (representing almost 80% of world GDP at the time) was based on ICP or ICP-equivalent estimates. The other 113 countries in the comparison ("non-sample") were covered by PPPs from PWT and by proxy estimates. The PWT PPPs were, in turn, estimated by Summers and Heston (1991) based on cost-of-living estimates for expatriates and foreign diplomats.

Maddison showed that using alternative estimates from different ICP rounds for non-OECD economies leads to larger differences than using alternative estimates for OECD countries. The range between different converters is generally larger in the non-OECD countries because their price structures differ more from those in the United States. China deserves special mention here, due to its size and because China did not participate in any of the ICP rounds before 2005. Maddison collected the data for his 1990 reference estimate from four different estimates of 1990 GDP per capita in GK dollars, ranging from $1135 to $4264. Later on, Maddison adjusted the Chinese per capita GDP level to $1,871 in 1990, based on an estimate for China by Ruoen (1997) to which he made significant adjustments (Maddison, 2007, pp. 154–155).

Even though there had been a new ICP round (ICP 1993[4]), Maddison continued to use 1990 as the reference year for his comparisons, as he saw considerable problems in putting together all the regional ICP estimates on a comparable basis. In his view, it would have been a complex exercise to update reference years on a consistent basis (Maddison, 2006, p. 172).

To summarize, the set of 1990 "Geary-Khamis PPPs" that serves as the "anchor" for the Maddison data has shortcomings due to data limitations. Most notably, data from actual ICP price surveys was available for fewer than one-third of the countries in his dataset, with PPP estimates for the remainder based on surveys designed for purposes other than comparing GDP levels across countries. For the countries with ICP data, the PPPs are drawn from several different benchmark estimates, often years apart, which all used different methods in collecting prices. Although Maddison's set of 1990 PPPs may have represented the best estimate at the time, measurement methodologies have since moved on.

3.2 | From 1990 to the 2011 global price comparison

In retrospect, the period during which Maddison developed his dataset was a low point in the history of the International Comparison Program: lacking sufficient funding and suffering from poor management, the 1993 round of ICP was only partially completed and widely regarded as imperfect (Ryten, 1998). This provided the impetus for a better-funded, more widely supported continuation of ICP, which led to the 2005 round of ICP. Compared to earlier global comparisons, this round was a major improvement, covering more countries than ever before (146), including China and India. Furthermore, the round featured more precise product specifications and a regional setup. Under the regional setup, countries would first compare prices within a region, based on a regional product list, and then different regional comparisons would be linked based on prices from a global price list to allow for a global price comparison. Both improvements should help mitigate one of the key challenges in international price comparisons—namely, to find and price products that are comparable across countries, yet representative of what consumers buy in each country. A more extensive discussion, as well as a critical reflection on measurement and methods in ICP 2005, can be found in World Bank (2008).

But although ICP 2005 represented an important step forward, concerns quickly arose about the results. Most notably, the relative income levels of lower-income countries, like China and India, were notably lower than had previously been thought. While that could have been the result

of improved measurements in ICP 2005, Deaton (2010) and Deaton and Heston (2010) raised the possibility of biases in measurement—particularly in regard to price surveys in China, which were only carried out in urban areas. If prices in rural areas are lower, this would introduce a source of bias relative to the "national average" prices used in other countries: see also Feenstra et al. (2013). Deaton (2010) and Deaton and Heston (2010) were also concerned about the second stage of ICP 2005, in which regional comparisons are linked, since any misalignment in this linking procedure would shift prices and income levels of whole regions relative to each other; Deaton (2010) refers to these as "tectonic regional PPPs." More recently, Inklaar and Rao (2017) have shown that the regional linking in ICP 2005 was indeed biased and that, as a result, the relative prices in African and Asian countries were too high, and thus their relative income levels too low.

Mindful of the weaker points of ICP 2005, measurement methods were modified and refined for ICP 2011 (see World Bank (2015)). Most notably, a more representative global product list was drawn up to be used in estimating the "tectonic regional PPPs." As shown in Inklaar and Rao (2017), the resulting regional linking in ICP 2011 does not suffer from the biases that were present in ICP 2005. Surveys in China were also more extensive, covering urban and rural areas in all provinces (Asian Development Bank, 2014). Finally, coverage was even more comprehensive, with complete data for 199 nations.[5] There are certainly areas of imperfect measurement: the prices for construction investment, housing, government, health, and education services are viewed as "comparison-resistant." The "output" of these activities is typically heterogeneous and difficult to define and/or price, so proxy methods are chosen (see World Bank (2015)). Still, these areas have been challenging since the start of ICP (see, e.g., Kravis et al. (1982)) and there is no indication that these measurement limitations have worsened. In summary, the ICP 2011 PPPs represent the best-measured set of relative prices for the largest set of countries finalized to date.[6] In terms of measurement methods, ICP 2011 is better—or at least no worse—in all areas.

3.3 | The PWT multiple benchmark approach

The third alternative that we explore in this paper is the methodology developed for version 8 of the Penn World Tables (Feenstra et al., 2015). This approach makes use of all available post-1950 benchmarks and adapts the growth rates of GDP from SNA to fit the changing relative position of countries as given by the various benchmarks. There have been global ICP rounds collecting prices in countries in the benchmark years 1970, 1975, 1980, 1985, 1993, 2005, and 2011. Over the years, an increasing number of countries has participated in these global collection rounds, with 199 participating countries in 2011 being the most comprehensive round (see Figure 1).

For each individual country, the PWT determines which benchmarks are available. These benchmarks are used directly as PPP converters. For the years between benchmarks, prices for final goods are interpolated by taking the benchmark price indexes as given and using the pattern of inflation from the national accounts in the years between benchmarks to determine how the overall change in price indexes between these benchmark years should be distributed (Feenstra et al., 2015, p. web appendix B: 5). For the years between 1950 and the first available benchmark for each country, prices are extrapolated using price deflators from the National Accounts. The same is done for years following the most recent benchmark (Feenstra et al., 2015, p. 3167). For the purpose of this paper, the 1950 estimates obtained for each country using this methodology serve as the anchor from which income estimates are extrapolated backwards in time using growth rates from existing series or from the newly included series as discussed below.[7]

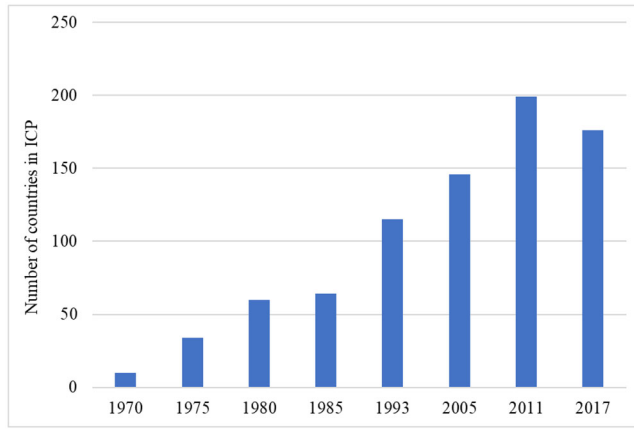

FIGURE 1 Historical global ICP participation. *Source*: https://www.worldbank.org/en/programs/icp/history. [Colour figure can be viewed at wileyonlinelibrary.com]

The PWT procedure gives a complete set of price indexes for all years and countries included in the dataset, providing the "best" estimate of the PPP for each year. The consequence of fitting the national income series to the various benchmark estimates is that the resulting series are no longer equal to the original official National Account series as a result of the inherent inconsistency between benchmarks and time series. The resulting pattern of income growth as published by the PWT sometimes differs from income growth as given by the National Accounts. For example, the decline in GDP per capita in Nigeria between 1977 and 1995 is much larger using the PWT approach than can be found in the official sources (other examples of substantial differences in growth rates between the PWT and National Accounts are Togo and Madagascar between 2005 and 2011).

4 | COMPARING THE THREE APPROACHES

In this section we compare the three benchmark options discussed above in terms of their ability to "predict" historical GDP estimates, or put differently, how closely relative income levels from the three scenarios are aligned to the direct historical benchmark estimates. Our first set of long run GDP per capita estimates link all GDP per capita series to the 1990 benchmark (the approach developed by Maddison and used in the original Maddison Database). The second set of estimates uses the 2011 PPP relative income estimates from PWT 9.0,[8] resulting in GDP per capita series expressed in 2011 (relative) prices; the third approach links the long run GDP per capita series to the (implicit) 1950 benchmark created by PWT (Bolt et al., 2018 for details on the latter). The series for individual countries are derived from the updated Maddison project dataset as presented below in the paper.

Linking alternative benchmark years to the original GDP per capita series involved two steps. We first calculated yearly growth rates of the original GDP series, based for each country on National Account Statistics and reconstructions of historical national accounts (resulting in estimates of the evolution of GDP and/or GDP per capita). Second, we used these growth rates to extrapolate the income series backwards using either the 2011 PPP converted income levels taken from the PWT, or to extrapolate income series backwards using the multiple benchmark approach (MBM) as employed by the PWT (Feenstra et al., 2015)—similar to Maddison (2003) and the 2013 Maddison Project update, which used the 1990 benchmark (Bolt & van Zanden, 2014).

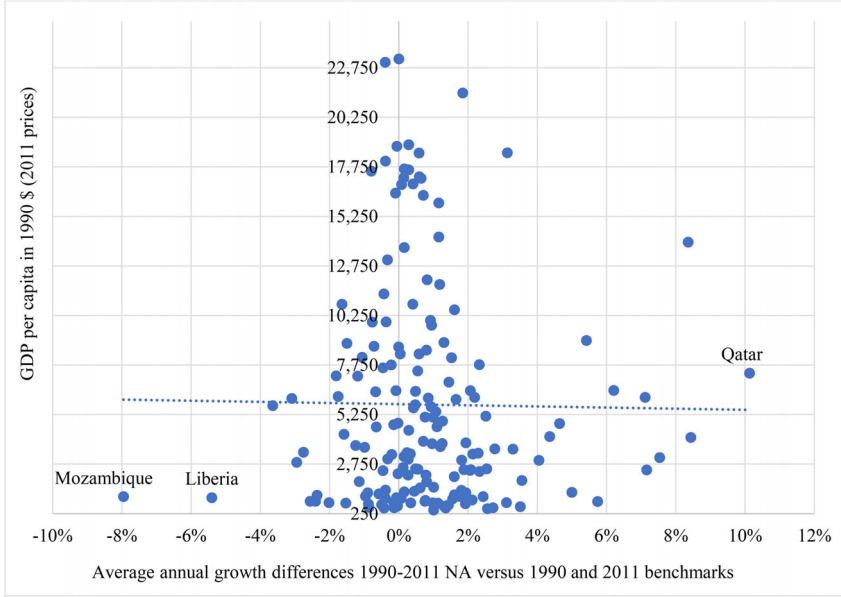

FIGURE 2 1990 income levels versus differences in growth rates 1990–2011. Comparison of GDP per capita in 1990 (1990 PPP $, 2011 prices; vertical axis) and the difference between growth rates obtained from the 1990 and 2011 benchmark levels of GDP, and the national accounts growth rates. [Colour figure can be viewed at wileyonlinelibrary.com]

The outcome of these rebasing exercises essentially entails a level shift in income estimates, depending on the relative income level for each country according to various benchmarks. As an example, for countries that experienced an increase in prices between 1990 and 2011 vis-à-vis the base country (the United States), original income series have been divided by higher relative prices in the 2011 PPP series, leading to a lower level of real income relative to the US. Conversely, for countries that have experienced a slower increase (or even a decrease) in prices between 1990 and 2011 relative to the US, income estimates shift upwards (as they are divided by lower relative prices in the 2011 PPP series). Similarly, for countries that have experienced an increase in prices vis-à-vis the US between the first benchmark available for that country and 1990, the level of real income relative to the US is lower in 1990 than their relative income level for the first benchmark year (as their incomes are divided by higher prices in the 1990 series). Using different benchmarks as the basis for long run income series leads thus to a reshuffling of the level of historical series for various countries and has consequences for important debates in economic history—such as when the US overtook the UK as the world-leading country, which we will go on to discuss in more detail.

In Figure 2, we illustrate these reshufflings between 1990 and 2011. On the y-axis, Figure 2 shows the level of GDP per capita using the 1990 PPP benchmark (expressed in 2011 dollars). On the x-axis, the figure shows the difference in average annual growth rate between 1990 and 2011 obtained from the benchmark levels of income versus using the national account growth rates (i.e., subtracting the average annual growth rate from the national accounts[9] from the average annual growth rate calculated between benchmark levels of income[10]). One of the clearest examples of a changing relative position in the global income distribution comes from Qatar, a country that was in the lower half of the income distribution in 1990 and by far the wealthiest country in 2011. This

shift in the income distribution appears, in Figure 2, through the large positive average growth difference between the growth rate from the national accounts and the growth rate from benchmark to benchmark (10% higher annual growth rate between benchmarks). At the other extreme, Mozambique and Liberia, two countries with extremely low levels of real income in 1990, had very negative growth differences between the benchmarks and the national accounts. Although both countries experienced some growth in the national account series, they did not manage to increase their relative and absolute income level (the 2011 benchmark incomes for both countries are lower than their 1990 benchmark incomes).

Another way of presenting the changes between benchmarks is by comparing the consequences of different benchmarks for historical income estimates. Figure 3 compares incomes based on the original 1990 PPPs to incomes based on the 2011 PPP benchmark for the years 1820, 1870, 1913, and 2011. In the 2011 panel, the most notable outliers represent oil-rich countries in the Middle East: Qatar, Kuwait, United Arab Emirates, Bahrein, Oman, Saudi Arabia, Iraq, and Syria. Their relative GDP per capita increases substantially when the 2011 benchmark is used rather than the 1990 benchmark. This may be explained by deficiencies in the Maddison 1990 estimate. These countries were in the original Maddison database as part of the "non-sample" group; that is, they were not included in the ICP rounds underlying Maddison's 1990 PPP calculations. Prevailing price estimates for 1990 were based on expat information and likely overestimated actual prices in these oil rich economies, leading to an underestimation of actual income levels. Using the more precise price estimates of the 2011 ICP rounds shows a substantial upward shift in incomes for these countries. The shift in the GDP levels of the Gulf States may also be caused by the significant increase in the price of oil between 1990 and 2011, when oil prices peaked; this shift drove up the relative income levels of these oil producing countries.

Extrapolating the very high 2011 incomes of these oil-rich countries back into the past on the basis of the estimated (NA) growth rates of these countries may therefore result in implausibly high income estimates for earlier periods, as Figure 3 demonstrates. For example, using the 2011 benchmark as the anchor from which to extrapolate backwards results in a model that posits Iraq as the wealthiest country in the world in 1820 (and close to that in 1870). Other notable outliers are Norway, Switzerland (1870 and 1913), and Lebanon (1820, 1870, and 1913). Outliers on the other extreme are the North Korea, Syria, and the Palestinian territory in 1820, 1870, and 1913, and Mozambique, Belarus, Armenia, Estonia, Trinidad and Tobago in 2011. The explanation for the group of negative outliers is not so obvious as the group of countries is quite heterogenous compared to the positive outliers (made up of the oil-rich countries). The difference between the two benchmarks is largest in 1820 and 2011 (the lowest R2 from the regression lines).

In Figure 4, we compare the 1990 benchmark to the MBM using the PWT approach. Again, as we discussed in regard to Figure 3, we find that for 2011, the most obvious outliers are the countries in the Middle East (most prominently Qatar), which show a much higher income when the MBM approach is used rather than the 1990 benchmark. Most countries in the Middle East did not participate in earlier ICP rounds, and therefore the MBM approach only integrates recent benchmarks (mostly 2005 or 2011) for these countries, leading to very similar results between the MBM series and the 2011 benchmark (see also the discussion around Figures 2 and 3). Countries that clearly fall below the regression line in 2011 are also, for this comparison, a less coherent group. They range from Mozambique, Palestine, and Syria to Armenia, Mauritius, Trinidad and Tobago, and Hong Kong. There is no clear pattern that helps explain why these countries deviate from the general trend.

For the years 1820 to 1913, the most obvious outliers are Cuba[11] —which has much higher income—and Argentina—which has much lower income—in comparing the MBM approach to

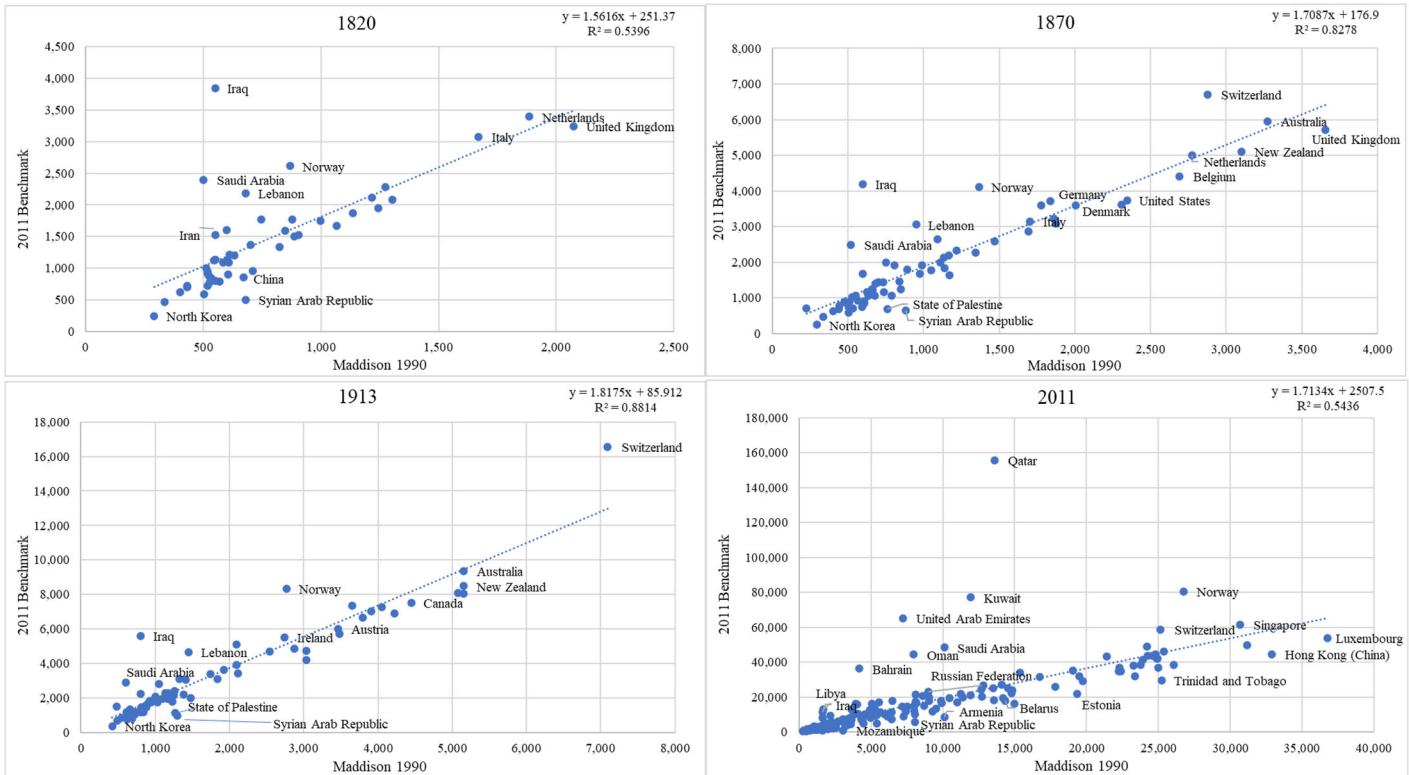

FIGURE 3 GDP per capita in various years based on Maddison 1990 versus ICP 2011 PPPs (in 2011 US dollars). The original GDP per capita values based on Maddison's 1990 PPPs have been converted to 2011 US dollar values (to make the numeraire comparable) by multiplying all observations by 1.59, which corresponds to US inflation between 1990 and 2011. [Colour figure can be viewed at wileyonlinelibrary.com]

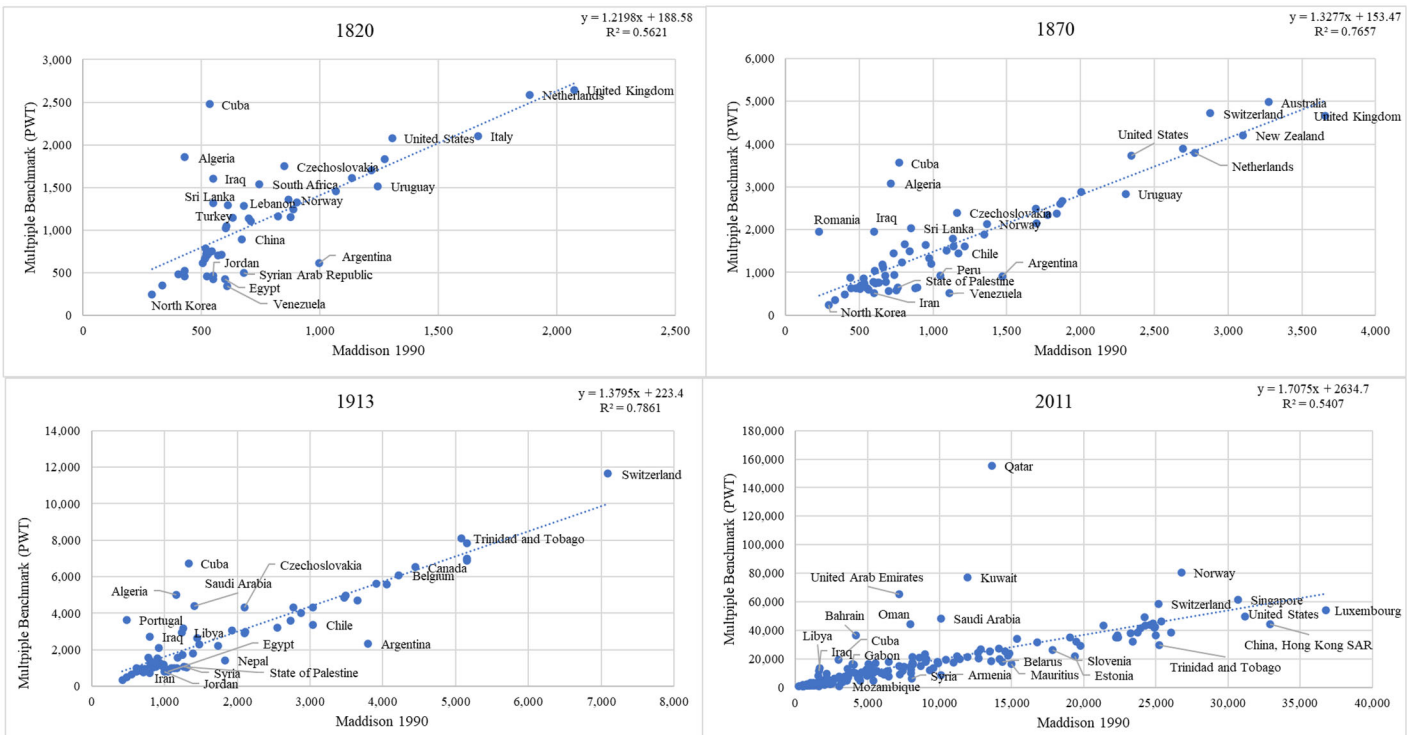

FIGURE 4 GDP per capita in various years based on Maddison 1990 versus multiple benchmark (MBM) approach (in 2011 US dollars). [Colour figure can be viewed at wileyonlinelibrary.com]

the 1990 benchmark level of income. Like Cuba, Algeria, Iraq, and Romania (1870) and Portugal and Saudi Arabia (1913) show substantially higher incomes when using the MBM approach rather than the 1990 benchmark. Similar to Argentina, incomes in Syria, Palestine, and Venezuela (1820 and 1870), Peru (1970 and 1913), Nepal (1913) and North Korea (most 1820, 1870, 1913) are also considerably lower when using the MBM approach rather than the 1990 benchmark. These results show that the choice of benchmark matters substantially when used to anchor historical income series, as different benchmarks sometimes lead to substantial reshuffling of groups of countries. Below, we discuss certain cases in which the choice of benchmark leads to implausibly low historical income estimates, such as for Peru during substantial parts of the 19th century, and for Korea prior to 1930.

4.1 | The first test: subsistence income

One of the implications of the rebasing exercise is that the poverty level has changed. With the 1990 price levels, subsistence level income was between 350 and 400 international dollars per year (Maddison, 2003). The poverty line therefore was equal to around 1 dollar a day, and was based on the first international poverty line, which was set at $1.01 per day using 1985 PPPs, and later updated to $1.08 per day using 1993 PPPs (Chen & Ravallion, 2001; Ravallion et al., 1991). This made the interpretation of historical income series very intuitive. By using other relative prices, this subsistence level of income changes. The World Bank defined the absolute poverty line at 1.90 US dollars a day, or 694 dollars per year, expressed in 2011 prices.

Rebasing the original Maddison estimates has notable effects for countries who experienced substantial price changes relative to the US between the benchmark years. China is an interesting case in point. When the 2005 PPPs were released, the prices for China had increased so much relative to the US that GDP per capita levels ended up around 40% lower than China's relative income based on earlier price estimates (Deaton & Heston, 2010, p. 3; Feenstra et al., 2013). This led to implausibly low historical income estimates for China, given that the original estimates were already very close to subsistence around 1950 (Maddison, 2007).

In the years after the release of the 2005 PPPs, a consensus arose about its shortcomings, most of which were corrected in the 2011 ICP round. Still, relative prices for China relative to the US were substantially higher in 2011 than in 1990, which lowers China's PPP adjusted income per capita in 2011 by 23%. Yet in this paper, we have updated the Maddison project income estimates for China based on Wu (2014), which shows a growth rate that is, on average, lower between 1952 and 2011 than in previous official estimates. Extrapolating backwards from the lower 2011 per capita income level using these lower growth rates leads to plausible historical income estimates for each alternative benchmark PPP (1990, 2011, and the multiple benchmark approach): see Figure 5. The 2011 benchmark gives a substantially lower level of average income due to the rapid increase in prices in China over the recent years. The MBM approach leads to results similar to the 2011 benchmark for the most recent years but deviates from that benchmark and comes closer to the 1990 benchmark for earlier years.

When we look more broadly into subsistence incomes, we see that the original Maddison project dataset includes 184 observations below the subsistence income level of 400 dollars per year (1% of a total of 17,872 observations), most of which are countries in times of civil war, such as Afghanistan, Liberia, Burundi, and the Democratic Republic of Congo. Moving towards the 2011 PPPs, the poverty line increases to 694 dollars (using 2011 prices), and the number of below-subsistence observations increases to 312 (1.8%). Using the Multiple Benchmark Method (MBM),

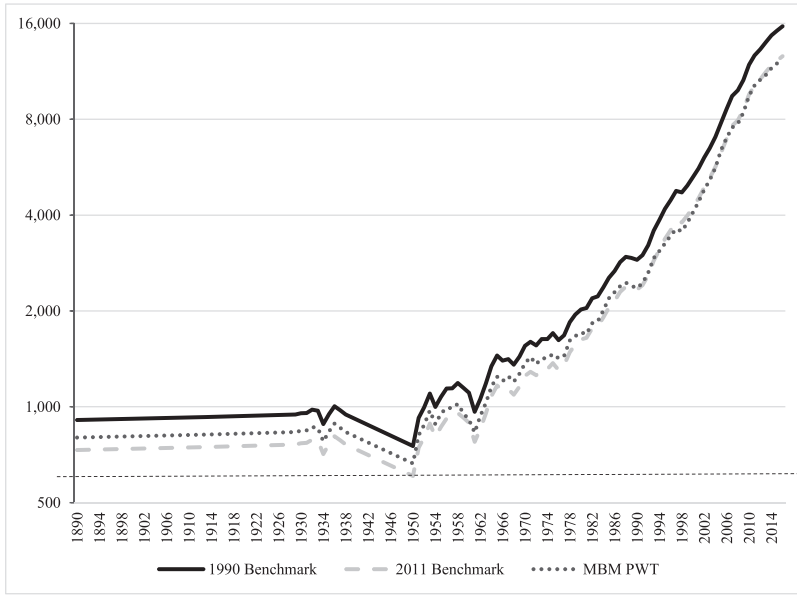

FIGURE 5 Historical income series for China using alternative benchmark PPPs. Note: The absolute poverty line is based on the World Bank poverty line of $1.90/day in 2011 US.

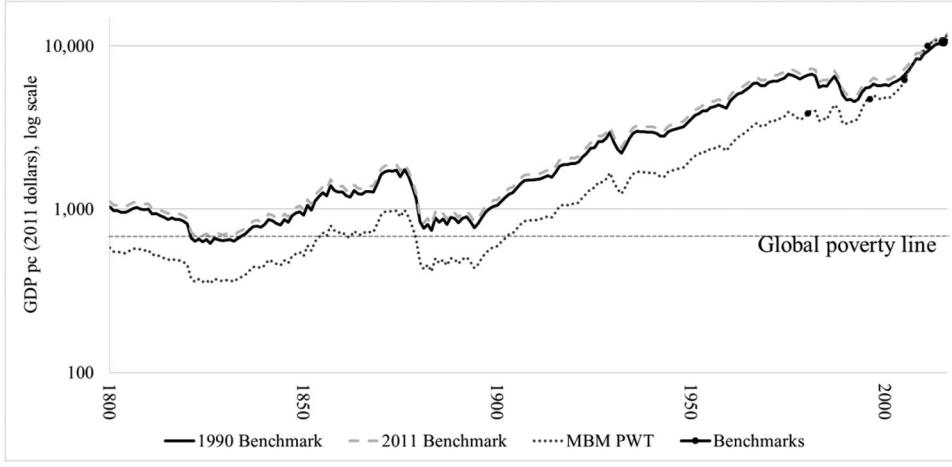

FIGURE 6 Historical income series for Peru using alternative benchmark PPPs. Note: absolute poverty line is based on World Bank poverty line of $1.90/day in 2011 US.

the number of observations below subsistence increases even more, to 386 (2.2%), and this includes a number of countries we would not expect: Peru during much of the 19th century (see Figure 6), Egypt and Chile in 1820, or Korea during most of the pre-1930, for example. Given the quality of the income series for some of those countries (the Peruvian historical series, e.g., is one of the best available, see Seminario (2015)), concerns arise about the quality of the earlier benchmarks used to calculate—in combination, of course, with the available time series of GDP per capita—these very low real income levels. The problem with the MBM is that the historical series are linked to

the earliest available benchmark of the ICP (of 1960 or 1970, for example), which may, in view of the further development of the project, have been rather crude.

4.2 | Comparing the accuracy of backward projection versus benchmarks

Another way to gauge the plausibility of the three benchmark scenarios for extrapolating income series backwards is to compare countries' relative income levels from the benchmark scenarios to available independent benchmark estimates from the literature. This comparison also confronts a common criticism of Maddison's work, as his dataset is based on the assumption that relative and absolute levels of real income could be measured in the relative prices of the base year 1990. Changes in relative prices were bound to lead to biases in the measured levels of GDP per capita; the longer the time period and the more dramatic the changes, the more problematic the approach would be. With the comparison between the benchmark scenarios and the independent benchmarks, we can actually assess to what extent the criticism is justified.

In this section, we therefore "test" the backward projections of the three scenarios (making use of the 1990, the 2011, and the multiple benchmarks, respectively): firstly, to see which method results in the best "predictions" of relative income levels in the 19th and early 20th centuries compared to the direct benchmarks from the literature; but also to discover the size of the biases resulting from the Maddison approach. In other words, do the critics have a point, or do changes in relative prices average out in the medium-long run? We developed two ways of testing the accuracy of the backward projections. First and foremost, we collected all independent benchmarks available in the literature. Scholars such as Ward and Devereux (2021), Broadberry (2006), and Fremdling (1991) have made detailed comparisons of the differences between GDP per capita in, for example, the UK and the US in 1872 and 1910 and between Germany and the UK in 1910 and 1935, which can be compared with the results of the three backward projection scenarios making use of recent PPPs. For all the independent benchmarks, we have calculated relative incomes for the same countries and years from the time series using the three different scenarios. Second, we have also gathered indirect (econometrically derived) estimates of relative income levels in the 19th and early 20th centuries from Prados de la Escosura (2000) in order to compare that to the relative income incomes for the same countries and years using time series for the three different scenarios.

Before we discuss the comparisons, it is important to realize that the price indexes underlying our income series on the one hand and the historical direct benchmarks and the indirect benchmark of Prados de la Escosura (2000) use different methodologies, which affects the comparison. Maddison's 1990 numbers were based on PPPs calculated using Geary-Khamis index numbers for 1990 or close to it, and the 2011 benchmark and the MBM approach we use in this paper are also expressed in Geary-Khamis dollars.[12] The earlier, pre-1950 income comparisons are often bilateral comparisons, or, for multi-country comparisons, cover only a limited number of countries. This explains why those comparisons tend to use bilateral Fisher indexes with the reference country (e.g., the US or UK) (Prados de la Escosura, 2000). Prados de la Escosura's (2000) indirect method for estimating historical income levels based on the contemporaneous price-income relationship is not based on an explicit index number approach, which also affects the comparison with the PPP benchmarks. When relative prices or the structure of economies differ across countries, different index number methodologies provide different outcomes and therefore a margin of statistical uncertainty (Deaton & Heston, 2010, table 2, p. 12). However, given that our three

TABLE 1 R2 of the regression line between scenarios and historical benchmarks.

	1990 benchmark	2011 benchmark	MBM (PWT)
Original dataset 2020[a]	0.65	0.57	0.58
All available benchmarks[b]	0.41	0.28	0.44
Adding Clark[c]	0.48	0.44	0.44
Adding Lindert and Williamson[d]	0.63	0.54	0.57
Adding Ward and Devereux 2021[e]	0.41	0.29	0.49

Notes: The R2 is taken from a linear regression line. We add the various studies separately.

[a]Dataset used in the 2020 Maddison Project Update, see Bolt and Van Zanden (2020).

[b]See appendix 1 for full list of sources.

[c]The 1929 benchmark from (Clark, 1932).

[d]Lindert and Williamson (2016a).

[e]Ward and Devereux (2021). The R2 for 2017 PPP benchmark and the independent benchmarks is 0.22 when including all available benchmarks.

scenarios are based on similar methodologies (GK dollars), the bias in the comparison with benchmarks using other index numbers (both historical benchmarks and the indirect benchmarks) would most likely also be relatively similar for the different scenarios and therefore not drive our results.

Table 1 shows the results of our comparisons expressed as the R2 of the linear regression model between the three scenarios for back-projections and the independent benchmark estimates from the literature. The first row in the table shows the R2 for the "original" benchmark dataset that was used in the 2020 update of the Maddison project (Bolt & Van Zanden, 2020). This shows that the 1990 Maddison benchmarks "outperforms" the two other approaches, with the 2011 ICP presenting the lowest R2. Expanding the analysis to include more recent works (and important historical work we previously overlooked) lowers the R2 for all scenarios, but also reverses the pattern as the MBM scenario scores a little better compared to the 1990 benchmark. To explore the influence of the various benchmark studies on the overall result, we add the most prominent studies separately to the 2020 dataset. Adding the benchmarks of the seminal work by Clark (1951) reduces the R2's but does not change the order in best performing scenarios with the 1990 with the highest R2 score. The differences between the different scenarios becomes much smaller though. This might be due to the fact that some of the Clark (1951) benchmarks are debatable—his benchmark for the ratio between South Africa and the UK was close to 1 which is much too high.[13] Adding Lindert and Williamson (2016b) results in a return to the results of the original sample, see Table 1. However, adding Ward and Devereux (2021) has a strong impact on the results, and causes the reverse in patterns found when we use all available benchmarks (row 2). Including Ward and Devereux (2021) (see row 5 in Table 1) leads to a higher R2 for the MBM by some margin to the original 1990 benchmark, while the 2011 benchmark performs significantly worse.

The second comparison we make is to the indirectly estimated benchmarks by Prados de la Escosura. This indirect method provides alternative real historical income levels by constructing (implicit) PPP converters based on an estimated relationship between price levels (PPP/Exchange rate ratios) and nominal product per head (exchange rate converted GDP pc) and a set of additional explanatory variables (Prados de la Escosura, 2000, p. 4). This is one of the ways to deal with the disadvantage of the extrapolation method which accepts one point of reference that could lead to distortions and ambiguities further away from the reference year. When we compare income levels of countries relative to the US as produced by Prados de la Escosura (2000) to the relative income levels as generated by the extrapolation method using the three scenario's for backward

TABLE 2 R2 of the regression line between scenarios and indirect benchmarks.

	1990 benchmark	2011 benchmark	MBM (PWT)
Indirect benchmarks	0.61	0.46	0.57

Source: the indirect benchmarks are taken from Prados de la Escosura (2000). The R2 for 2017 PPP benchmark and the indirect benchmarks is 0.39.

projection, we get a pattern similar to that based on the independent benchmarks (Table 2). The scenario based on the 1990 benchmark has the highest R2, followed by MBM (where the difference between the two is similar in magnitude but reversed in order with the 1990 benchmark having the highest R2), and the 2011 scenario has the lowest R2.

The conclusion we reach is that simply moving from the 1990 to the 2011 benchmark does not improve estimates for the historical periods, probably due to the greater distance between years when using 2011. The MBM approach, which uses all post-1950 benchmark estimates, does not suffer from the same bias; it does, however, result in a rather high share of countries with below-subsistence levels of GDP per capita, which is also problematic. Our explanation of these results is that the quality of the benchmarks is likely subject to two changes: one, the quality of the PPPs and the coverage of the various ICP rounds increases over time; two, the distance from the "historical" period in question increases over time as well, resulting in more biases due to changes in relative prices. Ergo, both a "perfect" but "distant" benchmark such as 2011 and the "less than perfect" but "early" benchmarks used by MBM approach have sizeable biases. It follows that Maddison's 1990 benchmark may well be the best compromise solution. It certainly produces the best results between the two tests that we carried out: it predicts historical benchmarks better than the 2011 benchmark (see Tables 1 and 2), and it does not result in a high share of countries with an implausibly low income level (below subsistence) as the MBM scenario does. This assessment therefore results in the conclusion that the best way forward is to stick to the 1990 benchmark for the overall architecture of the dataset.

4.3 | Integrating more benchmarks

In two ways, however, this analysis leads to departures from the original Maddison approach and closer to the multiple benchmark approach as developed by the PWT. There is, to begin with, no doubt that the 2011 PPPs and the related estimates of GDP per capita reflect the relative levels of GDP per capita in the world economy today better than the combination of the 1990 benchmark and growth rates of GDP per capita according to national accounts. This information should be taken into account. At the same time, the underlying rule within the current Maddison Database is that economic growth rates of countries in the dataset should be identical or as close as possible to growth rates according to the national accounts (which is also the case for the pre 1990 period). For the post-1990 period we therefore decided to integrate the 2011 benchmarks by adapting the growth rates of GDP per capita in the period 1990–2011 to align the two (1990 and 2011) benchmarks. We estimated (as shown in Figure 2) the difference between the combination of the 1990 benchmark and the growth rates of GDP (per capita) between 1990 and 2011 according to the national accounts, and annual growth rate from the 1990 benchmark to the 2011 benchmark. This difference is then evenly distributed to the growth rate of GDP per capita between 1990 and 2011; in other words, we added a country specific correction (constant for all years between 1990 and 2011) to the annual national account rate of growth to connect the 1990 benchmark to the 2011

benchmark. Growth after 2011 is, in the current update, exclusively based on the growth rates of GDP per capita according to national accounts.[14]

We also use the collected set of historical benchmark estimates to fine tune the dataset for the pre-1940 period, but only in those cases where the quality of the benchmark was high and there were multiple benchmarks to support a revision. The most important correction concerns the US/UK comparison. The conventional picture, based on the original 1990 Maddison estimates, indicated that the US overtook the UK as the world leader in the early years of the 20th century. This finding was first criticized by Ward and Devereux (2003), who argued, based on alternative measures of PPP-adjusted benchmarks between 1870 and 1930, that the United States was already leading the United Kingdom in terms of GDP per capita in the 1870s. This conclusion was criticized by Broadberry (2003).

New evidence, however, suggests a more complex picture: in the 18th century, real incomes in the US (settler colonies only, not including indigenous populations) were probably higher than those in the UK (Lindert & Williamson, 2016a). Until about 1870, growth was both extensive (incorporating newly settled territory) and intensive (considering the growth of cities and industry at the east coast), but on balance, the US may—in terms of real income—have lagged behind the UK.[15] After 1870, intensive growth becomes more important, and the US slowly gets the upper hand. This pattern is consistent with direct benchmark comparison of the income of both countries for the period 1907–1909 (Woltjer, 2015). This shows that GDP per capita for the United States in those years was 26% higher than in the United Kingdom. We have used Woltjer's (2015) benchmark to correct the GDP series of the two countries. Projecting this benchmark into the 19th century with the series of GDP per capita of both countries results in the two countries achieving parity in 1880. This is close to Prados de la Escosura's conjecture based on his shortcut method (Prados de la Escosura, 2000), and even closer to the Lindert and Williamson (2016a) results.

Changing the US/UK ratio on the basis of the new research by Woltjer (2015) raises the question of which country's GDP estimates should be adapted. In the current PWT approach, the growth of GDP per capita in the United States is the anchor for the entire system. For the 19th century, however, it is more logical to take the United Kingdom as the anchor, because it was the productivity leader, and because most research focused on creating historical benchmarks takes the United Kingdom as reference point. We have therefore adapted the UK series for the period 1908–1950 to fit the 1907–09 (Woltjer, 2015) benchmark in our view the best available benchmark for this period. The reason is that there are doubts about the accuracy of price changes and deflators for the UK for the period 1908–1950, given that it was characterized by two significant waves of inflation (during the two World Wars) and by large swings in relative prices and exchange rates (as documented in the detailed analysis by Stohr (2016) for Switzerland). Future research will have to assess whether this choice is justified.

5 | UPDATING HISTORICAL SERIES[16]

This new version of the MPD extends GDP per capita series to 2022 and includes all new historical estimates of GDP per capita over time that have become available since the 2013 update (Bolt & Van Zanden, 2014). As new work on historical national accounts appears regularly, a frequent update to include new work is important, as it provides us with new insights in long-term global development. Furthermore, we have incorporated all available annual estimates for the pre-1820 period instead of estimates per half-century, as was usual in the previous datasets.

A general "warning" is in place here. For the period before 1900 (and for parts of the world such as Sub-Saharan Africa before 1950), there are no official statistics that fully cover the various components of GDP; and the more one moves back in time, the more a scarcity of basic statistics becomes a problem for scholars trying to chart the development of real income and output. The statistics needed for reconstructing GDP are often produced in parallel to the process of state formation, but even large bureaucratic states such as China or the Ottoman Empire only rarely collected the data that allow us to estimate levels of output and income. Much of the work on pre-industrial economies makes use of the "indirect method," which links data on real wages and levels of urbanization to estimates of GDP per capita. But a few countries, during the Medieval and Early Modern periods, did collect the (tax) data to estimate GDP in the "proper" way (Tuscany in 1427, Holland in 1514, and England in 1086). These benchmarks, in combination with the many "indirect" estimates, allow us to create a tapestry of estimates which becomes—with the increase of the number of studies—increasingly robust. Where the original Maddison dataset included 158 observations for the pre 1820 period, the current 2023 MPD includes close to 2800 data points for the preindustrial period.

For the recent period, the most important new work is Harry Wu's reconstruction of Chinese economic growth since 1950. Inspired by Maddison, Wu's model produces state of the art estimates of GDP and its components for this important modern economy (Wu, 2014). Given the large role China plays in any reconstruction of global inequality, this is a major addition to the dataset. Moreover, as we will see below, Wu's revised estimates of annual growth are generally lower than the official estimates. Lower growth rates between 1952 and the present, however, substantially increases the estimates of the absolute level of Chinese GDP in the 1950s (given the fact that the absolute level is determined by a benchmark in 1990 or 2011). This helps to solve a problem that arises in switching from the 1990 to the 2011 benchmark: namely, that when using the official growth estimates, the estimated levels of GDP per capita between 1890 and the early 1950s are substantially below subsistence level, and therefore too low. Including the new series as constructed by Wu (2014) gives a much more plausible long-run series for China.

Most of the other additions to the Maddison project dataset are contributions to the period before 1914; see Table 3. Again, important new work has been done for China, in particular the papers by Xu et al. (2017) and Broadberry et al. (2018). It is reassuring that these two independent teams of scholars who set out to quantify Chinese economic growth before 1900 produced very similar estimates, showing a strong decline (by about one third) of GDP per capita in the 18th century and quasi-stability in the 19th century. We are now for the first time also able to include estimates that measure incomes for nine Sub-Saharan African countries. For eight countries the estimates cover the colonial period (1890–1950) (Broadberry & Gardner, 2022); for Liberia, the estimates cover 1845–1950 (Gardner, 2022).

Often, studies producing very early per capita GDP estimates—particularly work on the early modern period (1500–1800)—make use of indirect methods. The "model" or framework for making such estimates is based on the relationship between real wages, the demand for foodstuffs, and agricultural output (Álvarez-Nogal & De La Escosura, 2013; Malanima, 2011 among others). This model has now also been applied to Poland (Malinowski & van Zanden, 2017), Spanish America (Abad & van Zanden, 2016), and France (Ridolfi, 2017; Ridolfi & Nuvolari, 2021). In this update, we have now included annual estimates of GDP per capita in the period before 1800 for these countries.

For some countries during a period before 1870 or 1800, we only have series of a certain province or similar entity. The British series links to estimates for only England for the period before 1700; the series for the Netherlands links to estimates for only Holland for the period before 1807.[17]

TABLE 3 New additions to the Maddison project database.

Country	Period	Source
Latin America		
Bolivia	1846–1950	Herranz-Loncán and Peres-Cajías (2016)
Brazil	1850–1899	Barro and Ursúa (2008)
Chile	1810–2004	Díaz Lüders et al. (2016)
Cuba	1690–1895	Santamaria Garcia (2005)
Cuba	1902–1958	Ward and Devereux (2012)
Mexico	1550–1812	Arroyo Abad and Van Zanden (2016)
Mexico	1812–1870	Prados de la Escosura (2009)
Mexico	1870–1895	Bértola and Ocampo (2012)
Mexico	1895–2003	Barro and Ursúa (2008)
Panama	1906–1945	De Corso and Kalmanovitz (2016)
Peru	1600–1812	Arroyo Abad and Van Zanden (2016)
Peru	1812–1870	Seminario (2015)
Uruguay	1870–2014	Bértola (2016)
Venezuela	1830–2012	De Corso (2013)
Europe		
England	1252–1700	Broadberry et al. (2015)
Finland	1600–1860	Eloranta et al. (2016)
France	1276–1860	Ridolfi and Nuvolari (2021)
Germany	1500–1850	Pfister (2022)
Holland	1348–1807	Van Zanden and van Leeuwen (2012)
Iceland	1870–1945	Jonsson (1998)
Italy	1400–1861	Chilosi and Ciccarelli (2023)
Norway	1820–1930	Grytten (2015)
Poland	1409–1913	Malinowski and Van Zanden (2017)
Portugal	1530–1850	Palma and Reis (2019)
Romania	1862–1995	Axenciuc (2012)
Romania	1915–1919 & 1949	Voinea et al. (2019)
Russia	1860–1885	Kuboniwa et al. (2019)
USSR	1860–1885	Kuboniwa et al. (2019)
Spain	1277–1850	Prados de la Escosura et al. (2022)
Spain	1850–2020	Prados de la Escosura (2017)
Sweden	1300–1560	Krantz (2017)
Sweden	1560–1950	Schön and Krantz (2015)
Switzerland	1850–2011	Stohr (2016)
UK	1700–1870	Broadberry et al. (2015)
Asia		
China	1000–1661	Broadberry et al. (2018)

(Continues)

TABLE 3 (Continued)

Country	Period	Source
China	1661–1933	Xu et al. (2017) and Broadberry et al. (2018)
China	1952–2008	Wu (2014)
India	1600–1870	Broadberry et al. (2015)
Japan	724–1874	Bassino et al. (2019)
Japan	1874–1940	Fukao et al. (2015)
Korea, Republic of	1911–1990	Cha et al. (2022)
Korea, North	1911–1940; 1990–2015	Cha et al. (2022)
Malaysia	1900–1939	Nazrin (2016)
Turkey	1500–1820	Pamuk (2009)
Singapore	1900–1959	Sugimoto (2011)
Middle East		
Syria	1820, 1870, 1913, 1950	Pamuk (2006)
Lebanon	1820, 1870, 1913, 1950	Pamuk (2006)
Jordan	1820, 1870, 1913, 1950	Pamuk (2006)
Egypt	1820, 1870, 1913, 1950	Pamuk (2006)
Saudi Arabia	1820, 1870, 1913, 1950	Pamuk (2006)
Iraq	1820, 1870, 1913, 1950	Pamuk (2006)
Iran	1820, 1870, 1913, 1950	Pamuk (2006)
Africa		
Ghana	1885–1950	Broadberry and Gardner (2022)
Kenya	1904–1950	Broadberry and Gardner (2022)
Malawi	1904–1950	Broadberry and Gardner (2022)
Nigeria	1885–1950	Broadberry and Gardner (2022)
Uganda	1906–1950	Broadberry and Gardner (2022)
Zambia	1906–1950	Broadberry and Gardner (2022)
Zimbabwe	1914–1950	Broadberry and Gardner (2022)
Liberia	1845–1950	Gardner (2022)
Cape Colony/ South Africa	1700–1900	Fourie and Van Zanden (2013)

The switch from the national to the "partial" series is clearly indicated in the dataset, and the "correction" in terms of GDP per capita is indicated.

Finally, we have extended the national income estimates up to 2022 for all countries in the database. For this we use various sources. The most important is the Total Economy Database (TED) published by the Conference Board, which includes GDP per capita estimates for a large majority of the countries included in the Maddison Project Database. The 2013 MPD update took the same approach (Bolt & van Zanden, 2014). For countries unavailable through TED, we relied on UN national accounts estimates to extend the GDP per capita series. To extend the population estimates up to 2022, we used the TED and the US Census Bureau's International Database 2022.[18] The TED revised their China estimates from 1950 onwards based on Wu (2014). As discussed above, we also included Wu (2014)'s new estimates in this update. Finally, we have extended the series for the former Czechoslovakia, the former Soviet Union, and former Yugoslavia, based on GDP and population data for their successor states.

5.1 | Long-run global economic development since 1820

The new updated database offers a more precise understanding than we previously had of the global economy in 1820, as well as the pre-1820 divergence in economic development, which is evident in the disparities in 1820. 1820 marks the initial year for which we can obtain levels of GDP per capita for all major regions worldwide, although Sub-Saharan Africa's estimates remain uncertain. Notably, substantial differences in GDP per capita existed between world regions in 1820, with Western Europe and the Western Offshoots having three times the real GDP per capita income of Eastern Europe or Sub-Saharan Africa, while the rest of the world was relatively close to the latter level. This gap, especially when comparing the most advanced economies, such as the UK and the Netherlands, to the rest of the world, results primarily from economic growth that preceded the Industrial Revolution in the North Sea region, where GDP per capita tripled between 1347 and 1820, due to a slow but consistent rates of economic growth. In contrast, the most advanced parts of the world in 1200 or even 1500, including China, India, and the Middle East, experienced long-term stagnation or decline in the centuries leading to the Industrial Revolution; only Japan witnessed a slow process of GDP growth similar to that of Western Europe. The early divergence in per capita GDP between Europe and Asia sparked a fierce debate about whether the level of economic development in China, India, and Japan before industrialization was comparable to that of Western Europe (Pomeranz, 2002). Our estimates here, along with our earlier research, indicates that GDP per capita in Western Europe was much higher than in the rest of the world in 1820 suggesting that Western Europe's economic lead predated the Industrial Revolution (Bolt & van Zanden, 2014).

Since 1820, the world has seen a spectacular increase in GDP. The global average income has increased nearly 13-fold between 1820 and 2020. And we can see that economic growth has accelerated over time. During the 19th century, it took nearly 100 years for the average global income to double, to a little more than 2000 dollars in 1900.[19] During the first half of the 20th century, it took only 60 years from 1900 to 1960, for the average global income to double again to 4500 dollars. A third doubling occurred before the turn of the 21st century (1960–1995). Since then, global economic development has slowed down only marginally: the average global income has increased 1.5 times over the past 20 years: over 14,000 dollars per person per year in 2020.

For much of the 19th century, reliable and comparative data for Africa remains scarce, and therefore this region does not feature prominently in our estimate of global development. However, during the 20th century, and particularly since 1950, all world regions are well represented and all regions show real, sometimes accelerating GDP per capita growth. The most rapid economic growth worldwide was experienced during what is known as the "golden period of growth" during the 1950s and 1960s. During these decades, many countries managed to catch up or started to catch up to the frontier economies.

The most dynamic parts of the world economy since the 1960s have been East Asia and, to a lesser extent, South and Southeast Asia, although not all countries within these areas have experienced economic gains throughout the current period. Japan, for example, closed the income gap with the frontier economies during most of the 20th century, but entered a phase of slow economic growth in the 1990s (although incomes in Japan are still 20% higher in 2020 than in 1990). Prominent exceptions to the trend of increasing incomes in this period were the economies in the (former) USSR and Eastern Europe during the 1990s when the centrally planned economies were dismantled, causing the GDP per capita in for example the former USSR to fall by roughly 30 percent. Further, Sub-Saharan Africa and Latin America also experienced economic decline during the "lost decades" of the 1980s and 1990s.

As already described by Maddison (2003), world leadership in terms of GDP per capita and labor productivity has changed only a few times. During the 17th and 18th century, the Netherlands was the global productivity leader until the United Kingdom took over as the frontier economy around 1780. World leadership changed hands again between 1870 and 1880 when the United States surpassed the UK both in terms of labor productivity levels and in terms of GDP per capita. As the United States continued to forge ahead of the rest of the world, the resulting Trans-Atlantic income gap widened substantially between 1900 and 1950. And although Western Europe started to catch up to US levels after 1950, a gap in GDP per capita remained, partly due to higher levels of labor force participation and working hours in the United States (OECD, 2021).

6 | CONCLUSION

The measurement of (historical) national accounts builds on a long scholarly tradition; the first attempts to measure a nation's total economic activity date back to the 17th century. We have surveyed the literature on the developments in historical national accounting, including an overview of the literature on benchmarks and their importance for comparing incomes across countries and over time. The Maddison Project is an ongoing research project aimed at standardizing and updating the academic work in the field of historical national accounting in the tradition of the synthesis produced by Angus Maddison (1995, 2003). As PPP benchmarks are central to the methodologies for global income comparisons over time, and therefore vital to Maddison Project Database, we discuss the consequences of the various PPP benchmarks available for historical income estimates. This paper complements the update presented in 2018 and 2020 (Bolt et al., 2018; Bolt & Van Zanden, 2020), and is meant to explain the choices made in the 2023 update and explore the implications of these choices for charting long-run economic growth.

The Maddison Project Database, like Maddison's original global database (1995, 2001), has used the 1990 PPPs as a benchmark for its long-term income series. The primary aim of the 2018 revision was to incorporate the newly published and arguably better new benchmark of 2011 PPPs. However, the switch to the new benchmark had substantial and often unexpected consequences for the picture of long-run global development. We therefore felt it important, in this paper, to take a step back and explore more deeply the consequences of using alternative benchmarks, especially for historical development paths: the 1990 benchmark, the 2011 benchmark, and the multiple benchmark method following the recent PWT methodology.

This paper focuses on the pre-1950 period and studies the biases that emerge in using these various ways of estimating the relative levels of GDP per capita in earlier periods in two ways. First, we analyze the extent to which the combinations of benchmarks and time series result in plausible historical estimates, such that, for example, countries do not end up with incomes (far) below subsistence level for prolonged periods of time. Second, we collected all available direct and indirect estimates of relative income levels for groups of countries in the 19th and early 20th century. We then compare the relative income levels using the different PPP scenarios, with the relative income levels as indicated by the independent (direct) and indirect benchmarks. We subsequently analyze which would produce better estimates of relative levels of GDP per capita in the period before 1940: the "traditional" Maddison approach that uses his 1990 benchmark, or a revision based on either the 2011 PPPs or the MBM approach.

The conclusion of this analysis is that simply moving from the 1990 to the 2011 benchmark does not improve estimates for this historical period, likely due to the greater distance in

years between 2011 and the historical time series under review. However, the MBM approach, which uses all post-1950 official ICP benchmark estimates, does not suffer from this bias, but does result in a rather high share of countries with below-subsistence levels of GDP per capita, which is also problematic. Our explanation of these results is that the quality of the benchmarks is probably subject to two changes: the quality of the PPPs and the coverage of the various ICP-rounds increases over time, and the distance in time to the "historical" period in question increases as well. Therefore, while a more recent PPP improves the relative income estimates for contemporary periods, using a more recent benchmark for extrapolating income backwards results in more biases for historical periods due to changes in relative prices.

Given the focus of the Maddison Project Database on the pre-1950 period, the assessment in this paper leads to the conclusion that the best way forward is to stick to the 1990 benchmark for the overall architecture of the Maddison Project Database. However, our analysis also leads to departures from the original Maddison approach in two ways. There is, to begin with, no doubt that the 2011 PPPs and the related estimates of GDP per capita reflect the relative levels of GDP per capita in the world economy better than extrapolating forward from the 1990 benchmark based on national account growth rates. We therefore integrate the 2011 benchmark into the Maddison Project Database by adapting the growth rates of GDP per capita in the period 1990–2011 to link the 1990 benchmark to the 2011 benchmark. Growth after 2011 is, in the current update, again exclusively based on the growth rates of GDP per capita according to national accounts. For the historical period, we also incorporated high-quality historical benchmark estimates into the dataset, resulting in a firmer anchoring of the historical estimates of GDP per capita. In this way, the Maddison Project Database now moves to a multiple benchmark approach.

A final remark on the use of PPPs for historical economic analyses, such as the Maddison Project, is that there is no single PPP concept, method, or dataset that can be seen as optimal for satisfying all uses by economists and economic historians—or any researcher, for that matter. What the best set of PPPs is for comparative analysis of economic performance over time ultimately depends on the question one wants to answer.

ACKNOWLEDGMENTS

Thanks to Michalis Moatsos, Daniel Gallardo-Albarran and two anonymous reviewers for very helpful and constructive comments and Robert Inklaar for constructive feedback and access to the PWT data. Further we thank the members of the Advisory Board of the Maddison Project and participants of the session "On the metrics and drivers of long-term wellbeing since 1800" at the XIXth Word Economic History Congress in Paris 2022 for their input in the paper.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data for the MPD 2023 release are available from April 1, 2024 here: https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/

ORCID

Jutta Bolt https://orcid.org/0000-0002-6002-6937

ENDNOTES

[1] Due to a fundamentally different approach to estimating long-run estimates, we do not use the econometric or indirect method in this paper as one of the alternatives, but we do exploit its information to cross-check the plausibility of the alternative extrapolation methods.

[2] The Penn World Table before PWT 8 also used information from multiple benchmarks. In PWT 4 & 5, Summers and Heston (1988) do a "consistentization" where successive benchmarks are made consistent with intermediate growth rates, but what exactly is adjusted, and how much for which year, is not exactly clear. Rao et al. (2010) later proposed a method that does something similar in a more transparent way. In PWT 6 and 7, Summers and Heston abandoned this approach again and instead tried to make a good comparison for one year and link growth rates from National Accounts to it. Sometimes that one year is based on one benchmark, and sometimes data from previous benchmarks is also incorporated, similar to how Maddison arrived at the 1990 GK PPP.

[3] Originally produced by the World Bank (World Bank, 2015). In this paper we use the 2011 PPPs as available in the PWT data, which are expressed in Geary Kamis dollars to ensure comparability with the original 1990 PPPs from the Maddison database. The most recent 2017 PPPs would in principle also be an option for this. However, this means moving even further away from the historical time period central in the Maddison Project Database. Fortunately, the good news from the 2017 round was that it was consistent with the 2011 round ("the good news … is the lack of news, that the results are broadly in line with earlier results from 2011," (Deaton & Schreyer, 2022). Additionally, Inklaar et al. (2022) show that the 2011 and 2017 PPPs are broadly consistent and that existing inconsistencies between the benchmarks do not change the international income distribution. For some individual countries, however, adjustments between benchmarks are substantial. In Section 4.2 we show that the fit between the historical benchmarks and the 2017 PPPs is worse than the alternative benchmarks more thoroughly discussed in the paper (see notes Tables 1 and 2).

[4] Sometimes an ICP 1993/96 round is referenced. Data collection took place between 1993 and 1996 and the results of the 1993 round were presented in "1996 terms" that is, the 1993 BH PPPs are re-referenced to 1996 with (usually) one deflator (Biggeri & Laureti, 2011; Silver, 2010).

[5] See: https://www.worldbank.org/en/programs/icp/history

[6] As indicated before, the more recent 2017 round is considered equal in quality to the 2011 round but is not considered in great detail here; see footnote 3.

[7] In essence, this means that the multiple benchmark approach anchors the historical series to the ICP benchmark closest to 1950 for each country (the first available benchmark for each country, which varies per country). This approach thus does not compare the same benchmarks for each country.

[8] Originally produced by the World Bank (World Bank, 2015), we use here the 2011 PPPs as used in the PWT, which are expressed Geary Kamis dollars.

[9] Using the extrapolated 1990 benchmark series.

[10] The benchmark income in 1990 uses the 1990 PPP benchmark level of income expressed in 2011 dollars. The benchmark income in 2011 is taken from the 2011 PPP benchmark, also expressed in 2011 dollars.

[11] Income estimates for Cuba are taken from Bolt et al. (2018), based on based on Ward and Devereux (2012), Brundenius (1984), Pérez-López (1987); United Nations National Income Database (http://data.un.org/Data.aspx?d=SNA&f=group_code%3a101, accessed November 2017).

[12] We take them from the PWT, see footnote 8.

[13] In this "test" we do not distinguish between quality of the various benchmark and include all.

[14] And obtained from the Total Economy Database published by the Conference Board.

[15] South-Africa in the 19th century is another example of extensive economic growth dominated by the colonization of large inland territories and the incorporation of indigenous populations in this process, resulting in a long-term decline of GDP per capita; see Fourie and Van Zanden 2013.

[16] This section is based on Bolt et al. (2018) and Bolt and Van Zanden (2020).

[17] Thanks to the recent paper by Chilosi and Ciccarelli (2023) we could incorporate estimates for Italy as a whole in the new dataset – older versions of the dataset covered Northern Italy only before 1871.

[18] As Palestine is not included in US Census Bureau's International Database 2022, we used the growth rate for the West Bank to extend the population estimates for Palestine. Further, for Burundi, Benin, El Salvador, Guinea, Guinea-Bissau, Honduras, Montenegro, Serbia and Swaziland, we used the US Census Bureau's International Database 2022 for the period 1950–2021.

[19] All dollar incomes mentioned in this section are expressed in 2011 dollars.

REFERENCES

Abad, L. A., & van Zanden, J. L. (2016). Growth under extractive institutions? Latin American per capita GDP in colonial times. *The Journal of Economic History*, *76*(4), 1182–1215.

Álvarez-Nogal, C., & De La Escosura, L. P. (2013). The rise and fall of Spain (1270–1850). *The Economic History Review*, *66*(1), 1–37.

Asian Development Bank. (2014). *Purchasing power parities and real expenditures*. Asian Development Bank. https://www.adb.org/publications/purchasing-power-parities-and-real-expenditures

Axenciuc, V. (2012). *Produsul intern brut al Romaniei* (Vol. 1). Institutl de Economie Nationala.

Barro, R. J., & Ursúa, J. F. (2008). Macroeconomic Crises since 1870. *Brookings Papers on Economic Activity*, *2008*, 255–335.

Bassino, J.-P., Broadberry, S., Fukao, K., Gupta, B., & Takashima, M. (2019). Japan and the great divergence, 730–1874. *Explorations in Economic History*, *72*, 1–22. https://doi.org/10.1016/j.eeh.2018.11.005

Bértola, L. (2016). *El PIB per cápita de Uruguay 1870–2015: Una reconstrucción*. https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/27146

Bértola, L., & Ocampo, J. A. (2012). *The economic development of Latin America since independence*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199662135.001.0001

Biggeri, L., & Laureti, T. (2011). Understanding changes in PPPs over time. A proposal for the Extrapolation of PPPs for Household Consumption using CPI data. International Comparison Program 61th Technical Advisory Group Meeting October 3–4, 2011.

Bolt, J., Inklaar, R., De Jong, H., & Van Zanden, J. L. (2018). Rebasing "Maddison": New income comparisons and the shape of long-run economic development (Maddison Project Working Paper, 10). https://www.academia.edu/65438391/Rebasing_Maddison_new_income_comparisons_and_the_shape_of_long_run_economic_development

Bolt, J., & van Zanden, J. L. (2014). The Maddison Project: Collaborative research on historical national accounts. *The Economic History Review*, *67*(3), 627–651. https://doi.org/10.1111/1468-0289.12032

Bolt, J., & Van Zanden, J. L. (2020). *Maddison style estimates of the evolution of the world economy. A new 2020 update. Maddison Working paper WP 15*.

Broadberry, S. (2003). Relative per capita income levels in the United Kingdom and the United States since 1870: Reconciling time-series projections and direct-benchmark estimates. *The Journal of Economic History*, *63*(3), 852–863.

Broadberry, S. (2006). *Market services and the productivity race, 1850–2000: British performance in international perspective*. Cambridge University Press. https://doi.org/10.1017/CBO9780511495748

Broadberry, S., Campbell, B. M. S., Klein, A., Overton, M., & van Leeuwen, B. (2015). *British economic growth, 1270–1870*. Cambridge University Press. https://doi.org/10.1017/CBO9781107707603

Broadberry, S., Custodis, J., & Gupta, B. (2015). India and the great divergence: An Anglo-Indian comparison of GDP per capita, 1600–1871. *Explorations in Economic History*, *55*, 58–75. https://doi.org/10.1016/j.eeh.2014.04.003

Broadberry, S., & Gardner, L. (2022). Economic growth in Sub-Saharan Africa, 1885–2008: Evidence from eight countries. *Explorations in Economic History*, *83*, 101424. https://doi.org/10.1016/j.eeh.2021.101424

Broadberry, S., Guan, H., & Li, D. D. (2018). China, Europe, and the great divergence: A study in Historical National Accounting, 980–1850. *The Journal of Economic History*, *78*(4), 955–1000. https://doi.org/10.1017/S0022050718000529

Cha, M. S., Kim, N. N., Park, K.-J., & Park, Y. (2022). *Historical statistics of Korea | SpringerLink*. Palgrave Macmillan. https://link.springer.com/book/10.1007/978-981-15-3874-2

Chen, S., & Ravallion, M. (2001). How did the world's poorest fare in the 1990s? *Review of Income and Wealth*, *47*(3), 283–300. https://doi.org/10.1111/1475-4991.00018

Chilosi, D., & Ciccarelli, C. (2023). *Italy in the great divergence: What can we learn from Engel's Law* (SSRN Scholarly Paper 4514091). https://doi.org/10.2139/ssrn.4514091

Clark, C. (1932). *National Income 1924–1931*. Routledge. https://www.routledge.com/National-Income-1924-1931/Clark/p/book/9781138976757

Clark, C. (1940). *The conditions of economic progress* (1st ed.). MacMillan and Co Limited.

Clark, C. (1951). *The conditions of economic progress* (2nd ed.). MacMillan and Co Limited.

Corso, G. D. (2013). El crecimiento economico de Venuzuela, Desde la Oligarquia Conservadora Hasta La Revolucion Bolivariana: 1830–2012. Uno Vision cuantitativa *: Venezuelan economic growth from the conservative

oligarchy to the Bolivarian Revolution (1830–2012). *Revista de Historia Economica – Journal of Iberian and Latin American Economic History*, *31*(3), 321–357. https://doi.org/10.1017/S0212610913000190

Corso, G. D., & Kalmanovitz, S. (2016). Una estimación del PIB de Panamá: 1906–1945. *Tiempo & Economía*, *3*, (1).

Deaton, A. (2010). Price indexes, inequality, and the measurement of world poverty. *American Economic Review*, *100*(1), 5–34. https://doi.org/10.1257/aer.100.1.5

Deaton, A., & Heston, A. (2010). Understanding PPPs and PPP-based national accounts. *American Economic Journal: Macroeconomics*, *2*(4), 1–35. https://doi.org/10.1257/mac.2.4.1

Deaton, A., & Schreyer, P. (2022). GDP, wellbeing, and health: Thoughts on the 2017 round of the international comparison program. *Review of Income and Wealth*, *68*(1), 1–15. https://doi.org/10.1111/roiw.12520

Díaz, J. B., Schwarzenberg, R. L., & Wagner, G. (2016). *Chile 1810–2010: La República en Cifras: Historical statistics*. Ediciones Universidad Católica de Chile.

Eloranta, J., Voutilainen, M., & Nummela, L. (2016). *Estimating Finnish economic growth before 1860*. Available at SSRN: https://ssrn.com/abstract=4706862 or http://dx.doi.org/10.2139/ssrn.4706862

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the Penn World Table. *American Economic Review*, *105*(10), 3150–3182. https://doi.org/10.1257/aer.20130954

Feenstra, R. C., Ma, H., Peter Neary, J., & Prasada Rao, D. S. (2013). Who Shrunk China? Puzzles in the measurement of real GDP. *The Economic Journal*, *123*(573), 1100–1129. https://doi.org/10.1111/ecoj.12021

Fourie, J., & van Zanden, J. L. (2013). GDP in the Dutch Cape Colony: The national accounts of a slave-based society. *South African Journal of Economics*, *81*(4), 467–490. https://doi.org/10.1111/saje.12010

Fremdling, R. (1991). Productivity comparison between Great Britain and Germany, 1855–1913. *Scandinavian Economic History Review*, *39*(1), 28–42. https://doi.org/10.1080/03585522.1991.10408198

Fukao, K., Bassino, J.-P., Makino, T., Paprzycki, R., Takashima, M., & Tokui, J. (2015). *Regional inequality and industrial structure in Japan: 1874–2008*. Maruzen Publishing Co., Ltd. https://www.rieti.go.jp/en/publications/summary/15030018.html

Gardner, L. (2022). *Sovereignty without power: Liberia in the age of empires, 1822–1980*. Cambridge University Press.

Grytten, H. (2015). *Norwegian gross domestic product by industry 1830–1930* (Norges Bank Working Paper, 19/2015). https://www.norges-bank.no/en/news-events/news-publications/Papers/Working-Papers/2015/192015/

Herranz-Loncán, A., & Peres-Cajías, J. A. (2016). Tracing the reversal of fortune in the Americas: Bolivian GDP per capita since the mid-nineteenth century. *Cliometrica*, *10*(1), 99–128. https://doi.org/10.1007/s11698-015-0125-2

Inklaar, R., Marapin, R., Woltjer, P., & Timmer, M. (2022). Inconsistencies in comparing relative prices overtime: Patterns and facts. In *Advances in economic measurement: A volume in honour of D. S. Prasada Rao* (pp. 633–663). Palgrave Macmillan.

Inklaar, R., & Rao, D. S. P. (2017). Cross-country income levels over time: Did the developing world suddenly become much richer? *American Economic Journal: Macroeconomics*, *9*(1), 265–290. https://doi.org/10.1257/mac.20150155

Jonsson, G. (1998). *The gross domestic product of Iceland, 1870–1945*. Nordiska Historiska Nasjonalregnkaper, Artikkelsamling from Workshop IV, Solstrand 13–15 November 1998.

Krantz, O. (1983). Historical national accounts—Some methodological notes. *Scandinavian Economic History Review*, *31*(2), 109–131. https://doi.org/10.1080/03585522.1983.10408005

Krantz, O. (2017). Swedish GDP 1300–1560: A tentative estimate. *Swedish GDP 1300–1560*.

Kravis, I. B., Summers, R., & Heston, A. (1982). Comments on: "Selected Problems of Intercountry Comparisons on the Basis of the Experience of the Eec". *Review of Income and Wealth*, *28*(4), 407–410. https://doi.org/10.1111/j.1475-4991.1982.tb00625.x

Kuboniwa, M. (2019). Estimating GDP and foreign rents of the oil and gas sector in the Soviet Union and present-day Russia. In M. Kuboniwa, Y. Nakamura, K. Kumo, & Y. Shida (Eds.), *Russian economic development over three centuries: New data and inferences* (pp. 421–438). Springer. https://doi.org/10.1007/978-981-13-8429-5_12

Kuznets, S. (1941). *National income and its composition, 1919–1938, volume I, National Bureau of Economic Research*.

Lindert, P. H., & Williamson, J. G. (2016a). American colonial incomes, 1650–1774. *The Economic History Review*, *69*(1), 54–77.

Lindert, P. H., & Williamson, J. G. (2016b). *Unequal gains*. Princeton University Press. https://press.princeton.edu/books/hardcover/9780691170497/unequal-gains

Maddison, A. (1995). *Monitoring the world economy, 1820–1992*. Organization for Economic.

Maddison, A. (2001). *The world economy—A millennial perspective—OECD*. Organisation for Economic Co-operation and Development. https://www.oecd.org/dev/world-economy-millennial-perspective-historical-statistics.htm

Maddison, A. (2003). *The world economy: Historical statistics*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/development/the-world-economy_9789264104143-en

Maddison, A. (2006). *The world economy: Volume 1: A millennial perspective and Volume 2: Historical Statistics*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/development/the-world-economy_9789264022621-en

Maddison, A. (2007). *Contours of the world economy 1–2030 AD*. Oxford University Press. https://global.oup.com/ukhe/product/contours-of-the-world-economy-1-2030-ad-9780199227204

Malanima, P. (2011). The long decline of a leading economy: GDP in central and northern Italy, 1300–1913. *European Review of Economic History*, *15*(2), 169–219. https://doi.org/10.1017/S136149161000016X

Malinowski, M., & van Zanden, J. L. (2017). Income and its distribution in preindustrial Poland. *Cliometrica*, *11*(3), 375–404. https://doi.org/10.1007/s11698-016-0154-5

Nazrin, S. (2016). *Economic history of Malaya*. Oxford University Press.

OECD. (2014). *How was life? Volume I*. https://www.oecd-ilibrary.org/content/publication/3d96efc5-en

OECD. (2021). *How was life? Volume II*. https://www.oecd-ilibrary.org/content/publication/3d96efc5-en

Palma, N., & Reis, J. (2019). From convergence to divergence: Portuguese economic growth, 1527–1850. *The Journal of Economic History*, *79*(2), 477–506. https://doi.org/10.1017/S0022050719000056

Pamuk, Ş. (2006). Estimating economic growth in the Middle East since 1820. *The Journal of Economic History*, *66*(3), 809–828.

Pamuk, Ş. (2009). *Estimating GDP per capita for the Ottoman Empire in a European Comparative Framework, 1500–1820*. World Economic History Congress.

Pfister, U. (2022). Economic growth in Germany, 1500–1850. *The Journal of Economic History*, *82*(4), 1071–1107. https://doi.org/10.1017/S002205072200033X

Piketty, T., Saez, E., & Zucman, G. (2018). Distributional national accounts: Methods and estimates for the United States. *The Quarterly Journal of Economics*, 553–609.

Prados de la Escosura, L. (2000). International comparisons of real product, 1820–1990: An alternative data set. *Explorations in Economic History*, *37*(1), 1–41. https://doi.org/10.1006/exeh.1999.0731

Prados de la Escosura, L. (2009). Lost decades? Economic performance in post-independence Latin America. *Journal of Latin American Studies*, *41*(2), 279–307.

Prados de la Escosura, L. (2017). *Spanish economic growth, 1850–2015*. Palgrave Macmillan. https://link.springer.com/book/10.1007/978-3-319-58042-5

Prados de la Escosura, L., Álvarez-Nogal, C., & Santiago-Caballero, C. (2022). Growth recurring in preindustrial Spain? *Cliometrica*, *16*(2), 215–241. https://doi.org/10.1007/s11698-021-00232-7

Rao, D. S. P., Rambaldi, A., & Doran, H. (2010). Extrapolation of purchasing power parities using multiple benchmarks and auxiliary information: A new approach. *Review of Income and Wealth*, *56*, s59–s98.

Ravallion, M., Datt, G., & van de Walle, D. (1991). Quantifying absolute poverty in the developing world. *Review of Income and Wealth*, *37*(4), 345–361. https://doi.org/10.1111/j.1475-4991.1991.tb00378.x

Ridolfi, L. (2017). The French economy in the longue durée: A study on real wages, working days and economic performance from Louis IX to the Revolution (1250–1789). *European Review of Economic History*, *21*(4), 437–438. https://doi.org/10.1093/ereh/hex026

Ridolfi, L., & Nuvolari, A. (2021). L'histoire immobile? A reappraisal of French economic growth using the demand-side approach, 1280–1850. *European Review of Economic History*, *25*(3), 405–428. https://doi.org/10.1093/ereh/heab012

Ruoen, R. (1997). *China's economic performance in an international perspective*. Organization for Economic Cooperation and Development.

Ryten, J. (1998). *Evaluation of the international comparison programme*. United Nations Economic and Social Council.

Santamaria, G. A. (2005). *Las Cuentas nacionales de Cuba, 1960–2005*. Mimeo.

Schön, L., & Krantz, O. (2015). New Swedish Historical National Accounts since the 16th Century in Constant and Current Prices. *New Swedish Historical National Accounts since the 16th Century in Constant and Current Prices*.

Seminario, B. (2015). *El desarrollo de la economía peruana en la era moderna: Precios, población, demanda y producción desde 1700*. Universidad de Pacifio. https://fondoeditorial.up.edu.pe/producto/el-desarrollo-de-la-economia-peruana-en-la-era-moderna-precios-poblacion-demanda-y-produccion-desde-1700/

Silver, M. (2010). *IMF applications of purchasing power parity estimates* (Internation Monetary Fund Working paper WP/10/253).

Stohr, C. (2016). *Trading gains: New estimates of Swiss GDP, 1851 to 2008* (LSE Economic History Working Papers). Article 67032. https://ideas.repec.org//p/ehl/wpaper/67032.html

Sugimoto, I. (2011). *Economic growth of Singapore in the twentieth century: Historical GDP estimates and empirical investigations* (Vol. 2). Economic Growth Centre Research Monongraph series.

Summers, R., & Heston, A. (1988). A new set of international comparisons of real product and price levels estimates for 130 countries, 1950–1985. *Review of Income and Wealth*, *34*(1), 1–25. https://doi.org/10.1111/j.1475-4991.1988.tb00558.x

Summers, R., & Heston, A. (1991). The Penn World Table (Mark 5): An expanded set of international comparisons, 1950–1988. *The Quarterly Journal of Economics*, *106*(2), 327–368. https://doi.org/10.2307/2937941

van Zanden, J. L., & van Leeuwen, B. (2012). Persistent but not consistent: The growth of national income in Holland 1347–1807. *Explorations in Economic History*, *49*(2), 119–130. https://doi.org/10.1016/j.eeh.2011.11.002

Voinea, L., Cojocaru, A., Costache, B., Grigoraș, V., Lovin, H., Neagu, C., Orțan, D., & Tănase, A. (2019). *One hundred years of honesty: Recovering the lost memory of the Romanian economy 1918–2018*.

Ward, M., & Devereux, J. (2003). Measuring British decline: Direct versus long-span income measures. *The Journal of Economic History*, *63*(3), 826–851. https://doi.org/10.1017/S0022050703542000

Ward, M., & Devereux, J. (2012). The road not taken: Pre-revolutionary Cuban living standards in comparative perspective. *The Journal of Economic History*, *72*(1), 104–133.

Ward, M., & Devereux, J. (2021). New income comparisons for the late nineteenth and early twentieth century. *Review of Income and Wealth*, *67*(1), 222–247.

Woltjer, P. (2015). Taking over: A new appraisal of the Anglo-American productivity gap and the nature of American economic leadership ca. 1910. *Scandinavian Economic History Review*, *63*(3), 280–301. https://doi.org/10.1080/03585522.2015.1034766

World Bank. (2008). *Global purchasing power parities and real expenditures: 2005 International Comparison Program*. World Bank. https://openknowledge.worldbank.org/handle/10986/21558

World Bank. (2015). *Purchasing power parities and real expenditures of world economics: A comprehensive report of the 2011 International Comparison Program*. World Bank. https://doi.org/10.1596/978-1-4648-0329-1

Wu, H. (2014). China's growth and productivity performance debate revisited—Accounting for China's sources of growth with a new data set. *The Conference Board, EWP 14-01*. www.conference-board.org/publications/publicationdetail.cfm?publicationid=2690

Xu, Y., Shi, Z., van Leeuwen, B., Ni, Y., Zhang, Z., & Ma, Y. (2017). Chinese National Income, ca. 1661–1933. *Australian Economic History Review*, *57*(3), 368–393. https://doi.org/10.1111/aehr.12127

How to cite this article: Bolt, J., & van Zanden, J. L. (2024). Maddison-style estimates of the evolution of the world economy: A new 2023 update. *Journal of Economic Surveys*, 1–41. https://doi.org/10.1111/joes.12618

APPENDIX I

TABLE A1 List of benchmarks and sources included in the analysis.

Benchmark comparison	Year	Source benchmark
Australia/UK	1891	Haig, B. (1989). International Comparisons of Australian Gdp in the 19th Century. Review of Income and Wealth, 35(2), 151–162. https://doi.org/10.1111/j.1475-4991.1989.tb00587.x
Austrialia/UK	1891	Thomas, M. (1995). A Substantial Australian Superiority? Anglo-Australian Comparisons of Consumption and Income in the Late Nineteenth Century" Australian Economic History Review 35: 10–37, Table 5, p. 35
Australia/UK	1900	Haig, B. (1989). International Comparisons of Australian Gdp in the 19th Century. Review of Income and Wealth, 35(2), 151–162. https://doi.org/10.1111/j.1475-4991.1989.tb00587.x
Australia/US	1872	Ward, M. and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Australia/US	1910	Ward, M. and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Australia/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Austria/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Belgium/US	1872	Ward, M. and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Belgium/US	1910	Ward, M. and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Belgium/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Burma/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Burma/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880

(Continues)

TABLE A1 (Continued)

Canada/US	1872	Ward, M. and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Canada/US	1910	Ward, M. and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Canada/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Ceylon/Japan	1922	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Ceylon/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
China/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
China/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
China/Netherlands	1825	Li, B., & van Zanden, J. L. (2012). Before the Great Divergence? Comparing the Yangzi Delta and the Netherlands at the Beginning of the Nineteenth Century. The Journal of Economic History, 72(4), 956–989.
China/UK	1840	Broadberry, S., Guan, H., & Li, D. (2018). China, Europe, and the Great Divergence: A Study in Historical National Accounting, 980–1850. The Journal of Economic History, 78(4), 955–1000. https://doi.org/10.1017/S0022050718000529
China/US	1912	Ma, Y. and de Jong, H. (2017), Unfolding the Turbulent Century: A Reconstruction of China's Historical National Accounts, 1840–1912. Review of Income and Wealth (forthcoming). https://doi.org/10.1111/roiw.12314
China/US	1935	Fukao, K., Ma, D., & Yuan, T. (2007). Real GDP in pre-war East Asia: A 1934–36 benchmark purchasing power comparison with the U.S. Review of Income and Wealth, 53(3), 503–537. https://doi.org/10.1111/j.1475-4991.2007.00243.x

(Continues)

TABLE A1 (Continued)

Czechoslovakia/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Denmark/US	1872	Ward, M. and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Denmark/US	1910	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Denmark/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Egypt/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809−828, p. 7, table 1
Estland/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Finland/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
France/UK	1820	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1830	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1840	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1850	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1860	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1870	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1880	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1890	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/UK	1900	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.

(Continues)

TABLE A1 (Continued)

France/UK	1910	O'Brien, P., & Keyder, C. (1978). Economic Growth in Britain and France 1780–1914 (Routledge Revivals): Two Paths to the Twentieth Century. George Allen and Unwin.
France/US	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
France/US	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
France/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Germany/UK	1860	Fremdling, R., Productivity comparison between Great Britain and Germany, 1855—1913, Scandinavian Economic History Review 39 (1), 28–42
Germany/UK	1870	Fremdling, R., Productivity comparison between Great Britain and Germany, 1855—1913, Scandinavian Economic History Review 39 (1), 28–42
Germany/UK	1880	Fremdling, R., Productivity comparison between Great Britain and Germany, 1855—1913, Scandinavian Economic History Review 39 (1), 28–42
Germany/UK	1890	Fremdling, R., Productivity comparison between Great Britain and Germany, 1855—1913, Scandinavian Economic History Review 39 (1), 28–42
Germany/UK	1900	Fremdling, R., Productivity comparison between Great Britain and Germany, 1855—1913, Scandinavian Economic History Review 39 (1), 28–42
Germany/UK	1910	Fremdling, R., Productivity comparison between Great Britain and Germany, 1855—1913, Scandinavian Economic History Review 39 (1), 28–42
Germany/UK	1935	Broadberry, S.N. (2006), Market Services and the Productivity Race, 1850–2000: Britain in International Perspective. Cambridge: Cambridge University Press.
Germany/US	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Germany/US	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Germany/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Greece/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27

(Continues)

TABLE A1 (Continued)

HUN/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
India/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
India/Japan	1922	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
India/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
India/UK	1950/51	Broadberry, S., & Gupta, B. (2010). The historical roots of India's service-led development: A sectoral analysis of Anglo-Indian productivity differences, 1870–2000. Explorations in Economic History, 47(3), 264–278. https://doi.org/10.1016/j.eeh.2009.09.004
India/US	1870	Heston, A. and R. Summers, 1980. Comparative Indian Economic Growth: 1870 to 1970, American Economic Review, vol. 70 (2), pages 96–101
Indonesia/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Indonesia/Japan	1922	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Indonesia/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Indonesia/Netherlands	1820	Van Zanden, J.L. (2003). 'Rich and Poor before the Industrial Revolution: a comparison between Java and the Netherlands at the beginning of the 19th century', Explorations in Economic History 40, 1–23.
Indonesia/Netherlands	1860	Van Zanden, J.L. (2003). 'Rich and Poor before the Industrial Revolution: a comparison between Java and the Netherlands at the beginning of the 19th century', Explorations in Economic History 40, 1–23.
Ireland/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27

(Continues)

TABLE A1 (Continued)

Iran/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809–828, p. 7, table 1
Iraq/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809–828, p. 7, table 1
Italy/US	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Italy/US	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Italy/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Japan/US	1935	Fukao, K., Ma, D., & Yuan, T. (2007). Real GDP in pre-war East Asia: A 1934–36 benchmark purchasing power comparison with the U.S. Review of Income and Wealth, 53(3), 503–537. https://doi.org/10.1111/j.1475-4991.2007.00243.x
Jordan/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809–828, p. 7, table 1
Korea/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Korea/Japan	1922	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Korea/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Korea/US	1935	Fukao, K., Ma, D., & Yuan, T. (2007). Real GDP in pre-war East Asia: A 1934–36 benchmark purchasing power comparison with the U.S. Review of Income and Wealth, 53(3), 503–537. https://doi.org/10.1111/j.1475-4991.2007.00243.x
Libanon/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809–828, p. 7, table 1

(Continues)

TABLE A1 (Continued)

Netherlands/UK	1820	Frankema, E., P. Woltjer and J.P. Smits (2013). Changing Economic Leadership, a new benchmark of sector productivity in the United States and Western Europe, ca. 1910, Tijdschrift voor Sociale en Economische Geschiedenis (The low countries social and economic history), Vol 10, No. 3, pp. 80–113
Netherlands/UK	1850	Frankema, E., P. Woltjer and J.P. Smits (2013). Changing Economic Leadership, a new benchmark of sector productivity in the United States and Western Europe, ca. 1910, Tijdschrift voor Sociale en Economische Geschiedenis (The low countries social and economic history), Vol 10, No. 3, pp. 80–113
Netherlands/UK	1870	Frankema, E., P. Woltjer and J.P. Smits (2013). Changing Economic Leadership, a new benchmark of sector productivity in the United States and Western Europe, ca. 1910, Tijdschrift voor Sociale en Economische Geschiedenis (The low countries social and economic history), Vol 10, No. 3, pp. 80–113
Netherlands/UK	1890	Frankema, E., P. Woltjer and J.P. Smits (2013). Changing Economic Leadership, a new benchmark of sector productivity in the United States and Western Europe, ca. 1910, Tijdschrift voor Sociale en Economische Geschiedenis (The low countries social and economic history), Vol 10, No. 3, pp. 80–113
Netherlands/UK	1910	Frankema, E., P. Woltjer and J.P. Smits (2013). Changing Economic Leadership, a new benchmark of sector productivity in the United States and Western Europe, ca. 1910, Tijdschrift voor Sociale en Economische Geschiedenis (The low countries social and economic history), Vol 10, No. 3, pp. 80–113
Netherlands/US	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Netherlands/US	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Netherlands/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Norway/US	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Norway/US	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Norway/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27

(Continues)

TABLE A1 (Continued)

New Zealand/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Philippines/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Philippines/Japan	1922	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Philippines/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Poland/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Saudi Arabia/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809–828, p. 7, table 1
Spain/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Sweden/US	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Sweden/US	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Sweden/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Switzerland/US	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Switzerland/US	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
Switzerland/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
Syria/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809–828, p. 7, table 1

(Continues)

TABLE A1 (Continued)

Taiwan/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Taiwan/Japan	1922	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Taiwan/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Taiwan/US	1935	Fukao, K., Ma, D., & Yuan, T. (2007). Real GDP in pre-war East Asia: A 1934–36 benchmark purchasing power comparison with the U.S. Review of Income and Wealth, 53(3), 503–537. https://doi.org/10.1111/j.1475-4991.2007.00243.x
Thailand/UK	1958	Usher, D. (1963). The Thai National Income at United Kingom Prices. Bulletin of the Oxford University Institute of Economics & Statistics, 25: 199–214. https://doi.org/10.1111/j.1468-0084.1963.mp25003003.x
Thailand/Japan	1938	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Turkey/US	1913	Pamuk, Ş. (2006). Estimating Economic Growth in the Middle East since 1820. The Journal of Economic History, 66(3), 809–828, p. 7, table 1
Turkey/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
US/UK	1872	Ward, M. & Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
US/UK	1910	Ward, M., and Devereux, J. (2021), New Income Comparisons for the late Nineteenth and Early Twentieth Century. Review of Income and Wealth, 67: 222–247. https://doi.org/10.1111/roiw.12466, table 4, p. 235
UK/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
US/UK	1775	Lindert, P. H., & Williamson, J. G. (2016b). Unequal Gains. Princeton University Press.

(Continues)

TABLE A1 (Continued)

US/UK	1800	Lindert, P. H., & Williamson, J. G. (2016b). Unequal Gains. Princeton University Press.
US/UK	1850	Lindert, P. H., & Williamson, J. G. (2016b). Unequal Gains. Princeton University Press.
US/UK	1860	Lindert, P. H. & Williamson, J. G. (2016b). Unequal Gains. Princeton University Press.
US/UK	1870	Lindert, P. H., & Williamson, J. G. (2016b). Unequal Gains. Princeton University Press.
US/UK	1873	Ward, M., & Devereux, J. (2003). Measuring British Decline: Direct Versus Long-Span Income Measures. The Journal of Economic History, 63(3), 826–851. https://doi.org/10.1017/S0022050703542000
US/UK	1878	Ward, M., & Devereux, J. (2003). Measuring British Decline: Direct Versus Long-Span Income Measures. The Journal of Economic History, 63(3), 826–851. https://doi.org/10.1017/S0022050703542000
US/UK	1884	Ward, M., & Devereux, J. (2003). Measuring British Decline: Direct Versus Long-Span Income Measures. The Journal of Economic History, 63(3), 826–851. https://doi.org/10.1017/S0022050703542000
US/UK	1891	Ward, M., & Devereux, J. (2003). Measuring British Decline: Direct Versus Long-Span Income Measures. The Journal of Economic History, 63(3), 826–851. https://doi.org/10.1017/S0022050703542000
US/UK	1905	Ward, M., & Devereux, J. (2003). Measuring British Decline: Direct Versus Long-Span Income Measures. The Journal of Economic History, 63(3), 826–851. https://doi.org/10.1017/S0022050703542000
US/UK	1908	Woltjer, P. (2015). Taking over: A new appraisal of the Anglo-American productivity gap and the nature of American economic leadership ca. 1910. Scandinavian Economic History Review, 63(3), 280–301. https://doi.org/10.1080/03585522.2015.1034766
US/UK	1910	Broadberry, S. N., & Irwin, D. A. (2006). Labor productivity in the United States and the United Kingdom during the nineteenth century. Explorations in Economic History, 43(2), 257–279. https://doi.org/10.1016/j.eeh.2005.02.003
US/UK	1930	Ward & Devereux (2003)
US/UK	1935	Ward, M., & Devereux, J. (2003). Measuring British Decline: Direct Versus Long-Span Income Measures. The Journal of Economic History, 63(3), 826–851. https://doi.org/10.1017/S0022050703542000
US/UK	1937	Broadberry, S. (2006). Market Services and the Productivity Race, 1850–2000: British Performance in International Perspective. Cambridge University Press. https://doi.org/10.1017/CBO9780511495748
USSR/US	1913	Nutter, G.W. (1962), The growth of industrial production in the Soviet Union, NBER: Princeton. table 63, p. 238

(Continues)

TABLE A1 (Continued)

USSR/US	1929	Nutter, G.W. (1962), The growth of industrial production in the Soviet Union, NBER: Princeton. table 63, p. 238
USSR/US	1955	Nutter, G.W. (1962), The growth of industrial production in the Soviet Union, NBER: Princeton. table 63, p. 238
USSR/US	1955	Bergson, A. (1972). The Comparative National Income of the USSR and the United States, NBER, http://www.nber.org/chapters/c5095, pp. 145–224
Vietnam/Japan	1913	Bassino, J.-P., & van der Eng, P. (2020). Asia's 'little divergence' in the twentieth century: Evidence from PPP-based direct estimates of GDP per capita, 1913–69. The Economic History Review, 73(1), 185–208. https://doi.org/10.1111/ehr.12880
Yugoslavia/US	1929	Clark, C., 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27
South Africa/US	1929	Clark, C. 1951. The Conditions of Economic Progress. London: MacMillan & Co. Ltd, p. 27