How Was Life? Volume II

NEW PERSPECTIVES ON WELL-BEING AND GLOBAL INEQUALITY SINCE 1820

OECD

DEVELOPMENT CENTRE

OECD Better Life Initiative

How Was Life? Volume II

NEW PERSPECTIVES ON WELL-BEING AND GLOBAL INEQUALITY SINCE 1820

DEVELOPMENT CENTRE

Please cite this publication as:
OECD (2021), *How Was Life? Volume II: New Perspectives on Well-being and Global Inequality since 1820*, OECD Publishing, Paris, *https://doi.org/10.1787/3d96efc5-en*.

ISBN 978-92-64-80029-8 (print)
ISBN 978-92-64-40315-4 (pdf)

Corrigenda to publications may be found on line at: *www.oecd.org/about/publishing/corrigenda.htm*.

Preface

In 2014, the OECD published *How Was Life?: Global Well-being since 1820,* the result of a collaboration between the OECD and the OECD Development Centre, on one side, and an international group of economic historians organised around the Clio-Infra initiative and the Maddison Project, on the other (van Zanden et al., 2014[1]). This joint undertaking built on Angus Maddison's long career at the OECD and the OECD Development Centre, and on the close contacts that he maintained with these institutions. In the 1990s and early 2000s, the OECD Development Centre had published Angus' pioneering books on long-term economic growth in the world economy.

How Was Life? emerged from talks about how to continue the work of Angus Maddison and add an historical and global dimension to the OECD's Better Life Initiative, launched following the release of the Stiglitz, Sen and Fitoussi report (*Report by the Commission on the Measurement of Economic Performance and Social Progress*) (2009[2]). The goal was to provide an historical perspective to the evidence included in *How's Life?,* the OECD report published bi-annually since 2011 and now in its fifth issue, which relies on a large set of comparable well-being indicators for OECD countries and, to the extent possible, other major economies. *How Was Life?* added an historical dimension to this endeavour, by presenting data on a broad range of well-being dimensions for 25 large countries, eight regions and the world as a whole for the period since 1820. The community of economic historians organised around the Clio-Infra initiative supplied the historical data featured in *How Was Life?* and the expertise to assess the quality of these estimates and interpret the long-term trends in the world economy.

The present report continues this collaboration between the OECD and the Clio-Infra Initiative by presenting evidence from an even more ambitious research agenda. First, it extends the available historical evidence to other dimensions of people's well-being, by presenting new estimates of working hours, biodiversity loss (a key aspect of sustainability) and government social spending, as well as new estimates of GDP that account for the 2011 round on purchasing power parities prepared by the International Comparison Program. Second, it broadens the perspective of inequality by looking beyond income disparities. The rationale is that, if we are interested in multi-dimensional well-being, we should look beyond income inequalities (covered in a chapter of the 2014 *How Was Life?*) to the other dimensions of well-being that are included in the OECD Better Life Initiative. The present report pursues this approach by presenting historical evidence drawn for large datasets on inequalities in wealth, longevity and educational attainment, as well as gender disparities and extreme poverty. The report retains a "global" perspective, providing evidence on a large number of countries across the whole world. While the historical evidence presented often relies on partial and sometimes limited evidence, each of the chapters in this book includes an assessment of the quality of the data used and identifies areas for further historical research.

The publication of this report is testimony of the role played by the OECD in promoting the discussion on "GDP and beyond", and also witnessed by the recent release of the reports of the OECD-hosted High Level Expert Group on the Measurement of Economic Performance and Social Progress chaired by Joseph Stiglitz, Jean-Paul Fitoussi and Martine Durand (2018[3]; 2018[4]). The OECD Secretary-General Angel Gurría has put the notions of well-being and inclusive growth at the heart of his vision and efforts to enhance the OECD's relevance, responsiveness and impact during his tenure. The OECD Development

Centre has operationalised many of these ideas in its *Multi-Dimensional Country Reviews.* The new OECD Centre on Well-Being, Inclusion, Sustainability and Equality of Opportunity (WISE), created on 1 July 2020 by gathering under one roof a range of statistical and policy activities previously carried out in different parts of the organisation, aims to further strengthen and consolidate this work.

We believe that the long-term and global perspective provided in this report will be an essential reference for researchers, practitioners and general readers interested in knowing more on the historical development of people's lives.

Mario Pezzini,
Director, OECD Development Centre

Romina Boarini,
Director, OECD Centre for Well-being, Inclusion, Sustainability and Equal Opportunity (WISE)

Foreword

Jan Luiten van Zanden, Utrecht University and University of Stellenbosch
Marco Mira d'Ercole, OECD WISE Centre
Mikołaj Malinowski, Utrecht University and Universities of Lund and Humboldt in Berlin
Auke Rijpma, Utrecht University and the International Institute of Social History

In 2014, the OECD published the report *How Was Life?: Global Well-being since 1820* (van Zanden et al.[1]). Its aim was to provide an historical counterpart to the *How's Life?* report published bi-annually by the OECD since 2011. The latter report was a first attempt at the international level to go beyond the conceptual stage and to present a large set of comparable well-being indicators for OECD countries and, to the extent possible, other major economies. *How Was Life?* added an historical dimension to this pioneering work, presenting data for 25 large countries, eight world regions and the world as a whole for the period since 1820. *How Was Life*? was based on a collaboration between the OECD and the economic historical community organised around the Clio-Infra initiative and the Maddison Project. These initiatives supplied the historical data for the *How Was Life?* report, the expertise to assess the quality of the data and the possibilities for interpreting the long-term trends in the world economy. This volume continues the efforts initiated in *How Was Life?*.

The background to these reports was the "GDP and beyond" debate that, in recent years, received new impetus thanks to the seminal report by Stiglitz, Sen and Fitoussi (2009[2]). Economic historians have done a lot of work on estimating trends in GDP and GDP per capita for the world economy in the past thousand years, which had been integrated by Angus Maddison into one consistent dataset of GDP, population and GDP per capita for the past two millennia. Maddison had always maintained close contact with the OECD, which, in the 1990s and early 2000s, published his major books on long-term economic growth in the world economy. The *How Was Life?* volume emerged from talks about how to continue the Maddison tradition and how to add a historical dimension to the OECD's Better Life Initiative.

The present report continues this line of work, and develops an even more ambitious research agenda. One implication of the "beyond GDP" debate is that we should not measure inequality within countries – and at the global level – on the basis of estimates of income disparities only, but broaden our approach to other aspects of social and economic inequality. In other words, if we are interested in inequality of well-being, we have to try to measure the development of the inequalities of all dimensions of well-being that are, for example, included in the OECD Better Life Initiative. This has been the first goal of this report: to collect and standardise historical data about inequalities of various dimensions of well-being. This is a new and experimental area for historical research. Such data have not been collected systematically in the past, and their use requires the analysis of large new datasets (as presented in the various chapters of this report). The chapters in this book on inequality in wealth (Chapter 5), longevity (Chapter 6), educational attainment (Chapter 7), gender (Chapter 8) and the share of the population living in extreme poverty (Chapter 9) are all focused on the evolution – at the national but also at the global level – of these aspects of inequality.

A related aim of this second book is to extend the database, and discuss and present new data covering other dimensions of well-being featuring in the Better Life Initiative. The chapter on Working Hours covers one of the main gaps in the *How Was Life?* report based on new historical data (Chapter 3). New historical research on one of the key aspects of sustainability, Biodiversity Loss, is presented in Chapter 10. This volume also presents new historical estimates of GDP that account for the 2011 round on purchasing power parities prepared by the International Comparison Program (Chapter 2). In addition, a global dataset on social spending is presented in Chapter 4.

The final chapter synthesises the results of the report through two composite indices of well-being**;** first, an average of the average well-being measures of each country presented in the current and previous volumes (i.e. a “mean-of-means”); second, a composite of the within-country inequality in income, educational attainment, life expectancy and gender.

Acknowledgements

This report has been prepared by a team of historians participating in the Clio-Infra project. Individual chapters of the report have been prepared by Jutta Bolt and Jan Luiten van Zanden (Chapter 2); Oisín Gilmore (Chapter 3); Peter Lindert (Chapter 4); Guido Alfani and Sonia Schifano (Chapter 5); Lamar Crombach, Jeroen Smits and Christiaan Monden (Chapter 6); Bas van Leeuwen and Jieli Li (Chapter 7); Selin Dilli, Sarah Carmichael and Auke Rijpma (Chapter 8); Michail Moatsos (Chapter 9); Thomas van Goethem and Jan Luiten van Zanden (Chapter 10); and Auke Rijpma (Chapter 11). Statistical assistance in the production of tables and figures has been provided by Vararat Atisophon, from the OECD Development Centre, while editorial assistance was provided by Helena Cordasev. Anne-Lise Faron, WISE Publications and Event Coordinator, prepared the manuscript for publication; Sonia Primot designed the cover and Martine Zaida, WISE Communications, Partnership and Forum Manager, provided advice on communication aspects. Coordination of the report has been ensured by Marco Mira d'Ercole (OECD) and Jan Luiten van Zanden, Mikołaj Malinowski and Auke Rijpma (all at Utrecht University).

The authors of the individual chapters have benefitted from comments provided by the coordinators of the report, by other members of the Clio-Infra team, as well as by participants and discussants at a workshop organised by the OECD in cooperation with Clio-Infra held in Paris in 2019. Special thanks to Stephen Broadberry (University of Oxford), Guido Alfani (Bocconi University), Daniel Waldenström (Paris School of Economics), Rosie Seaman (University of Stirling), Jan Luiten van Zanden (Utrecht University), Mikolaj Szołtysek (University of Warsaw), Leandro Prados de la Escosura (Carlos III University), Aafke Schipper (Radboud University), Peter Van de Ven (OECD), Pascal Marianna (OECD), Willem Adema (OECD), Carlotta Balestra (OECD), Fabrice Murtin (OECD), William Thorn (OECD), Lorenzo Pavone (OECD), Alex Kolev (OECD), Katia Karousakis (OECD) and Romina Boarini (OECD) for commenting on various chapters.

Table of contents

Tables

Figures

Boxes

Readers' Guide

Conventions

- Through the report, data shown for world regions and the world as a whole are population-weighted averages for all countries included in the Clio-Infra database. The coverage of countries typically increases as more country data become available for more recent periods. To ensure more meaningful trends for world regions over time, imputations are made for missing countries.
- Depending on the chapter, the data shown in the tables and figures may refer to either individual years or decadal averages. In the latter case, the data are shown as, e.g. "1990s" and the corresponding values are computed as decadal averages of the 1990-1999 period.
- Data quality varies across domains, regions and periods: an assessment of data quality is provided in a summary table within each chapter.
- Historical data labelled RUS may sometimes refer to data covering the former Soviet Union. When this occurs, it is mentioned in notes to the tables and figures.
- Inequality is generally measured using the Gini coefficient. However, the chapter on wealth inequality also relies on measures of the share of wealth accruing to the richest 10% of the population, while the chapter on educational inequality also presents measures based on the standard deviation.
- Gini coefficients for world regions are population-weighted averages of Gini coefficients for individual countries. They are hence not interpretable as Gini coefficients computed across all people in the region, irrespectively of the country where they lived.
- (..) is used to indicate missing values.

Executive summary

Are we better off than our ancestors? Did quality of life improve over historical times? And, if so, did the tide of progress lift all boats? The Industrial Revolution was a critical break in human history that reshaped societies and initiated a new era of economic growth. According to the international estimates of Gross Domestic Product (GDP) produced by Simon Kuznets, Angus Maddison and their followers, material living standards have increased substantially since the adaptation of the steam engine and other path-breaking techniques that truly reshaped the world. The latest historical estimates of GDP presented in this book show that per capita GDP in the world economy expanded 13-fold between 1820 and 2016. The average human now is undoubtedly much richer than her or his counterpart two centuries ago.

However, GDP per capita provides an incomplete picture of the way societies evolve. Two limitations stand out. First, GDP accounts for the market value of goods and services produced in an economy. It captures the material aspect of human experience, but not other aspects essential for quality of life such as their health, skills or happiness. Moreover, although in poor societies, any increase in the availability of material goods is directly beneficial to the inhabitants' living conditions, in rich countries the relationship between material abundance and all-round prosperity is less clear. These and other concerns have long been voiced by economists and other social scientists. The "Beyond GDP" initiative aims to complement the picture based on GDP with a range of other well-being indicators such as poverty, resource depletion, health and quality of life. The *How Was Life?* report published by the OECD in 2014 was a first step towards building a more holistic picture of the global economy in the past 200 years, by presenting evidence on secular changes in educational attainment, personal security and health status, among others (van Zanden et al.[1]). This book continues this task by presenting historical evidence on other well-being dimensions such as working time and more.

The second limitation of GDP per capita as an overarching measure of development is that it represents the economic output available to the "average" citizen. It does not account for how that output is distributed within the population nor does it indicate how representative the identified average is for the total population. The issue of inequality has always been at the core of political economy, and has featured prominently in the public debate thanks to the works of, among others, Joseph Stiglitz, Angus Deaton, Thomas Piketty and Branko Milanović. Inequality can be describe either between or within countries. With reference to the former, while GDP per capita has increased everywhere around the world since 1820, some societies benefited from this expansion more than others. In particular, the 19th century has been described as the period of the "Great Divergence" between prosperous Western countries and the rest of the global economy. Much less is known about inequality within countries. While the *How Was Life?* previous volume presented evidence on historical income inequality within countries, this book moves "Beyond Income" to account for disparities in a range of well-being aspects within countries.

Economic historians and social scientists have been producing estimates of inequality in non-material dimensions of well-being for some years but, so far, these estimates have not been compared systematically over time and space. This report aims to fill this gap. It presents evidence on long-term trends in inequality in other dimensions of well-being, such as longevity and educational attainment across the global economy since 1820, providing estimates for a large number of countries. Where applicable, trends are charted for 25 countries, eight world regions and the world economy as a whole. This report presents, collects, harmonises and documents the estimates developed through the Clio-Infra and Maddison Projects, presenting the data, discussing sources and their limitations, providing an overview of trends and suggesting avenues for further research.

The report opens with new estimates of GDP per capita that form the backdrop of the analysis. Second, it presents two new indicators of average well-being in a country, namely, the average working week in manufacturing and social transfers. Third, it presents estimates of three types of within-country inequality, namely, disparities in wealth, in life expectancy and in educational attainment, as well as gender inequality. Lastly, the report presents new measures of extreme poverty and biodiversity loss.

This book provides evidence of historical trends in nine different dimensions of well-being and inequality. For some of these dimensions, the statistical correlation with the increase in the GDP per capita is high. The global average **GDP per capita** increased by a factor of 13 between the 1820s and the 2010s; in other terms, as the total population increased seven-fold, total GDP went up by a factor of 90. But this growth was spread very unevenly over the globe. Until about 1950, the rich countries grew faster than the poor ones, resulting in a huge increase in international income disparities. This does not mean that the poor countries did not grow economically. In 1950, the **global extreme poverty rate,** based on the inability to purchase a near-subsistence basket, was 53%. By 1990, the rate had dropped to 31%. Broadly speaking, economic expansion, especially in rich countries, also led to a better work-and-life balance. Amongst OECD countries, the length of the **working week in manufacturing** has declined dramatically since the 19th century. Full-time workers in manufacturing worked between 60-90 hours per week in the 19th century. Today they work roughly 40 hours. Economic expansion had environmental costs, however. **Biodiversity** has declined at least since 1500, and probably for the better part of the Holocene. A case study reconstructing biodiversity change in the Netherlands from 1900 onwards shows a long-term decline in biodiversity until 1970, followed by a partial recovery since then.

Economic expansion was associated with changes in inequality. In general, within-country **wealth inequality** increased in the 19th century, and continued to grow during the first decade of the 20th century. In the West, this tendency was stopped by World War I and II and by the troubled times in-between. However, there is no clear correlation between the level of GDP per capita and wealth inequality. This could partly reflect the development of the “welfare state”. Public spending on social transfers was tiny or non-existent until political developments assigned more and more of these tasks to the government in the last hundred years. In 1900, no national government transferred more than 3% of the country’s GDP via social programmes. This has increased gradually in most countries, rising to levels of between 15 and 30% today. Provision of public goods by the government links up with within-country **inequality of educational achievement** and **life expectancy**, which fell significantly over the 19th and 20th centuries.

A composite measure combining the average well-being measures of each country presented in the current and previous volumes shows that adding the new measures included in this book does not fundamentally change the global picture drawn in 2014: average well-being increased significantly throughout the 1820-2010 period, and these gains were more evenly distributed across countries than in the case of GDP per capita. When considering inequalities, the long-term declines of inequality in educational attainment and life expectancy as well as of gender inequality mean that this composite measure declined continuously in many parts of the world since 1820, more than offsetting episodes of rising income inequality. The 20th century therefore featured both large improvements of average levels of well-being and massive reductions in well-being inequality in the areas of life expectancy, educational attainment and gender relations. These results show the importance of trying to get the full picture of inequality within societies.

Overall, this book collects, summarises and critically discusses our current knowledge on long-term trends in global well-being and inequality over the past two centuries. While it provides a better view than one based on GDP per capita and income inequality alone, weak spots remain in our knowledge, in particular, concerning developments in Africa before 1950 and in major Asian countries in the 19th century. It is hoped that this book will stimulate further research into these areas.

References

Stiglitz, J., J. Fitoussi and M. Durand (2018), *Beyond GDP: Measuring What Counts for Economic and Social Performance*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264307292-en. [3]

Stiglitz, J., J. Fitoussi and M. Durand (eds.) (2018), *For Good Measure: Advancing Research on Well-being Metrics Beyond GDP*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264307278-en. [4]

Stiglitz, J., A. Sen and J. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, https://www.economie.gouv.fr/files/finances/presse/dossiers_de_presse/090914mesure_perf_eco_progres_social/synthese_ang.pdf (accessed on 4 February 2021). [2]

van Zanden, J. et al. (eds.) (2014), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-en. [1]

1 New perspectives on global inequality throughout history

Jan Luiten van Zanden, Utrecht University and University of Stellenbosch
Marco Mira d'Ercole, OECD WISE Centre
Mikołaj Malinowski, Utrecht University and University of Groningen
Auke Rijpma, Utrecht University and the International Institute of Social History, Amsterdam

This chapter provides an introduction to, and summary of, the contents of this book. It outlines the aim of the project and provides an overview of the indicators covered, comparing them with those used in the OECD Better Life Initiative. The chapter also presents the criteria used through the report to assess the quality of the indicators used and discusses practical and methodological issues. Finally, the chapter summarises the content of each chapter and its main highlights.

Introduction

"Black Lives Matter" is the battle-cry of the movement against racism that suddenly dominated the news and the public debate in many countries in mid-2020. It reminds us, if that were still necessary, how deep social inequalities are rooted in our societies, in this case going back to the tragic history of slavery and the slave trade centuries ago. At the same time, the refugee crisis, which was on top of political agendas in recent years, is a reflection of another dimension of inequality, the huge differences in "citizen rent" (the concept coined by Branko Milanovic to describe the economic benefits that accrue to people just by virtue of their residence in a given country) within the world economy. Global inequality in real incomes (and in citizen rent) has been a persistent feature of the world economy since, at least, the Industrial Revolution that began some 200 years ago. Around 1820, average real incomes in the richest regions in the world were at most about five times the levels in the poorest regions. And, as shown in this volume, the spread in terms of other welfare measures was even smaller. Since then a "Great Divergence" has resulted in the massive levels of cross-country inequality that characterise the world today.

But the picture of long-term trends in inequality is quite complex. In some parts of the world some types of inequality have declined substantially: gender inequality is a case in point, but inequalities in life expectancy and educational attainment have also shown a declining trend. The overall picture of changes in inequality in the world economy is highly complex, with many country-specific patterns – such as the dramatic rise of almost all dimensions of inequality in the United States in the period since 1980.

Most of what we know about long-term trends in welfare is based on the historical estimates of real Gross Domestic Product (GDP) by Angus Maddison, summarised in his books. In his view, which is confirmed by more recent research, Western Europe was already much more economically prosperous than the rest of the world at the start of the industrialisation process (i.e. around 1820).The gap between the United Kingdom, the "productivity leader" at the time, and the poorest parts of the world was at most 5 to 1. The 19th century was a period of divergence, during which the rich became richer – Western Europe and its offshoots (the United States, Canada, Australia) profited from the technological changes unleashed by the Industrial Revolution – whereas other parts of the world economy (China, India, Indonesia) saw their GDP per capita fall or stagnate at best, due in part to de-industrialisation and colonial exploitation. During the first half of the 20th century, it was mainly the United States that forged ahead, but the Transatlantic income gap that emerged as a result narrowed again during the so-called Golden Years of Capitalism that saw the Western European countries grow rapidly between 1950 and 1973.

Gradually, other parts of the world also began to participate in the process of modern growth, sometimes helped by decolonisation. In particular, East Asia became the most dynamic centre of the world economy after 1980, when China also joined the convergence club. On the other hand, economic growth in Africa has been much more incidental, and this continent has continued to lag behind the rest of the world.

In the previous *How Was Life* report, published by the OECD in 2014 (van Zanden et al.[1]), we tried to find out whether these patterns based on the reconstruction of GDP estimates were confirmed when we broadened the scope of analysis beyond GDP and included several other indicators reflecting trends in aggregate well-being at the level of the nation, such as earnings, health and personal security.

The first report was a contribution to the "beyond GDP" debate. The measurement of welfare based solely on GDP has been criticised for quite a long time, and this debate has received new impetus thanks to the seminal report by Stiglitz, Sen and Fitoussi (2009[2]). There is an increasing awareness that GDP provides only a partial perspective on the multidimensional nature of well-being – what matters to people's lives – and has limited power in explaining changes in other aspects. Many old and new indicators of quality of life such as life expectancy and level of schooling have been suggested to supplement GDP estimates. Much progress has also been made to derive complementary summary measures. The Human Development Index (HDI) of the United Nations Development Programme (UNDP, 1990[3]; 2013[4]) was a pioneering effort. More recently, scoreboards have been constructed based on headline indicators such as the one used by the OECD in its Better Life Initiative (OECD, 2020[5]).

An alternative measure of well-being that has been much discussed is subjective well-being, as measured by national of international surveys that ask people to evaluate their lives as a whole. There are, however, problems in international comparisons of this kind (language and cultural issues regarding what "happiness" or "life satisfaction" means in different parts of the world), which apply to an even larger degree to comparisons over time. And there is the obvious insurmountable problem that there are simply no social surveys for the period before 1970. Attempts to find a suitable proxy – such as the use of the word "happiness" in the literature - have so far been only partly successful (Hills, 2019[6]).

Against this background, the 2014 *How Was Life?* report (van Zanden et al.[1]) provided a synthesis of what was known about historical developments in various dimensions of well-being in the world economy since 1820. This co-operative effort by the OECD and its Development Centre, in collaboration with Clio-Infra and the Maddison Project, covered GDP per capita and Income Inequality, Real Wages, Educational Attainment, Life Expectancy, Human Height, Personal Security (Homicides), Political Institutions, Environmental Quality and Gender Inequality, together with a composite indicator summarising developments in all well-being dimensions. It presented as a result the first overview of the evolution of global well-being and global inequality over the long time-period 1820-2000.

Aim of this report

The present report continues this line of work, and develops an even more ambitious research agenda. One implication of the "beyond GDP" debate is that we should not measure inequality within countries – and at the global level – on the basis of estimates of income disparities only, but broaden our approach to other aspects of social and economic inequality. In other words, if we are interested in inequality of well-being, we have to try to measure the development of the inequalities of all dimensions of well-being that are, for example, included in the OECD Better Life Initiative. This has been the first goal of this report: to collect and standardise historical data about inequalities of various dimensions of well-being. This is a new and experimental area for historical research. Such data have not been collected systematically in the past, and their use requires the collection and analysis of large new datasets (as presented in the various chapters of this report).[1]

We know rather well what happened to income inequality in the various parts of the world, thanks to the huge literature about this topic (Milanovic, 2016[7]). But, with few exceptions (OECD, 2020[5]), trends in the inequality of other dimensions of well-being have not been studied in the same, systematic way. We list the major advances that are made in this report:

- **Inequality of Longevity** (or of life expectancy) is an obvious candidate for such a study. The recent work by Case and Deaton (2019) about "Deaths of Despair" in the United States, and the current Covid-19 pandemic, which has affected different social classes very differently, are examples of its huge relevance. Chapter 6, authored by Lamar Crombach, Jeroen Smits and Christiaan Monden, presents a dataset that makes it possible to chart this dimension of the inequality of well-being for many countries over the past 200 years (Chapter 6).
- **Inequality of Educational Attainment** has also changed dramatically over time, and is at the same time highly relevant for well-being; in a way, the social-economic structure of society, with all its inequalities, is reproduced by the educational system. Bas van Leeuwen and Jieli Li provide the data to measure this in the very long run (Chapter 7).
- **Wealth inequality** has received renewed attention thanks to the research by Thomas Piketty. But research on global patterns throughout history is still very scarce. Chapter 5 by Guido Alfani and Sonia Schifano present measures for parts of the world economy in the long run.

- The extent of **Extreme Poverty** is another aspect of within-country inequality that is highly relevant for any assessment of well-being, as reflected by its inclusion in both the Millennium Development Goals and, now, the Sustainable Development Goals agreed by world leaders in 2015. Research by Michail Moatsos presents new historical measures of extreme poverty based on the "basic diet" poverty line pioneered by Allen (2017[8]) (Chapter 9).
- **Gender Inequality** is another aspect of within-country inequality that affects the well-being of the whole society. While it was already covered in our 2014 report, it is taken up again in this new book because of its relevance for the research agenda. Selin Dilli, Sarah Carmichael and Auke Rijpma have compiled new data on the issue covering additional dimensions (Chapter 8).

These topics cover many dimensions of inequality, but not all. How unequal is personal security distributed in a society – when the rich live in gated communities and the poor in ghettos infected by crime and drugs addiction? How unequal is access to a clean environment and a rich nature – with, traditionally, the rich residing in clean neighbourhoods upstream of the urban pollution, and the poor in dirty, polluted slums? The long-term trends in hours worked may be surprisingly different – the rich developing from a 19th century leisured class to an elite that is obsessed by hard work, whereas the working class at the same time has drastically reduced its workload since the late 19th century (as shown in Chapter 3). At the same time, the new precariat that has emerged in the rich countries in recent times (and that has, as "informal labour", been the dominant form of labour in large parts of the developing world) does not have access to such privileges, and is often pressured to combine long working hours and very low wages. How have political institutions developed – from the egalitarian ideal "one man one vote" to political systems in which only money matters, and the social "underclass" is strongly discouraged from active political participation? These are some of the other inequalities that do matter for people's lives – to get the full picture of societal inequality of well-being all dimensions count, of course.

The second aim of the present volume is to broaden the scope of the 2014 *How Was Life?* report by including new datasets covering additional dimensions of well-being. This concerns:

- **Working Hours**. Oisin Gilmore presents a new dataset of the Working Week in Manufacturing, providing an insight into work-life balance, a dimension not covered by the first report (Chapter 3).
- New estimates of **GDP per capita**, created as part of the Maddison Project. Jutta Bult and Jan Luiten van Zanden discuss the consequences of integrating a new set of PPPs, the ICP-round of 2011, into the Maddison dataset (Chapter 2).
- **Biodiversity Development** is crucial for understanding the environmental costs of economic development. Thomas van Goethem and Jan Luiten van Zanden provide measures changes in biodiversity through history (Chapter 10).
- While **Social Transfers** are not a well-being dimension *per se*, they represent one of the key policy levers through which governments have attended to the living conditions of their citizen. Peter Lindert presents evidence on the historical evolution of social transfers as a share of GDP throughout most of the world (Chapter 4).
- Finally, **Multi-dimensional measures of welfare** (Chapter 11) are discussed by Auke Rijpma in the final chapter, which also helps to summarise some of the key results of the volume.

The methodological underpinnings for approaching well-being as done in this report are most clearly expressed by Sen (1993[9]), (1985[10]) and Nussbaum (2000[11]). Sen's theoretical framework is based on the distinction between functionings and capabilities. Functionings can be interpreted as actual achievements of a person, i.e. what he or she manages to do or be; they comprise an individual's activities and his or her states of being, for example, being in good health or being able to move freely. On the other hand, capabilities are the individual's abilities to achieve these functionings, e.g. the person's freedom to choose between different ways of living (Kuklys and Robeyns, 2005[12]). This approach moves away from the traditional utility or resource-based views of well-being that relates welfare to income, enabling us to view life as a combination of various "doings and beings", with quality of life assessed in terms of the capability to achieve valuable functionings (Sen, 1993[9]). Nussbaum (2000[11]) identified ten different groups of capabilities that fit Sen's framework.[2] Some of these capabilities relate to the dimensions of well-being used by the OECD in its *How's Life?* report. For example, Nussbaum's capabilities "Life" and "Bodily health" can, to some extent, be proxied by measures of life expectancy.

The choice of indicators presented in this report has been guided by three considerations: the theoretical literature summarised above, the OECD's Better Life Initiative, and the availability of historical data and international comparative datasets. All chapters in this book present state-of-the-art datasets on historical developments in various dimensions of well-being, many of which are the result of recent research by the Clio-Infra team in co-operation with groups of experts in specific fields (such as in the Maddison project). Most of this research has resulted in publications in international journals (see the lists of publications in the various chapters). While most chapters focus on just one key indicator (e.g. GDP), some present several indicators covering various aspects of the well-being dimension considered (e.g. gender inequality).

The topics covered in this volume mirror the dimensions of well-being covered by the OECD's *How's Life?* report (see Table 1.1). Because of data limitations, this book does not cover housing conditions, subjective well-being or social connections, which were part of the *How's Life?* report. The unit of observation is, in most chapters, the 19th-21st century nation-state.

The data and the datasets underpinning the series are available online on the Clio-Infra website.[3]

Table 1.1. Correspondence between the historical indicators included in *How Was Life?* (I and II) and the well-being dimensions in the OECD's *How's Life?* report

Historical indicators featuring in *How Was Life?* (volumes I and II)	Corresponding dimensions in the OECD *How's Life?* report
GDP (I, II) Income inequality (I) Wealth inequality (II) Extreme poverty	Income and wealth
(Measures not available)	Housing
Real wages (I)	Work and job quality
Life expectancy (I) Life expectancy inequality (II	Health (human capital)
Working week - manufacturing (II)	Work-life balance
Personal security (I)	Safety
Institutions (I)	Civic engagement
(Measures not available)	Social connections (social capital)
Educational attainment (I) Educational inequality (II)	Knowledge and skills (human capital)
Environmental quality (I) Biodiversity (II)	Environmental quality (natural capital)
(Measures not available)	Subjective well-being
Gender inequality (I, II)	
Social transfers (II)	
Composite indicator of well-being (I, II)	(Better Life Index)

Note. The framework underpinning the *How's Life?* report distinguishes between dimensions of current well-being, and resources (in the form of different types of capital) needed to ensure sustainability, while no such distinction is implemented in the *How Was Life?* reports; "resources" for sustainability are indicated (in parenthesis) in the right-hand column of the table. The *How's Life?* framework also considers "inequalities" as a cross-cutting aspect, rather than a separate dimension. While no composite indicator is included in *How's Life?,* the Better Life Index is a communication tool that allows users to set their weight to each of the 11 dimensions of the *How's Life?* framework, to derive a summary view of overall well-being in different countries; the Better Life Index is also indicated (in parenthesis) in the table above.

Practical issues

One of the recurrent issues in historical research is the changing borders of countries. For example, Germany in 2013 is different from Germany in 1989, 1938 or 1913 – not to mention Germany before 1871 or 1798. This applies to almost all the countries covered in this report, although not to the same extent. Maddison created a dataset for GDP and population that took the borders of 1990 as the starting point and tried to correct historical data for past changes in national borders whenever these occurred. The idea was to create a consistent set of estimates of countries based upon the 1990 borders. This report follows this approach whenever possible.

The unit of observation is, as in most comparable studies, the 19th-21st century nation-state. We present values for individual countries and, based on these, for broad world regions and, if possible, for the world as a whole. We aggregate country data weighted by their populations, thus giving China a much bigger weight than Belgium or Nepal. In the presentation of the data, we focus on long-term trends; hence, we present most estimates in the form of averages over a ten-year period, where the 1820s is the average from 1820 to 1829. In other cases, where we do have annual observations – such as GDP estimates – we present these as such (i.e. without averaging across the decade); this is always clearly indicated in the tables included in this book, i.e. "1820" relates to the annual observation of that year.

We also concentrate on regional averages, distinguishing between the eight world regions that already featured in Angus Maddison's analysis. These are Western Europe, Eastern Europe and the former Soviet Union, the Western Offshoots, East Asia, South- and Southeast Asia, the Middle East and North Africa, Sub-Saharan Africa, and Latin America and the Caribbean.[4] Eastern and Western Europe are divided by the former “iron curtain”, a somewhat arbitrary border that is, however, conveniently almost identical to Hajnal’s line used in many economic-historical and demographic studies as the border between different family systems in Europe. Eastern Europe, as defined in this report, also includes the territory of the former USSR, including its Asian parts. The Western Offshoots consists of the United States, Canada, Australia and New Zealand, regions of immigration from Western Europe that shared a common developmental path. East Asia consists of China, Japan, Taiwan, Korea and Hong Kong, while South- and Southeast Asia covers the rest of Asia, except for countries to the west of Afghanistan, which are part of the Middle East. The Middle East and North Africa region covers all African countries bordering the Mediterranean, Iran, the Arab world and Turkey. Finally, Latin America and the Caribbean consists of the Americas except for the United States and Canada.

Some of these regions are dominated by one large country (China towers over East Asia, the United States has a huge weight among the Western Offshoots). To deal with this, we present in this book tables with data for a sample of the 25 focus countries that together cover a large part of the world’s historical population. Regional averages are based on all country observations that are available in the Clio-Infra database for the region concerned. Going back in time, the number of observations is generally declining, and simple algorithms are used to create a set of consistent regional (and, on that basis, global) estimates. Usually it is assumed that “missing countries” develop in the same way as countries in the region for which data are available, and sometimes as countries with a similar economy.

To assess the quality of the historical data, each chapter includes a table that presents our assessment of the reliability of the statistics used. Data quality is broadly assessed in terms of ***credibility*** (the degree to which the sources of the data can be confidently relied on); ***accuracy*** (the extent to which the data are valid and reliably represent what they purport to measure); and ***comparability*** across countries (the extent to which data from different sources are collected under the same methodology and measure the same thing). Four classes of data are distinguished:

1. **High-quality** data are the product of official statistical agencies (national or international) or derived by techniques that ensure equivalent credibility; this implies that high standards of accuracy are maintained, and data are collected based on a consistent methodology across countries.
2. **Medium-quality** data are the product of historical research using the same sources and methods as applied by official statistical agencies; comparability is deemed to be generally good, but differences do exist (across countries and over time) that limit comparability.
3. **Moderate-quality** data result from historical research that relies on indirect data (for example, data of the proceeds of a tax on production are supposed to reflect the “underlying” output) and estimates, resulting in some loss of accuracy; also, not all country estimates are based on data collected with the same methodology.
4. **Low-quality** data are estimates based on various proxies. There may be significant inconsistencies between countries or gaps in coverage.

Much effort has been put into making the data comparable in time and space. However, we are aware of the fact that all data, contemporary and historical, have their weaknesses, and that international comparison – in particular on the global scale that we are doing here – is a “high risk, high gain” business. Even contemporary estimates of GDP are not beyond criticism, as regular revisions demonstrate, and margins or error are getting larger going back in time. We have always tried to collect and compare the best data available, but as Maddison’s work and the Maddison project demonstrate, this is an on-going and cumulative process: gradually, rough estimates (which Maddison called “guestimates”) are replaced by more accurate ones based on more research, and comparisons in time and space help to filter out the highly implausible ones. This report should therefore not be seen as the endpoint of this kind of research, as many data will be refined and improved by future work, but as a review of the current state of the evidence from historical research about long-term trends in well-being in the world economy over the past 200 years.

Introduction to individual chapters

The first part of the volume deals with between-country inequality. It investigates the aggregate levels of well-being of each country and compares countries with one another as if they were uniform entities. The volume opens with a discussion of the new estimates of GDP, the workhorse of all well-being studies. World **GDP per capita** increased by a factor of 13 between the 1820s and the 2010s. As the total population increased seven-fold, total GDP went up by a factor of 90. But this growth was spread very unevenly over the globe: until about 1950, the rich countries grew faster than the poor ones, resulting in a huge increase in international income disparities. Maddison based his seminal study of GDP on 1990 purchasing power parities and prices. For 2011, a new dataset of global prices is now available through the International Comparison Programme (ICP). The chapter investigates the consequences of using these alternative prices when computing historical estimates of GDP for the years before 1950. It argues that simply moving from the 1990 to the 2011 price benchmark does not improve estimates for the historical period, probably due to the greater distance between 2011 and the historical benchmark estimates in the dataset. The evidence summarised in the chapter suggests that the 1990 benchmark created by Maddison is probably the best compromise solution for historical estimates, in particular when used in combination with the 2011 PPPs for the most recent period.

Chapter 3 deals with one of the crucial missing elements in the first *How Was Life?* report (van Zanden et al., 2014[1]). Working hours and time off from work are central to people's work-life balance. When considered globally, we observe substantial disparities in working time, with people in some countries working significantly more than in others. Amongst OECD countries, the **working week in manufacturing** has declined dramatically since the 19th century. Full-time workers in manufacturing worked 60 to 90 hours per week in the 19th century. Today they work roughly 40 hours. In the second half of the 20th century, the Middle East and North Africa, as well as Sub-Saharan Africa, experienced an increase in working time in manufacturing. In the case of the Middle East and North Africa, the working week rose to close to 55 hours.

Chapter 4 presents estimates of the value of **social transfers.** While these transfers typically aim at addressing various risks that people face during their lives, they also play an important role in mitigating within-country inequalities. Such efforts were tiny or non-existent until political developments assigned greater responsibility for addressing more of these risks to the government in the last hundred years. In 1900, no national government transferred more than 3% of GDP via social programmes. This has increased gradually in most countries, to a level in some of them above 15% in 1970, and to around 30% today. While initially governments transferred most of these funds to support young people attending school, gradually the make-up of these transfers has shifted to old-age pensions and support to the elderly. Since the richest countries tend to have the most developed social programmes, social safety nets tend to be strongest where they are least needed on social and economic grounds.

Chapter 5 pushes forward the frontier of historical research on within-country economic inequality by investigating **wealth inequality,** as measured by the Gini coefficient and by the wealth share controlled by the richest 10 percent of the population. Different regions of the world have followed different long-term trends in wealth inequality. While in Western Europe wealth inequality has decreased in the past two centuries, in the Western Offshoots it increased, a tendency stopped only by World Wars I and II and by the troubled times in-between. Despite this difference, one can see a general pattern, with an overall tendency for wealth inequality to increase during the 19th century, and seemingly during the first decade of the 20th century. There is no clear correlation between the level of GDP per capita and wealth inequality.

Chapter 6 focuses on **inequality in** one of the crucial outcomes of economic development, i.e. length of life. Until 1900, **life expectancy** in the best-performing countries was limited to 65 years for men and 67 years for women. For men, it then increased to about 72 around 1950, where it plateaued until about 1970, after which it increased gradually to reach its current value of around 82. For women, the increase was more continuous, although a slight slowdown is visible after 1950, reaching the current value of about 87 in the best-performing countries. Throughout the period, women had a higher life expectancy than men, a gap that reached 13 years in Russia in the 2000s. Inequality in life expectancy, as measured by the Gini index, is much lower than for income, and has decreased over time with only few exceptions (Mexico and Egypt in the 2000s and 2010s). Moreover, inequality in life expectancy has changed relatively little in countries such as India, Indonesia, Nigeria, Argentina, Brazil and Mexico, while it is higher in the United States compared to other developed countries.

Chapter 7 investigates **inequalities in educational attainment**, a factor that has crucial importance for equality of opportunity. The Gini coefficient in the years of schooling of the population declined significantly over the 19th and 20th centuries. This reduction was caused mostly by the fall in the share of people without any educational qualification. While the Gini coefficient is a measure of relative inequality (when the educational attainment of everyone doubles, the Gini coefficient does not change), a measure of absolute inequality such as the standard deviation highlights the existence of an educational Kuznets's curve: whereas initially a rise in education led to more inequality, after a certain point inequality starts to decrease. While most countries reached this tipping point in the 20th century, other countries are still in the "rising part" of the curve, implying that they confront a trade-off between stimulating education and reducing educational inequality.

Chapter 8 investigates **gender inequality** – a major determinant of well-being for, at least, half of the population – providing new estimates dating back to 1900 and sometimes earlier. These include new historical measures of the gender gap in wages and unemployment in the 20th century as well as new data on the share of female-headed households. While the post-World War II era has been traditionally seen as a period of progressive narrowing of gender gaps, the new evidence suggests that progress towards greater gender equality had started already in the pre-War era.

Chapter 9 discusses perhaps the most important benefit of economic growth – the reduction of **extreme poverty** – using a metric (based on the inability to purchase a near-subsistence basket of goods and services) that overcomes some of the limitations of the traditional 1.9 USD-a-day standard. Based on this measure, 53% of the world's population lived in extreme poverty in 1950, which is 23 percentage points lower than the level in 1820. By 1990, the rate had dropped further to 31%. The fall of extreme poverty continued in the following years, and by 2018 global extreme poverty had dropped to 10%, with reductions in India and China as the main drivers. Countries continue to have differing levels of extreme poverty. In Western Europe and the Western Offshoots, extreme poverty fell below 1% as early as the 1970s. When considering other world regions, the biggest declines include Russia (from a staggering 98% in 1820 to 2% in 2018) and Japan (where extreme poverty was eradicated by 1975 from levels as high as 95% in 1820). However, by 2018 the absolute number of persons living in conditions of extreme poverty was on par with that in 1820, at about three-quarters of a billion people.

Chapter 10 pushes the frontier of knowledge about the historical interaction between human activity and the natural environment by focusing on **biodiversity loss.** Globally, species populations declined by 36% to 52% in abundance between 1970 and 2010. But historical records suggest that biodiversity has been declining at least since 1500, and probably for the better part of the Holocene. A case study reconstructing biodiversity change in the Netherlands from 1900 onwards based on different assemblages of species shows a long-term decline in biodiversity between 1900 and 1970 and a partial recovery since the 1980s. Population growth, the intensification of agriculture, pollution and the expansion of infrastructure are identified as the key human drivers of biodiversity loss in the Netherlands.

The final chapter synthesises the results of the report through two **composite indices of well-being**. The first is an overall indicator compiled as the average of the average well-being measures of each country presented in the current and previous volumes (i.e. a "mean-of-means"). Adding the new average well-being measures included in this book does not fundamentally change the global picture drawn in 2014: that there was great progress throughout the 1820-2010 period, which is more evenly distributed across countries than in the case of GDP per capita. The estimated indices of well-being show less divergence than GDP per capita, and therefore smaller global differences, as well as considerable convergence of large parts of the world (East Asia, Eastern Europe and the Middle East). However, in particular Africa did lag behind relative to the rest of the world, although well-being there did also increase substantially.

The second is a composite indicator of the within-country measures of well-being inequality (i.e. a "mean-of-inequalities"), which weights four different inequality measures (inequality of educational attainment, life expectancy, income and gender). The long-term declines in three aspects (income inequality being the exception) means that this composite inequality indicator has declined in many parts of the world since 1820, with a stronger reduction in Africa and Asia compared to Western Europe and its Offshoots. The 20th century therefore featured both large improvements of average levels of well-being and massive reductions in well-being inequality in life expectancy, educational attainment and gender relations. The decline in income inequality, which characterised the 20th century, lasted only until about 1970, and was followed by an often dramatic reversal. While we still live in a highly unequal world, where the place of birth – via the citizen rent – and other non-personal features – gender, race, religion, social class – strongly affect the choices that people have to lead a better life, this report has identified at least some changes that point to convergence and reductions of social and economic inequalities.

References

Allen, R. (2017), "Absolute poverty: When necessity displaces desire", *American Economic Review*, Vol. 107/12, pp. 3690-3721. [8]

Hills, T. (2019), "Historical analysis of national subjective wellbeing using millions of digitized books", *Nature Human Behaviour*, Vol. 3, pp. 1271-1275. [6]

Kuklys, W. and I. Robeyns (2005), "Sen's Capability Approach to Welfare Economics", in *Capability Approach. Studies in Choice and Welfare*, Springer, Berlin, Heidelberg, https://doi.org/10.1007/3-540-28083-9_2. [12]

Milanovic, B. (2016), *Global Inequality: A New Approach for the Age of Globalization*, Harvard UP, Cambridge, Mass. [7]

Nussbaum, M. (2000), *Women and Human Development: The Capabilities Approach*, Cambridge University Press, Cambridge. [11]

OECD (2020), *How's Life? 2020: Measuring Well-being*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9870c393-en. [5]

Piketty, T. (2014), *Capital in the Twenty-First Century*, Harvard University Press. [14]

Sen, A. (1993), *"Capability and well-being"*, Clarendon Press. [9]

Sen, A. (1985), "Well-being, agency and freedom: The Dewey lectures 1984", *The Journal of Philosophy*, Vol. 82/4, pp. 169-221, https://doi.org/10.2307/2026184. [10]

Stevenson, B. and J. Wolfers (2008), "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox", *Brookings Papers on Economic Activity*, Vol. 2008, 1, pp. 1-87, https://doi.org/10.1353/eca.0.0001. [13]

Stiglitz, J., A. Sen and J. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*. [2]

UNDP (2013), *Human Development Report 2013. The Rise of the South: Human Progress in a Diverse World*, http://hdr.undp.org/en/content/human-development-report-2013. [4]

UNDP (1990), *Human Development Report 1990: Concept and Measurement of Human Development*, http://www.hdr.undp.org/en/reports/global/hdr1990. [3]

van Zanden, J. et al. (eds.) (2014), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-en. [1]

Notes

[1] A focus on well-being inequalities (beyond income), their causes and consequences also features in the Deaton Review of Inequalities launched by the IFS (Institute for Fiscal Studies) in May 2019 (https://www.ifs.org.uk/inequality/about-the-review/).

[2] These capabilities are: (1) life, (2) bodily health, (3) bodily integrity, (4) senses, imagination and thought, (5) emotions, (6) practical reason, (7) affiliation, (8) other species, (9) play and (10) control over one's environment.

[3] www.clio-infra.eu.

[4] The regions used in this report are those used by Maddison in his own work for the OECD.

2 The long view on economic growth: New estimates of GDP

Jutta Bolt, University of Groningen and
Jan Luiten van Zanden, Utrecht University

GDP estimates form the backbone of our understanding of economic change in the past and the present. This chapter presents and discusses the results of an update of the Maddison Project, which aimed to incorporate the results of the 2011 round of the International Comparison Programme (ICP), and of the efforts of many scholars who have extended and deepened the work on historical national accounts. A method is presented to test the reliability of various approaches to back-project historical time series of GDP, and it is concluded that Maddison's 1990 benchmark still produces the most plausible results. The implications for this for the growth of the world economy since 1820 are discussed.

Introduction

This chapter presents long-term trends in GDP per capita in a global perspective. GDP per capita is both an important indicator for measuring the economic performance of countries and a crucial driver of economic well-being. In this chapter, we build on the work by Angus Maddison and the Maddison Project, and add to the existing work in two ways. First, we include all the new historical national account estimates that have become available since the publication of *How Was Life?* in 2014 (van Zanden et al.[1]). Second, we explore the consequences of using alternative measures of relative price levels for reconstructing the shape of long-run GDP growth at the global level, developing a way to assess the biases associated with the different alternatives. Finally, we present and discuss the global trends in GDP per capita since 1820 using all the new historical estimates that are currently available. As will become clear, using an alternative set of relative prices for reconstructing historical series of GDP per capita has important implications for our understanding of long-run economic development. This justifies a more in-depth discussion of relative price estimates and an analysis of the consequences of using different relative price estimates. In doing so, the structure of this chapter deviates slightly from that of the other chapters in this volume.

In order to compare GDP per capita levels both across countries and over time, it is necessary, first, to express all countries' GDP estimates in a common currency and, second, to take into account the differences in price levels and structures between countries. The original Maddison and Maddison Project series are based on a single set of cross-country comparisons of relative income levels in 1990 using 1990 Purchasing Power Parities (PPPs). These relative income levels are then projected forwards and backwards using historical data on the real growth of GDP per capita. Since then, newer and arguably better Purchasing Power Parity (PPP) benchmarks have become available, most specifically the 2011 benchmark (Deaton, 2011[2]). In this chapter, we compare historical GDP per capita estimates based on these 2011 PPPs with estimates of relative levels of GDP per capita in the period before 1940 based on the "traditional" Maddison approach, which makes use of 1990 PPPs. Further, we compare these results with a third approach, building on the estimates provided by the Penn World Tables (2011[3]) for the period from 1950 to the present, linking the historic (pre-1950) time series to the PWT estimates for that year. By comparing these three approaches with independent benchmarks of relative levels of GDP per capita for groups of countries in the 19th and early 20th century, we can get an idea of the biases involved in the three approaches, and of the consequences of using alternative relative prices for our understanding of long-run global development.

The focus of this chapter is on the consequences of using alternative measures of relative prices for measuring per capita GDP for the years before 1950. We test various ways of back-projecting historical estimates of GDP per capita against a dataset of independent benchmark estimates of relative levels of GDP per capita for the period before 1940. We find that simply moving from the 1990 to the 2011 PPP benchmark does not improve GDP estimates for the historical period, probably due to the greater distance between 2011 and the historical benchmark estimates in the dataset. The PWT approach, which uses all post-1950 PPP 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 in the earlier years, which is also problematic. We conclude that the 1990 benchmark created by Maddison is probably the best compromise solution for historical analysis, in particular when used in combination with the 2011 PPPs for the post-1990 period and with the various historical benchmark estimates of relative levels of GDP that are available in the literature. We propose such a hybrid dataset as the best way to use the available information on long-term economic growth and relative levels of GDP per capita. Moreover, the estimates presented here are used in the other chapters to analyse the correlation between the other variables presented and GDP per capita, and the underlying nominal values are, for example, used to estimate the share of social transfers in GDP in Chapter 4.

Description of the concepts used

This chapter centres on Gross Domestic Product (GDP) per head of the population for understanding long-term trends in economic well-being. We also include a discussion of prices between countries, as both the structure of relative prices and the absolute level of prices are correlated with levels of economic development. On average, price levels are higher in more developed countries. If these differences are not taken into account when comparing GDP levels between countries, the output and incomes of developed countries are overestimated and those of developing countries are underestimated. Finally, we include a short discussion of historical comparisons of relative income. These direct estimates of past relative income levels between countries provide benchmarks against which we can compare our GDP per capita series to get an idea of which relative price levels produce the most consistent historical income series for a global set of countries.

Gross Domestic Product per capita

There are three basic ways to measure a country's GDP: as the total income earned by residents in a country, as the sum total of final expenditures and as the total output of goods and services produced in a country.[1]

The Income approach to measuring GDP

The income approach to measuring a country's GDP sums the total income earned by all households and firms within a country in a given year. This includes incomes received by residents as wages, rents, profits and interest incomes. Due to the limits of historical information, no adjustments are generally made to account for income flows accruing to residents from abroad or for income from domestic production paid abroad.

The Expenditure approach to measuring GDP

The expenditure approach to GDP sums up the total value of all domestic expenditures made on final goods and services within a country in a given year. This includes consumption expenditure, government expenditure, investment expenditure and net exports. The expenditures made on final goods and services equal the incomes received as wages, rents, interest incomes and profits.

The Production approach to measuring GDP

The final method of determining GDP sums the value-added generated from the production of goods and services within a country in a given year. This method, also known as the Value-added method or the Net Product method, is based on the production of all sectors in the economy and involves three steps. First, the gross value of output is estimated for all sectors. Second, the intermediate consumption for each sector, such as the cost of intermediate inputs used in the production of final output, is derived. Third, net production is determined by deducting the intermediate consumption from the total gross output.

The construction of historical national accounts relies on all three methods discussed above and combines various pieces of information to obtain a GDP per capita measure.[2] Comparing GDP per capita levels both across countries and over time requires first that the data are expressed in a common currency, and second that differences in price levels and structures between countries are taken into account.

Relative prices

Price differences between countries are important as, generally, price levels are correlated with levels of economic development. This means that on average prices are higher in more developed countries, so that any currency can buy less of the same good in a developed country than in a developing country. As the trading of goods makes prices converge between countries, the price differences between developing and developed countries are typically largest for non-traded goods such as domestic transport, housing and many services. If these differences are not taken into account when comparing GDP levels between countries, the incomes of developed countries are overestimated, and those of developing countries are underestimated.

The original Maddison data are expressed in 1990 constant prices (Maddison, 2003[4]). In other words, a benchmark year PPP is estimated for the year 1990, which is used to compute GDP levels for all countries for that year. From this benchmark year, the GDP per capita series for all countries are extrapolated (backward and forward) by using volume growth rates of GDP for the countries included in the set. 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. To circumvent this problem, he used information from earlier rounds of the International Comparison Programme (ICP) of the World Bank, or he used estimates from the PWT or other proxies (Maddison, 2006[5]), p. 610. In his final set of estimates (Maddison, 2006[5]), data for 43 countries (representing almost 80% of world GDP at the time) were based on ICP or ICP-equivalent estimates; those for the other 113 countries ("non-sample") were based on PPPs from the Penn World Tables (2011[3]) and on proxy estimates. The PWT PPPs were, in turn, estimated by Summers and Heston (1991[6]) based on cost-of-living estimates for expatriates and foreign diplomats.

Since 1990, however, there have been multiple ICP rounds, which have resulted in more up-to-date and, especially in the case of the 2011 round, more accurate estimates of relative and absolute prices for a common basket of goods and services. One of the aims of this chapter is to explore whether these more up-to-date PPP benchmarks also lead to improved long-run comparisons of per capita GDP.

Finally, the most recent version of the Penn World Tables uses a new methodology that no longer relies on a single PPP benchmark but instead integrates all the official ICP PPP benchmarks that have become available since 1950 (Feenstra, Inklaar and Timmer, 2015[7]). For example, for 1980 the methodology takes as the "best" estimate of PPP the value from the ICP round for 1980, while for 2005 it relies on PPPs from the ICP round of 2005. Income estimates of different countries in a common currency for 1980 and 2005 are then generated based on these two sets of price data.[3] The GDP per capita series are subsequently tied to these relative income levels. An advantage of this approach is that it takes into account the relative price changes between the different benchmark years. The major disadvantage is that the growth rate of GDP per capita between two benchmarks is determined not only by the national accounts statistics of the country involved, but also by the differences between the PPPs of earlier and later years. Moreover, this methodology assumes that all ICP rounds are correct, which may be problematic, especially for the earlier ones. In this chapter, we explore the implications of using all available PPPs for constructing long-term GDP series, and analyse how this approach compares to the GDP per capita series based on a single benchmark PPP, e.g. the 1990 or the 2011 PPP rounds.

Historical GDP Benchmarks

Benchmark studies in this field essentially examine the ratios of incomes or output between two or more countries in the past. For each country in the comparison, the performance of the economy is measured in a consistent way for a given year, making it possible to compare levels of GDP per capita across countries for that year (for example, a benchmark study may compare GDP per capita in Germany, the United Kingdom and the United States for 1913). We collected the available benchmark studies for the pre-1940 period (the post-1950 benchmarks are already available in the PWT) and used them to anchor the estimates of GDP per capita, which were based on back-projections of the 1990 (Maddison, 2003[4]; Maddison, 2006[5]; Maddison, 2007[8]) or 2011 benchmarks.

Historical sources and data quality

The basis for the GDP per capita series presented in this chapter is the same as for the *How Was Life?* report, i.e. sourced directly from the Maddison Project (Bolt and van Zanden, 2020[9]). The present dataset additionally includes all new historical estimates that have become available since then, as well as estimates for the most recent years up until 2016. The underlying historical sources for this overview are the following. For the recent period, national statistical agencies produce estimates of GDP and its components that are harmonised and standardised by various international organisations (OECD, United Nations, World Bank), which then feed in secondary datasets such as the Penn World Tables[4] and the Maddison dataset. The PPPs used to convert GDP measures (expressed in current prices of a particular country) into an international standard are derived from the ICP programme organised by the World Bank.[5] Historical studies covering the period before 1950 (or before 1913, since some statistical agencies have done work on the first half of the 20th century as well) use a large variety of sources to reconstruct the development of national income and product. This includes labour force and production censuses, tax records, data on international trade, wage and price data from different sources, etc.

In general, the accuracy of GDP estimates decreases going back in time with the declining availability of modern statistical sources. But, as explained earlier, historians use a broad range of sources (on the total output of the economy, total expenditures and total income earned) as pieces in a large jigsaw puzzle: historians almost never have all the pieces, but quite often they have enough to get a good picture of the economy concerned.

For the recent period, the most important new work included in the dataset is Harry Wu's reconstruction of Chinese economic growth since 1950, which produces state-of-the-art estimates of GDP and its components for this country (Wu, 2014[10]). Given the large role that China plays in any reconstruction of global economic development, this is a major addition to the dataset. Moreover, as we shall see below, these revised estimates of GDP growth are in general lower than the official estimates produced by the Chinese Statistical Office. Lower GDP growth between 1952 and the present, however, substantially increases the estimates of the absolute level of Chinese GDP in the 1950s (given the fact that their absolute level is determined by a benchmark in 1990 or 2011) (see also Figure 2.3). We will revert to this point below.

Most of the other additions to the Maddison Project dataset relate to the period before 1914 (Table 2.1). Again, important new work has been done for China, in particular by Broadberry, Guan and Li (2018[11]) and Xu et al. (2016[12]). 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.

As is clear from this overview, historical research – in particular work on the early modern period (1500-1800) – is producing new time series of per capita GDP, often making use of indirect methods to estimate its long-term development. The "model" for making such estimates – based on the links between real wages, the demand for foodstuffs and agricultural output, as developed by Malanima (2011[13]), (Álvarez-Nogal and De La Escosura (2013[14]) and others – has now also been applied to Poland (Malinowski and van Zanden, 2017[15]), Spanish America (Arroyo and van Zanden, 2016[16]) and France (Ridolfi, 2016[17]). A detailed overview of these new estimates is provided by Bolt and van Zanden (2020[9]).

Table 2.1. New GDP estimates integrated into the Maddison Project Database

	Country	Period	Source
Latin America	Bolivia	1846-1950	Herranz-Loncán and Peres-Cajías (2016[18])
	Brazil	1850-1899	Barro and Ursua (2008[19])
	Chile	1810-2004	Díaz, Lüders and Wagner (2007[20])
	Cuba	1902-1958	Ward and Devereux (2012[21])
	Cuba	1690-1895	Santamaria Garcia (2005[22])
	Mexico	1550-1812	Arroyo and van Zanden (2016[16])
	Mexico	1812-1870	Prados de la Escosura (2009[23])
	Mexico	1870-1895	Bèrtola and Ocampo (2012[24])
	Mexico	1895-2003	Barro and Ursua (2008[19])
	Panama	1906-1945	De Corso and Kalmanovitz (2016[25])
	Peru	1600-1812	rroyo and van Zanden (2016[16])
	Peru	1812-1870	Seminario (2015[26])
	Uruguay	1870-2014	Bèrtola (2016[27])
	Venezuela	1830-2012	De Corso (2013[28])
Europe	England	1252-1870	Broadberry et al. (2015[29])
	Finland	1600-1860	Eloranta, Voutilainen and Nummela (2016[30])
	France	1250-1800	Ridolfi (2016[17])
	Holland	1348-1807	an Zanden and van Leeuwen (2012[31])
	Italy (north)	1310-1871	Malanima (2011[13])
	Norway	1820-1930	Grytten (2015[32])
	Poland	1409-1913	Malinowski and van Zanden (2017[15])
	Portugal	1530-1850	Palma and Reis (2019[33])
	Romania	1862-1995	Axenciuc (2012[34])
	Spain	1850-2016	Prados de la Escosura (2017[35])
	Sweden	1300-1560	Krantz (2017[36])
	Sweden	1560-1950	Schön and Krantz (2015[37])
	Switzerland	1850-2011	Stohr (2016[38])
	United Kingdom	1700-1870	Broadberry et al. (2015[29])
Asia	People's Republic of China	1952-2008	Wu (2014[10])
	People's Republic of China	1661-1933	Xu et al. (2016[12]); Broadberry, Guan and Li (2018[11])
	India	1600-1870	Broadberry, Custodis and Gupta (2015[39])
	Japan	724-1874	Bassino et al. (2018[40])
	Japan	1874-1940	Fukao et al. (2015[41])
	Korea, Republic of	1911-1990	Cha et al. (2020[42])
	Korea, DRP of	1911-1940; 1990-2015	Cha et al. (2020[42])
	Malaysia	1900-1939	Nazrin (2016[43])
	Turkey	1500-1820	Pamuk (2009[44])
	Singapore	1900-1959	Sugimoto (2011[45])
Middle East	Syria	1820, 1870, 1913, 1950	Pamuk (2006[46])
	Lebanon	1820, 1870, 1913, 1950	Pamuk (2006[46])
	Jordan	1820, 1870, 1913, 1950	Pamuk (2006[46])
	Egypt	1820, 1870, 1913, 1950	Pamuk (2006[46])
	Saudi Arabia	1820, 1870, 1913, 1950	Pamuk (2006[46])
	Iraq	1820, 1870, 1913, 1950	Pamuk (2006[46])
	Iran	1820, 1870, 1913, 1950	Pamuk (2006[46])
Africa	Cape Colony/ South Africa	1700-1900	Fourie and van Zanden (2013[47])

The quality of the National Accounts estimates made by official statistical agencies is in general high (Table 2.2). However, some problems do remain even for the most recent period. Regular revisions of GDP estimates by these agencies – the result of new information and/or revisions of the internationally accepted System of National Accounts (SNA) – sometimes result in breaks in historical time series that limit comparison in time and space. The quality of the official GDP statistics, in turn, is related to the quality of censuses and, more generally, to the capacity of the state to register and “monitor” the population and its economic activities. In particular new, relatively weak states may be unable to produce state-of-the-art estimates of national accounts, or may have an incentive to underestimate their economic performance, for example, to qualify for certain forms of international aid.

Moreover, as will be discussed below, international comparisons are constrained by the limitations of the PPP approach and of the various ICP rounds. Modern statistical work in this field began on a global scale in the 1950s. Almost all estimates for the period before the 1940s are the result of historical research in this area, which began in earnest in the 1950s. The classification of the quality of the historical GDP estimates presented in Table 2.2 is subjective, based on what is known about available sources and studies (for a recent overview see Bolt and van Zanden (2014[48])). Most problematic are the estimates for sub-Saharan Africa, which, with the exception of the estimates for South Africa, are highly speculative for the pre-1950 period and, also, comparatively weak for the second half of the 20th century. Western Europe, the Western Offshoots and Japan have the highest quality data for the 19th century.

Table 2.2. Quality of the sources of GDP data by decade and world region

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Sub-Saharan Africa	Middle East and North Africa	East Asia	South and Southeast Asia
1820	3	4	3	4	4	4	3	4
1870	2	3	3	3	4	4	3	4
1920	2	3	2	3	4	3	3	3
1950	1	1	1	1	1	1	1	1
1970	1	1	1	1	1	1	1	1
2000	1	1	1	1	1	1	1	1

Note:1. High quality: the product of official statistical agency (national or international); 2. Medium quality: the product of economic-historical research using the same sources and methods as applied by official statistical agencies; 3. Moderate quality: economic historical research, but making use of indirect data and estimates; and 4. Low quality: estimates based on a range of proxy information. In case of multiple sources, the lowest quality source is given.

Long-run global developments in per capita GDP: Comparing the three approaches

In this section we compare three different long-term GDP series, each one based on different estimates of relative prices. The baseline series is the one included in the original Maddison Project, integrated with the new historical estimates that have become available since 2014. These series are based on the 1990 PPPs, extrapolated backwards and forwards using official national accounts GDP growth rates, in a similar fashion to Maddison (2003[4]) and the Maddison Project 2013 update (Bolt and van Zanden, 2014[48]).

The first alternative series is based on the latest PPP estimates produced by the ICP for 2011 (World Bank, 2014[49]), extrapolated backwards and forwards using the same GDP growth rates (derived from SNA) as for our baseline series using the 1990 benchmark. The second alternative series uses the Penn World Tables approach for the series between 1950 and 2016, extrapolated backwards from 1950 using the national accounts GDP growth rates.

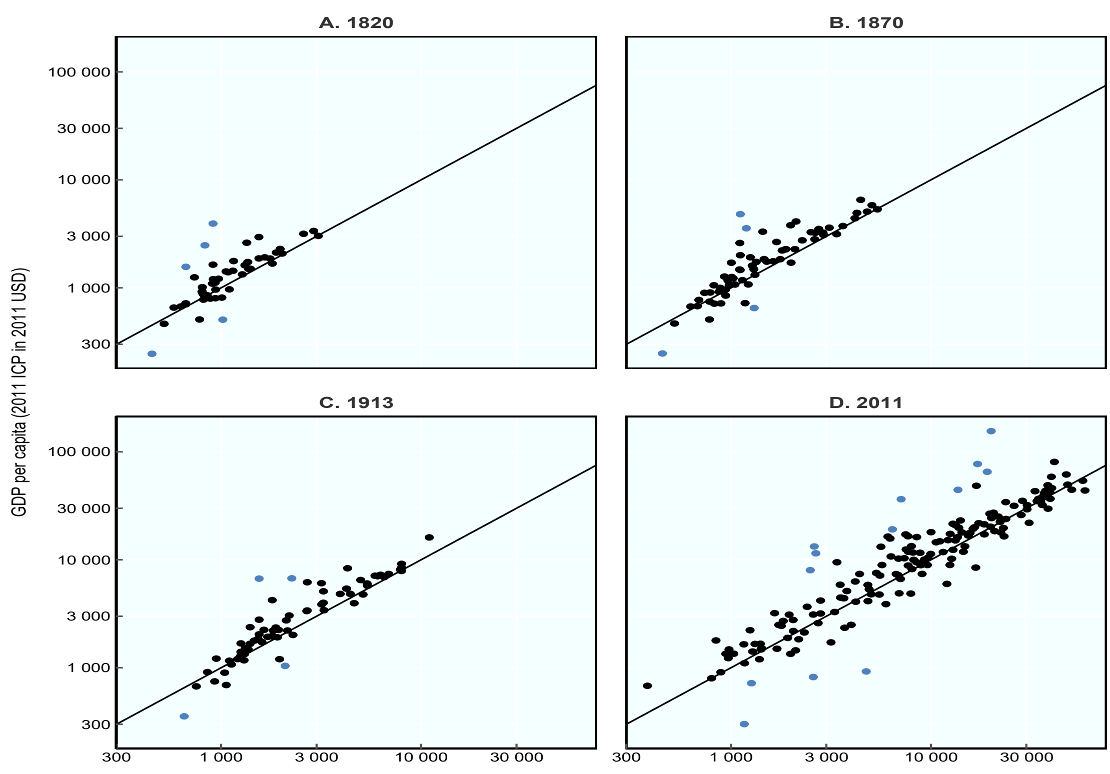
This provides us with three PPP-based estimates of GDP per capita: the first one expressed in 1990 USD using 1990 constant prices; the second one expressed in USD using 2011 constant prices; and the third one expressed in USD using constant prices from the first available benchmark for each country (which varies between individual countries).

The difference between the series using the 1990 PPP and the series using the 2011 PPP is essentially a level shift in GDP estimates. For example, for those countries that experienced an increase in prices between 1990 and 2011 vis-à-vis the United States, the original GDP series in nominal prices produced by the statistical agency of each country have been divided by the higher relative prices in the 2011 PPP series, leading to a lower level of real GDP relative to the United States. And for those countries that have experienced a decrease in prices during that period relative to the United States, GDP estimates shift upwards. In a similar fashion, we can compare the series based on 1990 PPPs and the PWT series using multiple PPP benchmarks. For those countries that experienced an increase in prices relative to the United States between the first benchmark available for that country and 1990, the level of real GDP relative to the United States for the first benchmark was higher than their relative GDP in 1990. This leads to a reshuffling of the level of historical GDP series for various countries.

One way of presenting this comparison is by looking at the consequences of making use of the two benchmarks (2011 and 1990) for the historical GDP estimates (Figure 2.1).

Figure 2.1. GDP per capita in various years based on Maddison 1990 and ICP 2011 PPPs

Values in 2011 US dollars



Note: 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. Darker dots standing out from the general pattern: Angola (2011); United Arab Emirates (2011); Burundi (2011); Bahrain (2011); Cuba (2011, 1820, 1870, 1913); Iraq (2011, 1820, 1870, 1913); Kuwait (2011); Liberia (2011); Libya (2011); Mozambique (2011); Oman (2011); Qatar (2011); Somalia (2011); Algeria (1820); North Korea (1820, 1870, 1913); Syria (1820, 1870, 1913). Blue dots are all the remaining countries.

Source: Bolt and van Zanden (2020[9]), "Maddison style estimates of the evolution of the world economy. A new 2020 update", *Maddison Project Working paper 15*.

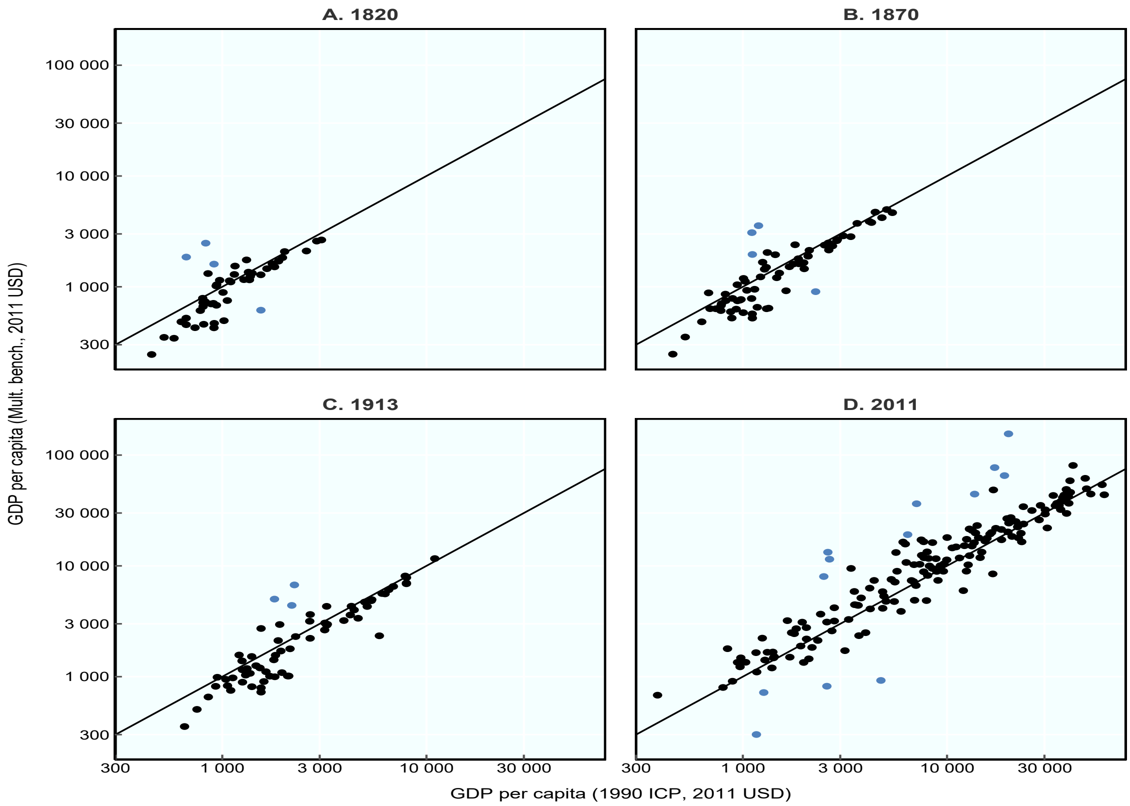
StatLink https://stat.link/rag2bd

The most notable differences between the two estimates relate to the oil-rich countries in the Middle East (Qatar, Kuwait, United Arab Emirates, Bahrain, Oman, Saudi Arabia, Iraq and Syria), highlighted with darker dots in Figure 2.1, whose GDP per capita increases significantly when switching to the 2011 benchmark. This may reflect limits in the Maddison 1990 estimates. These countries were included in the "non-sample" group in the original Maddison database, as they were not covered by the ICP rounds underlying Maddison's 1990 PPP calculations. The price estimates for 1990 used by Maddison (based on information on the prices paid by expatriates in different countries) probably overestimated the actual prices in these countries, and hence underestimated their GDP levels. However, the shift of the GDP levels of the Gulf states may also reflect the strong increase in the relative price of oil between 1990 (when oil prices were rather low) and 2011 (when they peaked), which drove up the relative income levels of these countries. Extrapolating the 2011 GDP levels back into the past, on the basis of the estimated GDP growth rates of these countries, may therefore result in implausibly high incomes in earlier periods, as shown in Figure 2.1. According to this combination of data (the 2011 PPPs and the estimated growth rates of GDP per capita), Iraq would be the wealthiest country in the world in 1820 (and close to that position in 1870), which is implausible. Other notable outliers are, for 1820, 1870 and 1913, Korea (PRK), the Palestinian territory and Lebanon; and Angola, Mozambique, Uzbekistan and Armenia for 2011.

Comparing the 1990 GDP benchmark estimates to the multiple benchmark estimates based on the PWT approach, we find that most countries are located very close to the 45-degree line for all years analysed (Figure 2.2). Notable outliers for 1820, 1870 and 1913 are Argentina and Cuba. For Cuba, the benchmark value of GDP per capita based on the 1990s PPP estimate is around one-third of the level based on the PWT approach. In contrast, for Argentina the benchmark GDP level based on 1990 PPPs is around three times larger than the one based on the PWT approach. Likewise, GDP per capita in Syria and Iran are also much higher using the 1990 PPPs compared to the PWT approach.

Figure 2.2. GDP per capita in various years based on Maddison 1990 and the multiple benchmark (PWT) approach

Values in 2011 US dollars



Note: Values in 2011 USD. Darker dots standing out from the general pattern: Angola (2011); United Arab Emirates (2011); Burundi (2011); Bahrain (2011); Cuba (2011, 1820, 1870, 1913); Iraq (2011, 1820, 1870); Kuwait (2011); Liberia (2011); Libya (2011); Mozambique (2011); Oman (2011); Qatar (2011); Somalia (2011); Argentina (1820, 1870); Algeria (1820, 1870, 1913); Russia (1913).

Source: Bolt and van Zanden (2020[9]), "Maddison style estimates of the evolution of the world economy. A new 2020 update", *Maddison Project Working paper 15.*

StatLink https://stat.link/b06own

The first test: Subsistence income

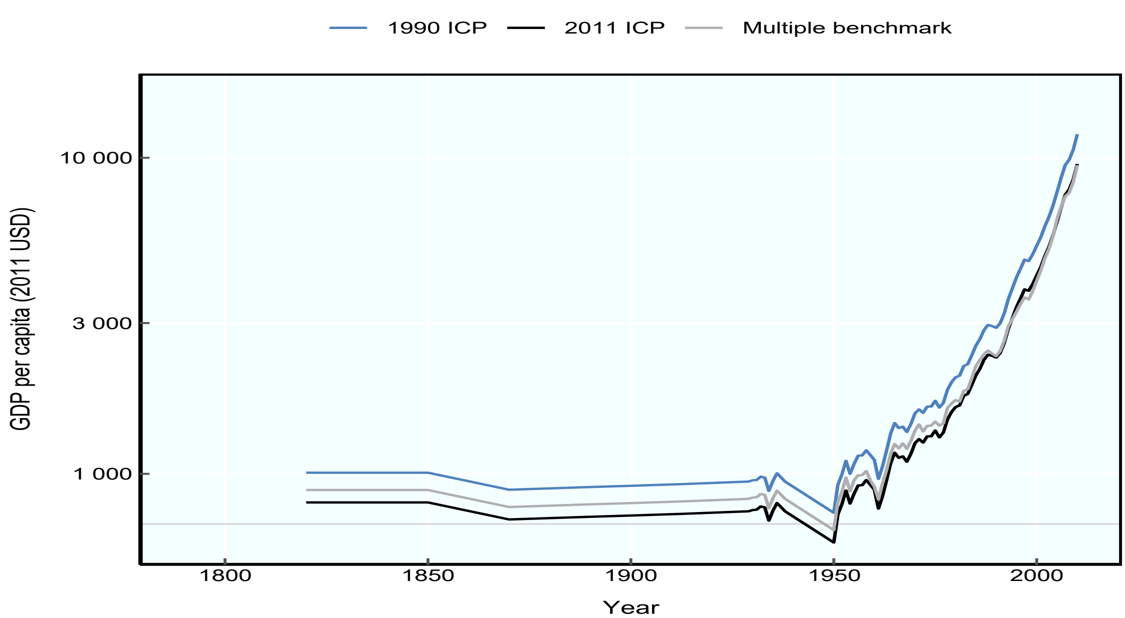
One of the implications of moving from PPPs in 1990 prices to PPPs in 2011 prices is that estimates of subsistence income change. With the 1990 price levels, subsistence income was between USD 350 and 400 per year (Maddison, 2003[4]). The poverty line therefore was equal to around USD 1 a day, as reflected in the first international poverty line set in 1990 by the World Bank at USD 1.01 per day using 1985 PPPs, later updated to USD 1.08 per day using the 1993 PPPs (Ravallion, Datt and van de Walle, 1991[50]; Chen and Ravallion, 2001[51]). When using other relative prices, this level of subsistence income changes. In 2015, the World Bank moved to a threshold for extreme poverty of USD 1.90 a day (or USD 694 per year) based on the 2011 PPPs.

The effects of using alternative PPPs instead of the original Maddison estimates are most notable for countries that experienced substantial price changes relative to the United States between the benchmark years. China is a case in point. According to the 2005 PPPs, prices for China had increased so much since 1990 relative to those in the United States that its level of GDP per capita was around 40% lower than the level based on earlier price estimates (Deaton and Heston, 2010[52]; Feenstra et al., 2013[53]). This led to implausibly low historical GDP estimates for China, given that the original Maddison's estimates were already very close to subsistence around 1950 (Maddison, 2007[8]).

In the years since the release of the 2005 PPPs, a consensus arose about the shortcomings of the 2005 ICP round, most of which were corrected in the 2011 ICP round. Still, prices in China relative to the United States were substantially higher in 2011 compared to 1990, which lowers China's PPP-adjusted GDP per capita in 2011 by 23%. In this chapter, we have updated China's GDP estimates of the Maddison Project based on Wu (2014[10]), which show lower GDP growth between 1952 and 2011 than the previous (official) estimates. Extrapolating backwards from this lower 2011 base 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 2.3. The 2011 PPP benchmark gives a much lower level of GDP per capita due to the rapid increase in prices over the recent period, but never falls substantially below subsistence for earlier periods. The PWT approach leads to similar results as the 2011 benchmark for the most recent years, but its estimates for China are closer to the 1990 benchmark results for earlier years.

Figure 2.3. Historical series of GDP per capita for China using alternative benchmark PPPs

Values in 2011 USD



Note: The absolute line for extreme poverty line is based on the current World Bank poverty line of $1.90/day in 2011 USD.
Source: Bolt and van Zanden (2020[9]), "Maddison style estimates of the evolution of the world economy. A new 2020 update", *Maddison Project Working paper 15.*

StatLink https://stat.link/48injw

Looking more broadly into subsistence incomes, the original Maddison Project dataset includes 184 observations of GDP per capita below the subsistence level of USD 400 per year (out of 17 872 observations, i.e. 1% of the total). Most of these relate to countries in times of civil war such as Afghanistan, Liberia, Burundi and the Democratic Republic of Congo. When moving to the 2011 PPPs, the poverty line becomes USD 694, and the number of GDP per capita estimates below subsistence increases to 312 (1.8% of the total). Using the multiple benchmark method developed by PWT, the number of observations below subsistence increases even more, to 386 (2.6%), and this includes countries such as Peru during substantial parts of the 19th century, Egypt and Chile in 1820, and Korea during most of the period. Because of the high quality of the GDP series for some of those countries (the Peruvian historical series, for example, is one of the best available), concerns arise about the quality of their earlier benchmarks, which may explain – in combination with the available time series of GDP per capita – these very low levels of GDP per capita. The main problem of the multiple benchmark method used by the PWT is that the historical GDP series are linked to the earliest available benchmark of the ICP (of 1960 or 1970 for example). This may, in view of the further development of the ICP Project, have been rather crude, therefore implying lower quality benchmarks and implausibly low historical estimates of GDP per capita for some countries.

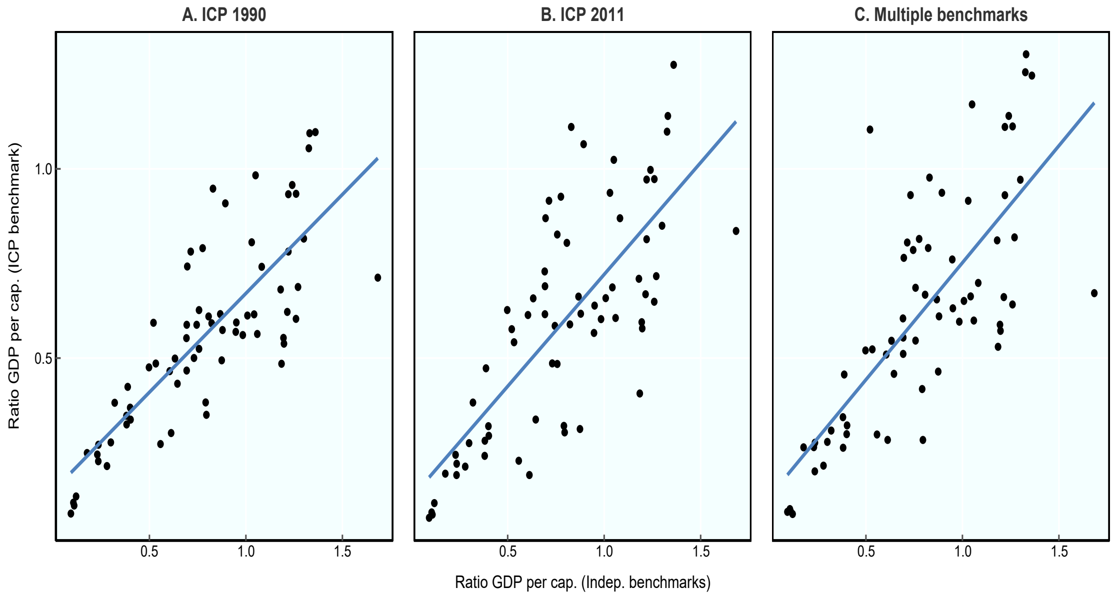
Analysing the accuracy of backward projections through comparisons with historical GDP benchmarks

Even though the effect of the new relative prices on the level of GDP per capita is, for some countries and regions, substantial, this does not necessarily mean that the view of past living standards provided by different methods completely changes, as that view depends largely on relative income levels, i.e. how well off certain countries were compared to other countries in the same year. So another way to assess how plausible the three approaches are is to calculate relative levels of GDP per capita for different countries for the years for which independent GDP benchmark estimates are also available. This comparison also addresses an often-heard criticism of the work by Maddison, i.e. that his dataset keeps relative and absolute levels of real GDP constant at the level of the base year 1990. In reality, prices change over time, and these changes in relative prices lead to biases in measured levels of GDP per capita; the longer the time period considered, and the larger the changes in relative prices, the more problematic an approach based on keeping prices constant would be. This criticism has resulted in various attempts to reconstruct the (implicit) PPPs for historical periods (Prados de la Escosura, 2000[54]) or to use alternative proxies of the economic well-being of nations, such as real wages.

In this section we analyse the backward projections of the three approaches (i.e. making use of the 2011 PPP, the 1990 PPP and the multiple benchmark PPPs, respectively) to see which method results in the best “predictions” of relative income levels in the 19th and early 20th century, but also to find out how large are the biases of the Maddison approach. In other words, do the critics have a point, or do these changes in relative prices cancel out in the long run?

For this analysis, we collected all the independent historical comparisons of GDP per capita levels between sets of countries that are available in the literature. For all of these independent benchmarks, we calculated relative GDP per capita for the same countries and years from the time series using the three different methods. Figure 2.4 shows that the R^2 in panel A (1990 PPP benchmark) is 0.65, whereas the fit for both the 2011 PPP benchmarks (panel B) and multiple benchmarks (panel C) is much lower, with an R^2 of approximately 0.57. It appears that the 1990 benchmark produces better fit historical GDP estimates than the other two approaches.

Figure 2.4. Comparing GDP per capita estimates based on the three PPP approaches with independent benchmarks



Note: Points represent the ratio of GDP per capita of two countries measured using an independent benchmark and one of the ICP benchmarks (1990 (panel A), 2011 (Panel B), or multiple ICP benchmarks (Panel C)).
Source: Bolt and van Zanden (2020[9]), "Maddison style estimates of the evolution of the world economy. A new 2020 update", *Maddison Project Working paper 15*.

StatLink https://stat.link/a6if97

Additionally, we also gathered alternative estimates of relative levels of GDP per capita in the 19th and early 20th century that are based on an *econometric* (indirect) method used by Prados de la Escosura (2000[54]). We compared these indirect estimates with our estimates of GDP per capita based on the three approaches described above, in a similar fashion as we did with the direct benchmarks. The results of these comparisons are very similar – see for details Bolt and van Zanden (2020[9]). The 1990 PPPs and the multiple benchmarks are quite close and compare favorably to the results for the 2011 PPPs.

The conclusion is that simply moving from the 1990 to the 2011 PPPs benchmark does not improve GDP estimates for the historical period, probably due to the greater distance between 2011 and the historical year under review. The PWT approach, which uses all post-1950 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 over a prolonged period of time, which is problematic. The idea of a subsistence minimum is that people cannot survive if their income is below this threshold, so a country can fall below this minimum for one or two years (during a famine, for example) but not permanently or for long periods. An explanation of these results is that the PPP benchmarks are probably subject to two changes: both the quality of the PPPs and the coverage of the various ICP rounds are improving over time, but the distance from the "historical" period is increasing as well, resulting in more biases due to changes in relative prices. It follows that both a "perfect" but "distant" benchmark such as 2011, and the "imperfect" but "early" benchmarks used by the multiple benchmark method, have sizeable biases, and that the 1990 benchmark created by Maddison may well be the best compromise solution. It certainly produces the best results in terms of the two tests carried out in this chapter: first, it predicts historical benchmarks rather well (almost as well as the PWT approach) and, second, it does not result in a high share of countries with an implausibly low (below-subsistence) GDP per capita level. This assessment leads to the conclusion that the best way forward is to stick to the 1990 benchmark for the overall architecture of the Maddison dataset.

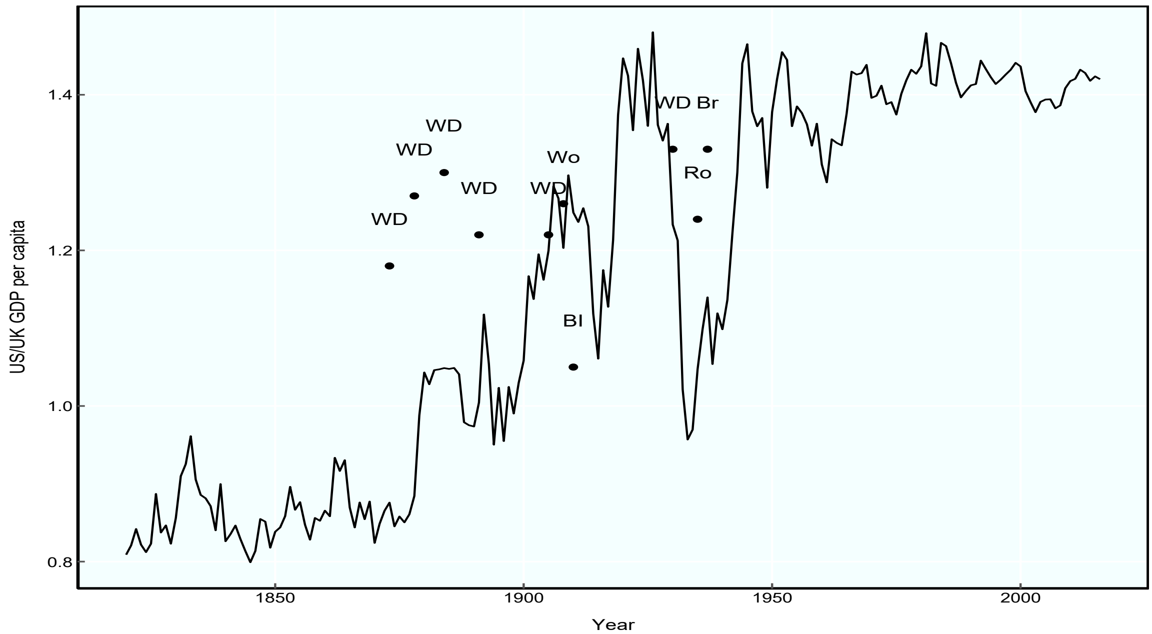
The dataset presented in this chapter, however, departs in a number of respects from the Maddison approach. There is, to begin with, no doubt that the 2011 PPPs and the related estimates of GDP per capita reflect the relative levels of GDP per capita in the world economy better than the combination of the 1990 benchmark and growth rates of GDP according to the national accounts. Hence, we adapted the growth rates of real GDP per capita in the period 1990-2011 to get a close fit between the two (1990 and 2011) benchmarks. Data on GDP growth after 2011 in the current dataset are based exclusively on the growth rates of GDP according to national accounts (but in the future this might be changed when new PPPs become available). A second change, i.e. developing a crude way to incorporate the available historical GDP benchmarks, is described below.

Incorporating important historical benchmarks of GDP per capita

Having collected all available GDP benchmarks (i.e. historical comparisons of GDP per capita between pairs of countries), it makes sense to use the information contained in them. We have therefore used the historical benchmark estimates to fine-tune the dataset for the pre-1940 period. When the quality of the benchmark was considered good, and the difference between the backward projected estimates based on the 1990 benchmarks was larger than 10%, we corrected the time series of GDP growth rates to fit the benchmark. It turns out that, when the Maddison 1990 benchmark is used, most historical benchmarks collected for this chapter do not show differences of more than 10% between the backward projections of the historical series, so the number of modifications is limited.

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 was first criticised by Ward and Devereux (2003[55]), who argued based on alternative measures of PPP-adjusted benchmarks between 1870 and 1930, that the United States led the United Kingdom in term of GDP per capita already in the 1870s, see Figure 2.5. This conclusion was criticised by Broadberry (2003[56]). New evidence however also suggests that the overtaking by the United States indeed already happened in the 19th century (Lindert and Williamson, 2015[57]). This is also confirmed by direct benchmark comparison of the income of both countries for the period 1907-09 (Woltjer, 2015[58]). This shows that GDP per capita for the United States in those years was 26% higher than in the United Kingdom (Woltjer, 2015[58]). We have used this 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, see Figure 2.5. This is close to Prados de la Escosura's conjecture based on his short-cut method (Prados de la Escosura, 2000[54]) and even closer to the Lindert and Williamson (2015[57]) results.

Figure 2.5. GDP per capita in the United States relative to the United Kingdom, 1850-1950



Note: WD: Ward and Devereux (2003); Wo: Woltjer (2015); BI: Broadberry and Irwin (2003); Ro: Rostas (1948); Br: Broadberry (2003).
Source: Bolt and Van Zanden (2020[9]), "Maddison style estimates of the evolution of the world economy. A new 2020 update", *Maddison Project Working paper 15*.

StatLink https://stat.link/dbnk4c

Changing the US/UK ratio on the basis of the new research by Woltjer (2015[58]) 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 the reference point. We have therefore adapted the US series for the period 1908-50 to fit the 1907-09 (Woltjer, 2015[58]) benchmark. The reason for selecting this period is that there are doubts about the accuracy of price changes and deflators for this period, which was characterised by two big 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[38]) for Switzerland). Future research will have to assess whether this choice is justified.

Table 2.3. Main highlights of trends in GDP per capita, regional estimates

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Carribean	East Asia	South and Southeast Asia)	Middle East and North Africa	Sub-Saharan Africa	World
1820	2 307	818	2 513	953	1 089	929	974	800	1 174
1830	2 384	942							
1840	2 580	907							
1850	2 678	985	3 474	1 081	900	929	1 000	800	1 225
1860	3 034	1 358	4 214	1 588					
1870	3 301	1 575	4 647	1 319	989	850	1 165	800	1 498
1880	3 585	1 886	6 019						
1890	4 079	2 204	6 481	1 673		951			
1900	4 724	2 700	7 741	1 751	1 086	994	1 300	850	2 212
1910	5 135	2 283	9 355	2 194		1 143			
1920	4 884	1 343	9 741	2 331	1 160	1 117	900	950	2 241
1930	6 409	2 464	10 297	2 700	1 273	1 299	1 600		
1940	7 185	3 209	11 621	3 024	1 361	1 235	2 146	1 100	3 133
1950	7 263	4 082	14 773	3 713	1 122	1 070	2 393	1 323	3 351
1960	10 974	5 779	17 472	4 751	1 735	1 295	3 110	1 574	4 386
1970	16 161	8 241	23 210	6 286	3 042	1 546	4 801	1 958	5 952
1980	20 950	9 933	28 787	8 728	4 212	1 897	6 742	2 026	7 233
1990	25 440	10 344	35 619	8 132	6 121	2 574	6 435	1 801	8 222
2000	32 536	8 986	44 329	10 225	8 164	3 437	9 640	1 981	9 915
2010	37 318	17 021	48 090	13 453	12 853	5 367	16 716	3 156	13 179
2016	38 511	19 446	51 668	14 090	15 698	6 991	18 010	3 491	14 700

Note: All estimates expressed in 2011 USD.

StatLink https://stat.link/j73ntz

Table 2.4. Main highlights of trends in GDP per capita, 25 major countries

	GBR	NLD	FRA	DEU	ITA	ESP	SWE	RUS	POL	AUS	CAN	USA	ARG	BRA	MEX	KEN	NGR	ZAF	CHN	JPN	IND	IDN	THA
1820	3 306	3 006	1 809	1 572	2 665	1 600	1 415		818	826	1 441	2 674	1 591	867	1 007			1 188	1 070	1 317		827	909
1830	3 550	3 038	1 898		2 657		1 468		942	1 352	1 594	3 039										822	
1840	4 018	3 623	2 276		2 711		1 568		907	2 190	1 852	3 319						1 173				883	
1850	4 332	3 779	2 546	2 276	2 611	1 706	1 715		985	3 148	2 120	3 632	1 994	867	1 054			1 042	858	1 436	947	724	
1860	5 086	3 840	3 016	2 613	2 573	1 930	1 941		1 358	4 613	2 313	4 402	2 160	991	921			1 294				819	
1870	5 829	4 422	2 990	2 931	2 826	1 809	2 144		1 575	5 217	2 702	4 803	2 340	1 084	1 046			1 286	945	1 580	850	810	969
1880	5 997	4 666	3 379	3 174	2 796	2 520	2 359		1 886	6 830	2 895	6 256	2 557	1 058				2 294				1 041	
1890	6 845	5 078	3 787	3 870	2 974	2 463	2 606	1 380	2 204	7 106	3 790	6 665	3 851	1 084	1 568			1 830	964	1 854	931	1 030	1 250
1900	7 594	5 306	4 584	4 758	3 264	2 676	3 320	1 906	2 700	6 397	4 640	8 038	4 583	874	1 822			1 494	972	2 123	955	1 151	
1910	7 718	6 030	4 726	5 337	3 829	2 823	4 053	2 149	2 694	8 305	6 481	9 637	6 092	990	2 240			1 835		2 317	1 111	1 264	
1920	7 017	6 727	5 144	4 457	3 789	3 244	4 788	917	3 000	7 597	6 154	10 153	5 536	1 242	2 552					2 974	1 012	1 409	
1930	8 673	8 931	7 224	6 333	4 631	3 923	6 755	2 308	3 178	7 504	7 669	10 695	6 503	1 350	2 233			2 252	1 012	3 334	1 157	1 704	
1940	10 928	7 701	6 443	8 612	5 099	3 170	7 739	3 417		9 828	8 557	12 005	6 633	1 610	2 560			3 419		4 882	1 093	1 766	
1950	11 061	9 558	8 266	6 186	5 582	3 464	10 742	4 529	3 900	11 815	11 622	15 240	7 949	2 236	3 510	1 038	1 200	4 041	799	3 062	987	1 280	1 302
1960	13 780	13 209	11 792	12 282	9 430	5 037	13 849	6 288	5 125	14 013	13 952	18 057	8 861	3 398	4 723	1 157	1 307	4 847	1 057	6 354	1 200	1 613	1 718
1970	17 162	19 075	18 187	17 277	15 492	9 511	20 269	8 886	7 058	19 166	19 207	23 958	11 639	4 635	6 873	1 458	1 744	6 448	1 398	15 484	1 384	1 882	2 700
1980	20 612	23 438	23 537	22 497	20 959	14 008	23 809	10 245	9 149	22 972	25 784	29 611	13 080	8 249	9 929	1 675	2 080	6 998	1 930	21 404	1 495	2 981	4 071
1990	26 189	27 515	28 129	25 391	26 003	19 215	28 068	10 989	8 150	27 373	30 082	36 982	10 254	7 842	9 699	1 780	1 773	6 111	2 982	29 949	2 087	4 007	7 385
2000	31 946	37 900	33 410	33 367	32 717	26 995	34 203	8 194	12 732	36 603	36 943	45 886	14 369	9 834	12 613	1 915	2 145	7 583	4 730	33 211	2 753	5 384	9 627
2010	34 754	43 812	36 087	41 110	34 766	31 786	42 635	16 345	20 609	45 400	41 209	49 267	18 980	14 216	14 697	2 580	4 905	11 319	9 658	35 011	4 526	8 386	13 344
2016	37 334	45 600	37 124	44 689	33 419	30 110	44 659	18 635	24 838	48 845	43 745	53 015	18 875	13 873	16 133	3 169	5 360	12 1390	12 569	37 465	6 125	10 911	15 454

Note: All estimates expressed in 2011 USD.

StatLink https://stat.link/0sl8io

The world in 1820

The new estimates allow us to draw a more accurate picture of the world economy at about 1820, and also of the process of divergence in economic development before that year, of which we see the results as reflected in disparities in 1820 (). 1820 is the first year for which we can estimate levels of GDP per capita for all major regions in the world (but those for sub-Saharan Africa are still very weak). There were already sizable differences in GDP per capita between regions in 1820: the real GDP per capita in Western Europe and the Western Offshoots was three times that of Eastern Europe or sub-Saharan Africa, and the rest of the world was also quite close to this latter level. This gap, which is even larger when we compare the most advanced countries (the United Kingdom and the Netherlands) with the rest of the world, is (first and foremost) the result of the economic growth that preceded the Industrial Revolution in the North Sea region, where GDP per capita tripled between 1347 and 1820 due to a more or less consistent process of (slow) economic growth. On the other hand, the most advanced parts of the world in 1200 or even 1500 – China, India and the Middle East in particular – showed long-term stagnation and even decline in the centuries leading to the Industrial Revolution. In Asia, only Japan witnessed a process of slow GDP growth similar to that of Western Europe. In Latin America, real GDP increased from the 16th century onwards – mainly due to the dramatic decline of the population, but also thanks to the import of new crops and livestock – but some of these gains disappeared after 1750, when the population began to increase again. This early divergence in per capita GDP between Europe and Asia has been the topic of a fierce debate, the central question being whether the level of economic development in China (and India and Japan) before industrialisation was comparable to Western Europe (Pomeranz, 2002[59]). New research by Bolt and van Zanden (2014[48]), however, clearly indicates that GDP per capita in Western Europe was much higher than in the rest of the world already in 1820, and given the size of the difference, predates the Industrial Revolution.

At the same time, there existed in 1820 substantial differences in levels of GDP per capita within Western Europe itself. The richest country in 1820 was the United Kingdom with around USD 3 300 in 1820; the poorest country was Finland, with around USD 1 300 per capita. North America (with a GDP per capita of almost USD 2 700 in the United States) and the southern cone of Latin America (with GDP per capita of USD 1 500 in Argentina and USD 1 800 in Uruguay) came very close to the Western European average of USD 2 300 (or even surpassed it, as in the case of the United States). GDP per capita for Latin America as a whole, at around USD 650, was much lower. The other main southern hemisphere countries experienced GDP per capita levels that were clearly above subsistence: the average GDP per capita for the Cape Colony in 1820 was about USD 1 500; in Australia this was much lower, at around USD 800.

Developments since 1820

Over the past two centuries the world has witnessed a spectacular growth of GDP and production capacity. On a global scale, GDP per capita has risen 12.5-fold since 1820 (Table 2.3). The average level of global GDP per capita almost doubled in the 19th century (from USD 1 174 in 1820 to USD 2 212 in 1900). GDP growth then accelerated in the 20th century: the next doubling of per capita GDP took about 60 years (1900-60), and a similar doubling between 1960 and 2000 happened in (slightly less than) 40 years.

In the long run, all countries and world regions saw their levels of real GDP per capita rise, as Table 2.3 and Table 2.4 demonstrate. Comparable GDP data on Africa are so scarce that it is difficult to draw any general conclusions (the sub-Saharan Africa estimates for the 19th century presented in Table 2.1 relate only to South Africa and are therefore unrepresentative for the region as a whole). During the 20th century, and in particular after 1950, all parts of the world show real, sometimes accelerating GDP growth. The 1950s and 1960s were the years with the most rapid economic growth worldwide, mainly because many countries managed to catch up (or started to catch up) with the United States. After 1970, and to be more precise after 1973, economic growth often decelerated (see Table 2.5), which partly reflects the end of catching up for Western European countries, where US levels of productivity had been achieved. Declines in material standards of living over longer time periods are exceptional for the post-1950 period; the most extreme case was the result of the dismantling of the centrally planned economies in the USSR and Eastern Europe in the early 1990s, causing GDP per capita in the former USSR to fall by about one-third between 1990 and 2000. But other regions also experienced disappointing developments during those years. Sub-Saharan Africa, for example, had its "lost decades" during the 1980s and 1990s, and the growth performance of Latin America was also quite poor during these first decades of "globalisation".

East Asia, and to a lesser extent, South and Southeast Asia, were the most dynamic parts of the world economy since the 1960s, although even this part of the world knew exceptions; Japan, after catching up in the 1980s, switched to a very slow growth path in the 1990s (but still, in 2018, has a GDP per capita that is 20% higher than in 1990).

As described by Maddison (2003[4]), world leadership in terms of GDP per capita and labour productivity has changed only a few times. The Netherlands was the productivity leader in the 17th and 18th century, until the United Kingdom took over world leadership after about 1780. The United Kingdom, in turn, was surpassed by the United States between 1870 and 1880, as discussed earlier, both in terms of productivity and GDP per capita. During this period, the United States became increasingly more productive compared to the rest of the world, including Europe. As a result, the Transatlantic productivity gap widened substantially between 1900 and 1950, and Western Europe started to catch up only after 1950. A gap in GDP per capita between the United States and Western Europe persisted, however, which is partly due to higher levels of labour force participation and working hours in the United States (see Chapter 3 on working hours).

Up to now, the focus of our attention has been the average GDP per capita of the world population. There is, however, another way of measuring global GDP growth, i.e. by estimating the average growth rate of GDP per capita experienced by the world's population. In the first case (the series we have been discussing so far), countries' GDP growth rates are weighted according to their share in global GDP. In the second series, the growth rates of GDP per capita are weighted according to the share of a country's population in the world total. The latter series, in a way, measures the average rate of GDP growth experienced by the "average" world citizen. When rich countries grow more rapidly than poor countries, the GDP-weighted world series records faster growth simply because rich countries account for a large share of world GDP. Conversely, when poor countries grow more rapidly than rich countries, the population-weighted world series records faster GDP growth, because the majority of the world's population still live in low-income countries.

Economic theory predicts that GDP growth is a cumulative process driven by people and ideas. One would, therefore, expect to observe an almost continuous acceleration of economic growth, as the number of people (and their ideas) grows continuously. This does not appear to be the case, however, when looking at global growth weighted by GDP (the second column in Table 2.5): in this case, GDP growth peaks in the "golden years" between 1950 and 1970, and strongly decelerates after 1970, to recover after 1990. The series weighted by population (the first column in Table 2.5) is much closer to theoretical expectations, with the most recent period also featuring the fastest GDP growth. The two series convey different perspectives on the development of global inequality between countries: initially, the rich countries grow more rapidly than the poor ones, resulting in an increase in income disparities between countries. But gradually the tables turn, and from the middle of the 20th century onwards poor countries start to catch up, with the population-weighted GDP series showing much more rapid growth than the GDP-based series.

Table 2.5. Growth of global GDP per capita, 1820-2010

	Weighted by population	Weighted by GDP	Difference
1820-1850	0.02	0.14	0.12
1850-1870	0.34	1.01	0.67
1870-1900	0.74	1.31	0.57
1900-1920	0.25	0.07	-0.19
1920-1940	1.27	1.69	0.42
1940-1950	-0.06	0.67	0.73
1950-1960	3.00	2.73	-0.27
1960-1970	3.53	3.10	-0.43
1970-1980	2.48	1.97	-0.51
1980-1990	2.10	1.29	-0.81
1990-2000	2.42	1.89	-0.53
2000-2010	4.19	2.89	-1.30

StatLink https://stat.link/fw0cyv

Priorities for future research

There are two main priorities for future research on historical GDP. The first is producing more and better GDP estimates, most notably for China and Africa. In this chapter we have discussed and integrated the important work by Wu (2014[10]) on charting Chinese economic development for the post-1950 period together with a series of recent estimates that document Chinese growth before the 20th century. An important region we know little about is Africa, especially prior to 1950. Nearly all available estimates start in 1950, and even those are sometimes of dubious quality. However, promising work is currently under way to produce the first GDP per capita estimates for Africa in the pre-1950 period (Broadberry and Gardner, 2019[60]).

The second priority for future research is creating more historical benchmarks of relative income levels, as our understanding of comparative income levels becomes sparser as we move back further in time. This is particularly pressing in regions such as Africa and large parts of Asia, but there also important gaps for Latin America in the 19th century. More fine-grained information and more comparative studies are crucial to broadening and deepening such understanding.

References

Álvarez-Nogal, C. and L. De La Escosura (2013), “The rise and fall of Spain (1270–1850)”, *The Economic History Review*, Vol. 1/66, pp. 1-37. [14]

Arroyo, A. and J. van Zanden (2016), “Growth under extractive institutions? Latin American per capita GDP in colonial times”, *Journal of Economic History*, Vol. 76/4, pp. 1182–1215. [16]

Axenciuc, V. (2012), *Produsul intern brut al Romaniei, Vol. 1*, Institutl de Economie Nationala. [34]

Barro, R. and J. Ursua (2008), “Macroeconomic Crises since 1870”, *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution*, Vol. 39/1, pp. 255-350. [19]

Bassino, J. et al. (2018), *“Japan and the Great Divergence, 730-1874,” CEI Working Paper Series 2018-13, Center for Economic Institutions, Institute of Economic Research, Hitotsubashi University*. [40]

Bèrtola, L. (2016), *El PIB per Capita de Uruguay 1870 - 2016: una reconstruccion. PHES working paper No 48*. [27]

Bèrtola, L. and J. Ocampo (2012), *The economic development of Latin America since independence*, Oxford: Oxford University Press. [24]

Bolt, J. and J. van Zanden (2020), *"Maddison style estimates of the evolution of the world economy. A new 2020 update", Maddison Project Working paper 15*. [9]

Bolt, J. and J. van Zanden (2014), "The Maddison Project: collaborative research on historical national accounts", *The Economic History Review*, Vol. 67/3, pp. 627-651. [48]

Broadberrry, 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", *Journal of Economic History*, Vol. 63, pp. 852-863. [56]

Broadberry, S. et al. (2015), *British Economic Growth 1270-1870*, Cambridge: Cambridge University Press. [29]

Broadberry, S., J. Custodis and B. Gupta (2015), "India and the great divergence: an Anglo-Indian comparison of GDP per capita, 1600–1871", *Explorations in Economic History*, Vol. 55, pp. 58-75. [39]

Broadberry, S. and L. Gardner (2019), "Economic Growth in Sub-Saharan Africa, 1885-2008", *Oxford Economic and Social History Working Papers*, Vol. 169. [60]

Broadberry, S., H. Guan and D. Li (2018), "China, Europe and the Great Divergence: a Study in Historical National Accounting, 980-1850", *Journal of Economic History*, Vol. 78/4, pp. 955-1000. [11]

Cha, M. et al. (2020), *Historical Statistics of Korea. Studies in Economic History*, New York: Springer Publishing. [42]

Chen, S. and M. Ravallion (2001), "How Did the World's Poor Fare in the 1990s?", *Review of Income and Wealth*, Vol. 47/3, pp. 283-300. [51]

De Corso, G. (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 History*, doi: 10.1017/S0212610913000190, pp. 321-357. [28]

De Corso, G. and S. Kalmanovitz (2016), "Una estimación del PIB de Panamá: 1906-1945", *Tiempo&Economía*, Vol. 3/1. [25]

Deaton, A. and A. Heston (2010), "Understanding PPPs and PPP-based National Accounts", *American Economic Journal: Macroeconomics*, Vol. 2, pp. 1-35. [52]

Díaz, J., R. Lüders and G. Wagner (2007), *Economía Chilena 1810 - 2000, Producto total y sectorial una nueva mirada, Pontificia Universidad Católica de Chile, Instituto de Economía, Documento de Trabajo no. 315*. [20]

Eloranta, J., M. Voutilainen and I. Nummela (2016), *"Estimating Finnish Economic Growth Before 1860" mimeo*. [30]

Feenstra, R., R. Inklaar and M. Timmer (2015), “The Next Generation of the Penn World Table”, *American Economic Review*, Vol. 105/10. [7]

Feenstra, R. et al. (2013), “Who Shrunk China? Puzzles in the Measurement of Real GDP”, *Economic Journal*, Vol. 123/573, pp. 1100-1129. [53]

Fourie, J. and J. van Zanden (2013), “GDP in the Dutch Cape Colony: the Nationals Accounts of a Slave-Based Society”, *South African Journal of Economics*, Vol. 81/4, pp. 467-490. [47]

Fukao, K. et al. (2015), *Regional Inequality and Industrial Structure in Japan: 1874-2008*, Tokyo: Maruzen Publishing. [41]

Grytten, O. (2015), *Norwegian gross domestic product by industry 1830-1930, Norges Bank Working paper 19/2015. Population from Maddison (2006)*. [32]

Herranz-Loncán, A. and J. Peres-Cajías (2016), “Bolivian GDP per capita since the mid-nineteenth century”, *Cliometrica*, Vol. 10, pp. 99-128. [18]

Krantz, O. (2017), “Swedish GDP 1300-1560. A Tentative Estimate”, *Lund Papers in Economic History: General Issues*, Vol. 152. [36]

Lindert, P. and J. Williamson (2015), “American colonial incomes, 1650-1774”, *The Economic History Review*. [57]

Maddison, A. (2007), *Contours of the World Economy, 1-2030 AD*, Oxford University Press. [8]

Maddison, A. (2006), *The World Economy: Historical Statistics*, Paris: Organisation for Economic Cooperation and Development. [5]

Maddison, A. (2003), *The World Economy: a Millennial Perspective*, Paris: Organisation for Economic Cooperation and Development. [4]

Malanima, P. (2011), “The long decline of a leading economy: GDP in central and northern Italy, 1300-1913”, *European Review of Economic History*, Vol. 15, pp. 169-219. [13]

Malinowski, M. and J. van Zanden (2017), “National income and its distribution in preindustrial Poland in a global perspective”, *Cliometrica*, Vol. 11/3, pp. 375-40. [15]

Nazrin, S. (2016), *Economic history of Malaya*, Oxford University Press, Kuala Lumpur, https://www.ehm.my/data/ (accessed on 1 March 2020). [43]

Palma, N. and J. Reis (2019), “From Convergence to Divergence: Portuguese Economic Growth, 1527-1850”, *The Journal of Economic History*, doi:10.1017/S0022050719000056, pp. 477-506. [33]

Pamuk, S. (2009), *“Estimating GDP per capita for the Ottoman Empire in a European Comparative Framework, 1500-1820”*, paper presented at the XVth World Economic History Congress, August 2009, Utrecht. [44]

Pamuk, S. (2006), “Estimating Economic Growth in the Middle East since 1820”, *The Journal of Economic History*, Vol. 66/3, pp. 809-828. [46]

Penn World Tabels (2011), *Description of PWT 7 0*, https://www.rug.nl/ggdc/docs/pwt-7-web2011a.pdf (accessed on 4 January 2020). [3]

Pomeranz, K. (2002), *The Great Divergence: China, Europe, and the Making of the Modern World Economy*, Princeton: Princeton University Press. [59]

Prados de la Escosura, L. (2017), *Spanish Economic Growth, 1850-2015*, London: Palgrave Macmillan. [35]

Prados de la Escosura, L. (2009), “Lost decades? Economic performance in post-independence Latin America”, *Journal of Latin America Studies*, Vol. 41, pp. 279-307. [23]

Prados de la Escosura, L. (2000), “International Comparisons of Real Product, 1820-1990”, *Explorations in Economic History,*, Vol. 37, pp. 1-41. [54]

Ravallion, M., G. Datt and D. van de Walle (1991), “Quantifying absolute poverty in the developing world”, *Review of Income and Wealth*, Vol. 37/4. [50]

Ridolfi, L. (2016), *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)”*, Dissertation IMT School for Advanced Studies, Lucca, http://e-theses.imtlucca.it/211/1/Ridolfi_phdthesis.pdf (accessed on 4 January 2020). [17]

Santamaria Garcia, A. (2005), *Las Cuentas nacionales de Cuba, 1960-2005, mimeo*. [22]

Schön, L. and O. Krantz (2015), “New Swedish Historical National Accounts since the 16th Century in Constant and Current Prices”, *Lund Papers in Economic History*, Vol. 140. [37]

Seminario, B. (2015), *El Desarrallo de la Economía Peruana en la Era Moderna*, Universidad de Pacifico, Lima. [26]

Stohr, C. (2016), “Trading Gains: new estimates of Swiss GDP, 1851-2008”, *Economic History Working Papers/London School of Economics and Political Science, Economic History Department, London, UK*, Vol. 245/2016. [38]

Sugimoto, I. (2011), *Economic growth of Singapore in the Twentieth century: Historical GDP estimates and empirical investigations*, Economic Growth Centre Research Monograph ser., 2, https://doi.org/10.1142/7858 (accessed on 30 January 2013). [45]

Summers, R. and A. Heston (1991), “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988”, *The Quarterly Journal of Economics*, Vol. 106, pp. 327-368. [6]

van Zanden, J. et al. (eds.) (2014), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-en. [1]

van Zanden, J. and B. van Leeuwen (2012), “Persistent but not consistent: the growth of national income in Holland 1347-1807”, *Explorations in Economic History*, Vol. 49/2012, pp. 119-30. [31]

Vogel, F. and P. Rao (eds.) (2011), *Reshaping the world: The 2005 round of the International Comparison Program, forthcoming*, Washington, DC: World Bank. [2]

Ward, M. and J. Devereux (2012), “The Road Not Taken: Pre-Revolutionary Cuban Living Standards in Comparative Perspective”, *The Journal of Economic History*, Vol. 72/1, pp. 104-132. [21]

Ward, M. and J. Devereux (2003), “Measuring British Decline: Direct Versus Long-Span Income Measures”, *The Journal of Economic History*, Vol. 63, pp. 826-851. [55]

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*, pp. 1-22. [58]

World Bank (2014), *Purchasing Power Parities and Real Expenditures of World Economies Summary of Results and Findings of the 2011 International Comparison Program World Bank: Washington DC*. [49]

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 Economics Program Working Paper Series EWP#14-01. [10]

Xu, Y. et al. (2016), “Chinese National Income, ca. 1661-1933”, *Australian Economic History Review*, Vol. 57/3, pp. 368-393. [12]

Notes

[1] “Domestic” refers to the fact that the estimates relate to what is happening within the borders of the country involved, whereas “national” would refer to the income or expenditure of the citizens of those countries.

[2] See the discussion of historical sources below for a more detailed explanation.

[3] PPP estimates between benchmarks are interpolated; when no earlier or later ICP benchmark is available, the extrapolation method is used.

[4] http://www.rug.nl/research/ggdc/data/penn-world-table.

[5] http://icp.worldbank.org/.

3 The working week in manufacturing since 1820

Oisín Gilmore, University of Groningen

Despite the substantial implications of working time for considerations of living standards and economic output, historical data on working hours are sparse. This chapter presents a new dataset on the length of the working week in manufacturing globally from 1820-2010[1]. The dataset contains some 4 300 observations and covers 120 countries or political units. It shows that workers in manufacturing worked 60 to 90 hours per week in the 19th century, compared to around 40 hours today. This is a reduction of 20-50 hours, that is, 50-125% of today's average working week. The data also show that weekly working hours declined rapidly after World War I, with the introduction of the eight-hour day, and again later in the 20th century, with the generalisation of the five-day week. This decline in weekly working hours stalled since the 1950s-60, and seem to have reversed in the most recent period.

Introduction

Working time and leisure time should be central concerns of any attempt to evaluate people's well-being. As the OECD notes in the flagship *How's Life?* publication, working hours and time off work are central to work-life balance. When considered globally, we continue to see substantial disparities in working time, with workers in some countries working significantly more than in others. This is equally true over time.

Amongst OECD countries the amount of time people spend working in a year has declined dramatically since the 19th century, and, with that, the amount of free time people have away from work has also substantially increased. This reduction in working time has two major components: first, more time spent on holidays and days off work, and second, a dramatic reduction in the length of the average working week (Huberman and Minns, 2007[1]).

This chapter presents a new dataset on the length of the working week in manufacturing globally from 1820-2010. The dataset contains some 4 300 observations and covers 120 countries or political units.[2] This sectoral focus on hours per week in manufacturing allows us to ensure that the data are broadly comparable between countries and over time.

This chapter shows that workers in manufacturing worked 60 to 90 hours per week in the 19th century, as compared to roughly 40 hours today. That is a reduction of 20-50 hours, i.e. 50-125% of today's average working week. The welfare implications of a reduction in working time of this scale are of course substantial. Both the reduction in the pain and toil of work and the increase in leisure time have substantial positive welfare implications (de Jong, 2015[2]). However, working time is often ignored in historical studies due to the inadequate quality of existing international datasets. This chapter advances our understanding of these issues by presenting a new international dataset on working hours in manufacturing since 1820.

The content of this chapter is as follows. The chapter starts by first describing the concept of "working time" and considering some of its historical and contemporary ambiguities. It then presents the existing datasets on working time in the long run and the historical sources used to build the new dataset presented in this chapter, and it assesses the quality of the data. The subsequent sections describe some of the main highlights of trends in working time between 1820 and 2010, and then some extensions of the measure of working time from a weekly measure to an annual measure for a restricted number of countries. The final two sections consider how working time correlates with real GDP per capita and what the priorities are for future research on this topic.

Description of the concepts used

Working time might appear a rather simple and uncomplicated concept. It would seem to be just the hours that you work. Unfortunately, it is a rather more complex concept with many ambiguities.

The first issue is what is "work" and what separates it from "non-work". To a worker in a contemporary developed economy, this might seem like a straightforward question, but if we think historically and consider the different approaches to "work" over history, the complexity becomes clear. Likewise with "time": while this may seem a simple uncontroversial concept, if we look at different attitudes to working time over the historical long run, it turns out that the idea of clearly delineated working time is a very modern concept, particular to our historical moment.

Marshall Sahlins' acclaimed anthropological work, *Stone-Age Economics* (Sahlins, 1972[3]) on hunter-gatherer societies, describes how concepts of both work and time were fundamentally different in the !Kung hunter-gatherer society of southern Africa. First, he noted that the time the !Kung spent at work was dramatically less than the modern Western worker, but he also noted that the concept of work was something that researchers were statistically imposing on the activities of the hunter-gatherers. The clear distinction between work and non-work was not present in the !Kung people's self-description of their activities (Sahlins, 1972[3]). This observation has been repeatedly confirmed in other studies of hunter-gatherer people. Again and again the average "working time" of hunter-gatherers, if it can be called that, has been found to be far lower than in advanced Western societies (Clark, 2005[4]).

Looking at the Middle Ages, Jacques Le Goff, in his classic essay on working time in medieval Western Europe (Le Goff, 1982[5]), argues that before the 14th century the approach to working time was fundamentally different to ours. He writes that "on the whole, labour time was still the time of an economy dominated by agrarian rhythms, free of haste, careless of exactitude, unconcerned by productivity – and of a society created in the image of that economy, sober and modest, without enormous appetites, undemanding, and incapable of quantitative efforts" (Le Goff, 1982[5]).

Clearly, across time and between different regimes of labour, conceptions both of work and of working time can be radically different, to the point that neither concept properly applies. Table 3.1 lays out an incomplete, schematic description of how regimes of working time have differed over the long run.

Table 3.1. Different patterns of time use in hunter-gather, agrarian, proto-industrial, industrialising and industrialised countries

	Hunter-gatherer	Agrarian	Proto-industrial	Industrialising	Industrialised
Working time	3-8 hours per day	Varies; ~250 days per year	10-12 hours per day, 150-200 days per year	10-12 hours per day, ~300 days per year (60-80 hour week)	8 hours per day, 40-48 hour week, paid holidays
Patterns and Conceptions of time use	Casual approach to work, little distinction between work/non-work, long breaks from work	Varies widely, work full days, dependent on season	Full day, many holidays/holy days/days without work	Full day, lack of holidays	Restricted day, clear idea of "free time": leisure time and holidays
Sources	Sahlins (1972[3]); Clark (2005[4])	Allen and Weisdorf (2011[6]); Isett and Miller (2016[7])	Voth (1998[8]); Allen and Weisdorf (2011[6]); Humphries and Weisdorf (2017[9]); Stephenson (2018[10])	Voth (1998[8]); Allen and Weisdorf (2011[6])); Humphries and Weisdorf (2017[9]); Stephenson (2018[10])	See text

When we try to develop estimates of working time historically, we find radical differences in the organisation of time in relation to work. In agrarian societies, work is highly seasonal and dependent on land quality and on the type of farming being done.[3] Several studies, primarily looking at the United Kingdom, have found that prior to the Industrial Revolution urban workers appear to have worked full days (i.e. 10-12 hours) but far fewer days a year than subsequently. Prior to the Industrial Revolution, workers worked in the region of 150-200 days per year. During the Industrial Revolution this increased to over 300 days per year.[4] This contrasts with industrialised economies in the 20th century when restrictions on hours of work with the introduction of the eight-hour day and the emergence of substantial paid holidays emerges as the norm. A clear distinction between working time and free time emerged in that period.

The fact that the work time/free time distinction is a highly modern one is not the only conceptual problem related to working time. Even in developed economies where this distinction is widely understood, the concept of work time contains other ambiguities. One obvious problem is that much work is unpaid. For example, when calculating the work time of a mother, what consideration should we give to unpaid domestic work? Another problem is whether all time devoted to work should be considered work time. Time spent commuting to and from work is not generally considered as work time, but it is, effectively, time devoted to working. But what about time moving to and from work sites while on the job? For example, in the case of coal miners, should the time spent travelling from the opening of a mineshaft to the coal face be considered work? Or what about meal times? Whilst a two-hour break for a siesta might not be considered part of work time, what about a 30-minute break for a coffee and sandwich? The assessment of these issues is ultimately contingent on social convention (Gershuny and Sullivan, 2019[11]).

These issues have been of concern to labour statisticians for a long time. The first International Conference of Labour Statisticians (ICLS), convened by the International Labour Organization (ILO) in 1923, agreed on a set of guidelines for the collection of statistics on working time (1923, 1924). These guidelines were then revised in 1962, at the tenth ICLS, and yet again more recently at the 18th ICLS in 2008. Under the latest guideline, working time is no longer confined to hours in paid employment and includes self-employment. And a distinction is made between two broad categories of working time data (International Labour Organization, 2008[12]).

The first establishes the preferred form of data: hours actually worked. This includes three elements: hours spent on an activity directly related to work (time spent working, moving between work locations, in-work training, etc.); in-between time (the inevitable interruptions to work processes that occur); and resting time (coffee breaks etc.). It explicitly excludes commuting time, long breaks such as meal breaks, annual leave, holidays, time spent in education or in training that is not part of the job and other reasons for not working (maternity and paternity leave, parental leave, slack in business, bad weather, etc.).

The second category of working time data is slightly broader and includes measures of working time calculated in the employer/employee relationship. These include in particular the concepts of normal hours, (usual) working hours and hours paid. Normal hours are the hours above which overtime would be paid, and are often set down in legislation, collective agreements, etc. Usual hours include in addition regular overtime. Hours paid simply refers to the hours for which a worker is paid. Note that the primary difference between these is overtime that is not paid. And as the 1962 guidelines note, "because of the wide difference among countries with respect to wage payments for holidays and other periods when no work is performed, it does not seem feasible at this time to adopt international definitions of hours paid for" (International Labour Organization, 1962[13]).

Historical sources

This chapter presents a newly collected global dataset on working time in manufacturing. It improves the three existing international datasets on working time (see below) in a number of ways. First, it covers a longer time period. Second, it has a far greater number of observations. Third, unlike all three previous datasets on working time, it does not rely on a large number of interpolated estimates. Finally, unlike previous datasets, it focuses exclusively on manufacturing: as explained below, it does so to ensure that the data are of the highest possible quality and broadly comparable over time and between countries.

The first of the three previously developed international datasets on working time is the dataset developed by Angus Maddison and presented in slightly different forms in a series of publications (Maddison, 1964[14]; 1995[15]; 1982[16]; 2001[17]). In these publications, Maddison takes decadal observations for a small number of countries from the ILO *Yearbooks of Labour Statistics*, and from a single observation for the average working week in the United Kingdom in 1860 he constructs very rough estimates for the years between 1860 and 1938. The second dataset is the Total Economy Database, produced by the Conference Board (2019[18]), which provides estimates of annual (rather than weekly) hours worked for 64 countries from 1950 onwards. This dataset draws on data from the OECD, Eurostat, the Asian Productivity Organization and a small number of research papers (Crafts, 1997[19]; Hoffman, 1998[20]). A substantial amount of the data in the Total Economy dataset, especially for earlier years, is based on interpolations and extrapolations. The third historical dataset is the one compiled by Huberman and Minns (2007[1]), perhaps the best of these previously available datasets, although it covers only 15 countries. From 1929 onwards, it takes most of its data from the prior two datasets, adding to them a new set of data for 1870-1900 collected by Huberman (2004[21]; 2012[22]), from the Fifteenth Annual Report of the US Commissioner of Labor (U.S. Department of Labor, 1990[23]).

The data presented in this chapter represent a significant improvement on previously available long-term series on working time. The dataset covers a much longer time frame and involves no interpolations or extrapolations except those in the source datasets. However, unlike the above three datasets, the data in this chapter focus exclusively on manufacturing, rather than attempting to estimate the average working week across the entire economy. By focusing exclusively on manufacturing, we can ensure that the data are broadly comparable over time and between countries.

While it would be preferable to have data on average working time for the entire economy, there are significant difficulties in estimating this over the long run. For much of the economy, little data on working time was recorded. Despite agriculture accounting for the vast majority of working time expended in most economies over history, we have reliable data on working time in agriculture for only a handful of countries. There is also very little data on working time in the service industry. While there is reasonably good data on working time in construction, mining and transport, country coverage for these series varies greatly.

Some estimates of the average working time outside agriculture do exist. However, the industries included in these estimates often differ by country, especially for earlier years.

Ideally, we would have reliable data on average working time across the entire economy, or at least data on the average working week outside agriculture. But neither of these datasets is achievable given the historical data available. Therefore, a choice needs to be made. On the one hand, it would be possible to present data that cover the greatest possible proportion of the workforce. However, because different parts of the workforce would be included at different times and in different countries, this data would not be comparable across time and between countries. On the other hand, we could choose a specific section of the workforce to focus on. This would allow for the data to be comparable across time and between countries, but it would be less comprehensive. In this chapter I choose the latter strategy by focussing on working time in manufacturing. For nearly all countries, the best time series on working time refer to manufacturing. Focussing on manufacturing thus ensures that the data is of the highest possible quality.

The new dataset presented in this chapter has been collected from three historical sources. The first and most important source is the series of *ILO Yearbooks of Labour Statistics*. A version of the *Yearbook* has been published since 1934 covering countries from 1927 onwards. All data on working time from each annual publication for 1934-70 has been extracted. Unlike Maddison, who took decadal observations for a small number of countries from the ILO *Yearbooks*, I take all the annual data available from every single *Yearbook* from 1934 to 1970. From 1970 onwards, I drew on the data that is available from the ILO website. However, this data oddly does not contain all the data of the ILO *Yearbooks*. I therefore completed these data by adding data from ILO *Yearbooks* for 1980, 1990, 1995, 2000 and 2009, covering all years from 1970-2008.

The second historical source is country-level studies. I collected these for 16 countries: Australia, Belgium, Canada, China, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Spain, Sweden, Switzerland, the United Kingdom and the United States. These studies are primarily in two forms. First, there are government publications from the respective country. Second, there are academic studies on the history of working time in the respective country. These data sources are fully described in Gilmore (forthcoming[24]). For this chapter, this type of data is used only for the period 1820-1939. After that, all the data come from the ILO.

The third source is the data from the Fifteenth Annual Report of the US Commissioner of Labor collected by Michael Huberman (2004[21]; 2012[22]), covering the years 1870-1900. It should be noted that this presents estimates for the entire economy, not exclusively manufacturing. However, these estimates are dependent largely on data on working time in manufacturing. Further, for the countries where Huberman presents data for both manufacturing and the entire economy, the figures are broadly similar to each other (Huberman, 2004, pp. 970-971[21]).

The data from the various historical sources were combined in the following manner. From the ILO *Yearbooks*, if available, I used the data series on the average working week for total manufacturing. If this was not available, I constructed an unweighted average of working time in various manufacturing industries. If no data from the ILO was available and the year was 1939 or earlier, I used data from national sources. Once again, if data for the average working week for total manufacturing were available, I used those. If not, I constructed an unweighted average of working time in various manufacturing industries. If none of this data was available and the years were between 1870 and 1900, I used data from Huberman.

Data quality

Generally, the data used in this chapter are of good quality and refer to actual hours worked by workers in manufacturing, although some data refer to hours paid for or normal hours worked. Almost all data come from establishment surveys or household and labour force surveys where employers or workers are asked how many hours were worked over the previous week or month.

The most significant problems that might affect the consistency of the data are in relation to, first, whether or not paid non-working time (e.g. breaks and holidays) is included; and second, whether or not unpaid working time (e.g. unpaid overtime) is included. As noted above, measurement of actual working time according to ICLS guidelines measures only hours of actual work. It does not include breaks, holidays and other forms of paid non-working time, and it does include unpaid working time such as unpaid overtime. For later years, most data were collected in line with the ICLS guidelines and are therefore highly comparable. For earlier years, most data are also on hours of actual work, but as the ICLS guidelines did not always exist, actual working time might be measured differently in different countries. However, as workers generally received fewer paid holidays in the past, estimates of actual working time are likely to include overtime and unlikely to be biased by the inclusion of substantial paid leave.

Table 3.2 describes the data quality using the four levels of quality described in *How Was Life?* (van Zanden et al., 2014[25]).[5] From the 1930s onwards, nearly all data come from the ILO, i.e. “an official statistical agency …[that]… uses techniques that ensure equivalent credibility” (Ibid.). Prior to that year, the data come from a range of national level sources. But again, nearly all are measures of actual hours of work and are based on survey data. Little data are available prior to the 1870s, and the data are of poorer quality.

Table 3.2. Quality of data on working time by region and benchmark year, 1820-2010

	East Asia	Eastern Europe	Latin America and Caribbean	Middle East and North Africa	South and Southeast Asia	Sub-Saharan Africa	Western Europe	Western Offshoots
1820							3	3
1870							2	2
1913	3						2	2
1930	1	1	1	1	1		1	1
1950	1	1	1	1	1	1	1	1
1970	1	1	1	1	1	1	1	1
2000	1	1	1	1	1	1	1	1

Note: 1. High quality: the product of official statistical agency (national or international); 2. Medium quality: the product of economic-historical research using the same sources and methods as applied by official statistical agencies; 3. Moderate quality: economic historical research, but making use of indirect data and estimates; and 4. Low quality: estimates based on a range of proxy information. In case of multiple sources, the lowest quality source is given.

Some issues affecting data quality should be flagged. First, although almost all of the data were reported in the original sources in the form of hours per week, this was not always the case. A small amount of data was given in other forms and had to be adjusted (e.g. data reported in the form of hours worked per fortnight were divided by two; those given in the form of hours per month were multiplied by 12/52; and data given in the form of hours per day were converted to hours per week by assuming that workers worked 5, 5 ½, or 6 days a week – depending on the time period considered). This application of these assumptions should not be particularly controversial. Frequently, the normal number of days in the working week could be easily calculated from the available data. Occasionally, it was assumed that the number of days in the working week was similar to neighbouring countries.[6]

A second larger problem of data quality is the problem of representativeness. In one case, this problem is very stark. Most of the data on working time in sub-Saharan Africa prior to the 1980s refer to the average working week of white male South Africans. For the entire dataset a major problem is how representative this dataset is of women workers. In most cases the original sources did not state whether the data are for male or female workers. But where data are not given for both sexes, unless a figure for all workers is given, I have used the data for male workers. I made this decision because I often do not have reliable consistent data on the gender composition of the workforce in manufacturing, so I am unable to generate an estimate of the average working-time weighted by sex. Fortunately, from available data we know that during this entire period most workers in manufacturing were male, and therefore the average working time for male workers should be broadly representative of the average worker (Mitchell, 2013[26]).

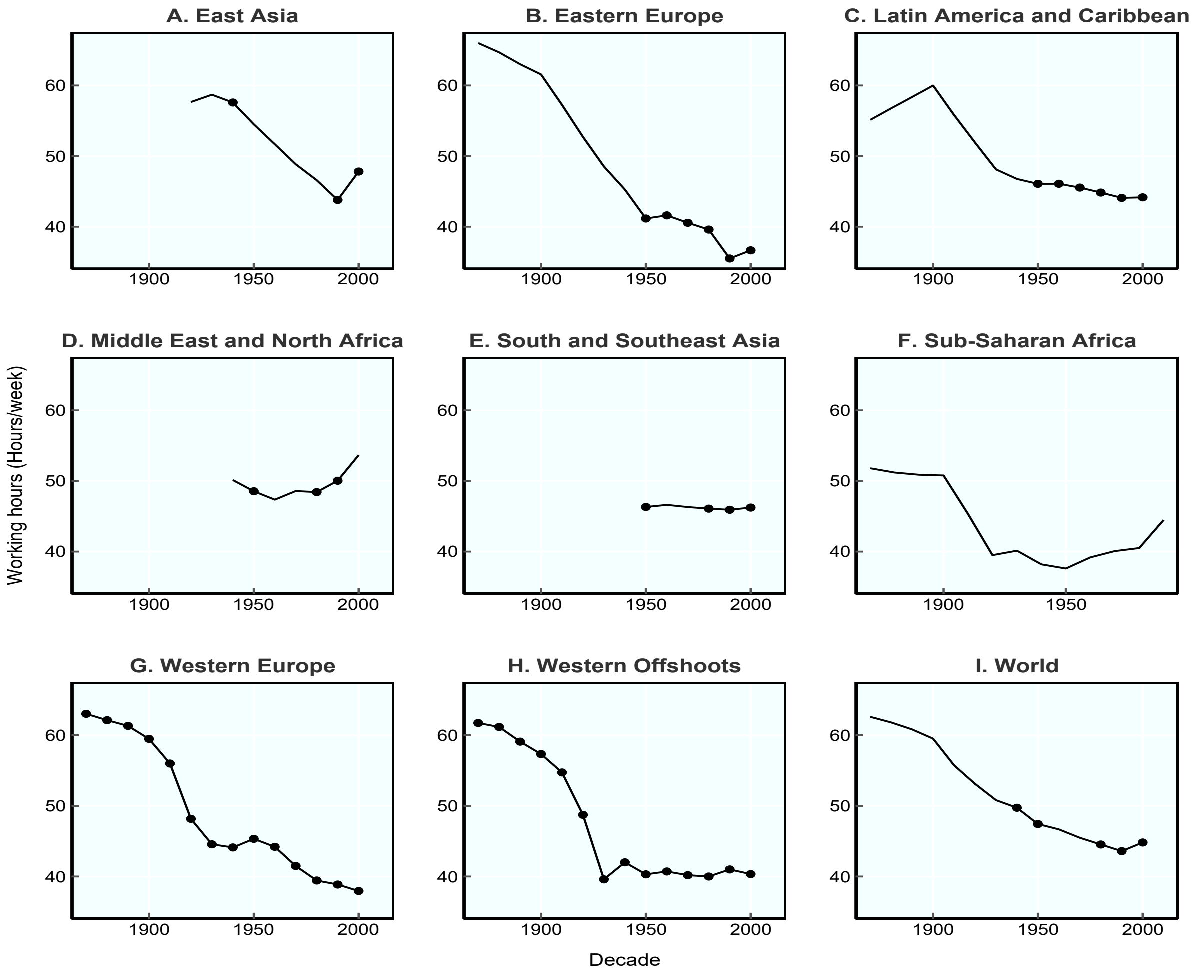
For series early in the dataset, there are further problems of representativeness. First, the data often refer to a few manufacturing industries, which may not be representative of all manufacturing workers. Second, especially in the 19th century, there was substantial seasonal variation in hours of work, which are not fully accounted for in this dataset due to data constraints.

Despite these problems of representativeness, the data presented in the chapter are deemed to be of high quality for historical data and provide significant new insights into the history of working time.

Main highlights of historical trends in working hours

Figure 3.1 presented below shows a dramatic decline in the average working week in manufacturing across the world. In the late 19th century, working hours were in the region of 60 per week while, by the start of World War I, working hours were around 55 hours per week. After World War I there was a dramatic decline in working hours in the West, with the introduction of the eight-hour day and the forty-eight-hour week in both Western Europe and in the Western Offshoots (including the United States, Canada, Australia and New Zealand). By the start of the second half of the century, working hours in manufacturing were generally somewhere between 40 and 48 hours a week. Since then, the decline in working hours has to a substantial degree stalled. In the most recent period, there is some evidence that they are beginning to increase again. Figure 3.1 also shows the decline in working hours outside the West. This generally followed a similar trend as in the West but condensed into a shorter time period.

Figure 3.1. Average working hours in manufacturing across world regions, 1870-2010



Note: Dots indicate regional averages with population coverage above 40%.

StatLink https://stat.link/7a15ce

Table 3.3 below allows us to look more closely at these changes in working time by presenting developments at a national level since 1820. In both Western Europe and the Western Offshoots, the early 18th century was characterised by an increase in working time from high levels to even higher levels, followed by a decline. This suggests a familiar Kuznets (inverted U) shape to the development of working hours relative to GDP (Spoerer and Streb, 2008[27]), i.e. one where the initial stages of industrialisation bring a number of negative side effects but, as economic development proceeds, those negative effects diminish. As with most other types of Kuznets curve, the relationship appears to break down in the 1980s and 1990s, with the decline in hours plateauing in the West and increasing in medium-income countries like Turkey, Egypt and China.

Table 3.3. Average working week for manufacturing workers in selected countries, 1800-2010

	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1800-10				72.0																					
1810-20	63.0																								
1820-30				77.5																					
1830-40	63.0			90.0								65.4													
1840-50	62.0			90.0								68.7													
1850-60	62.6			90.0								67.8													
1860-70	60.9			81.0								65.1													
1870-80	57.9	64.6	66.0	61.0	63.3	64.2	69.0		68.8	55.4	57.7	62.4	60.6							59.4					
1880-90	54.5	63.0	66.0	62.3	63.5	62.6	63.6			52.6	59.2	61.8													
1890-00	54.9	61.5	65.7	61.1	63.7	60.1	60.4			49.8	61.4	59.5													
1900-10	55.0	61.4	62.2	57.6	63.8	58.4	59.5		64.5	48.9	58.3	57.8	69.0							58.5	60.0				
1910-20	51.7	60.0	58.1	54.9	60.7	56.4	54.3			48.7	55.9	55.0													
1920-30	47.1	48.2	48.0	49.7	46.1	48.5	45.1	46.2		46.3	50.8	48.7								46.8		54.7			
1930-40	47.3	47.6	41.5	45.2	39.5	47.5	45.9	43.6		45.3	48.1	38.5	45.2							47.2		55.8			
1940-50	46.7	47.9	43.0	41.9	43.9		47.1			43.0	47.1	41.5	45.6		46.3	51.6				45.0	57.8	54.7			
1950-60	45.9	48.8	44.8	47.4	44.3		40.3		39.9	40.0	41.1	40.3		45.3		50.4	45.0			44.4		49.4	46.2		
1960-70	46.2	46.2	45.8	44.0	43.7	44.2	37.4	42.6	40.5	43.5	40.6	40.6	45.7	46.7		49.5				45.8		45.4			47.8
1970-80	43.8	42.2	42.6	42.1	38.5	43.4	38.8	38.3	40.6	42.4	39.2	40.2	45.8	45.6	44.2	56.1				46.7		41.2			
1980-90	42.8	40.3	39.2	40.6	38.7	36.9	38.0	35.0	40.5	37.1	38.5	40.4	46.3	44.8	41.3	58.3	39.5	42.3		47.1		41.3	46.2		47.8
1990-00	41.7	39.0	38.6	38.0	40.5	36.7	37.0	39.4	37.7	38.5	38.5	41.4	45.4	42.8	45.1	56.7	45.3	43.1		45.4	44.1	38.1	46.5	43.1	48.5
2000-10	41.0	38.2	36.1	37.9	37.8	36.1	37.4	41.2	36.4	38.0	38.0	40.8	45.2	41.7	45.0	55.8	52.4				48.8	38.5	47.0	43.4	

StatLink https://stat.link/ah9wuc

The early stage of this curve has been well studied by historical research. The increase in working time during industrialisation has been a major issue in both social history (Thompson, 1967[28]; Reid, 1976[29]) and economic history (Bienefeld, 1972[30]; Clark and Van der Werf, 1998[31]; Voth, 1998[8]; Allen and Weisdorf, 2011[6]; Humphries and Weisdorf, 2017[9]). It should be noted that the very high working times observed during this period often involved a different rhythm of work to what might be considered as normal today. The early 19th century in the United Kingdom saw the transition from the "workshop system" – where workers had substantial autonomy around their work time, being free to enter and leave work at their leisure, but often spending their entire working day with their family at their workplace – to the "factory system" in the early 19th century, with workers subject to "factory discipline" and having little control over work time, set times for meals, and fixed hours for starting and ending work.[7]

The high levels of work time led to workers' demands for their reduction relatively early in the Industrial Revolution. A major concern at the time was the impact of long working hours on children and family life. As early as 1802, legislation was introduced in the United Kingdom restricting the working time of children serving as "parish apprenticeships"[8] to 12 hours a day excluding breaks. In 1819, working hours for all children under 16 were restricted to 12 hours in cotton mills. In 1833, this was reduced to eight hours for children under 13, with a requirement that children attend two hours of schooling a day. In the 1840s, similar restrictions were placed on the working time of women. The result of these legislative changes was that by the end of the 1840s women and anyone under 18 could work no more than 12 hours per day. The introduction of restrictions on the working time of men however progressed much more slowly.

In other industrialising countries such as France, Germany, the United States and the Benelux, the process of industrialisation happened later and faster, as did the introduction of restrictions on working time, with many basic reforms coming in the 1890s and early 20th century (Huberman, 2012[22]).

By the start of World War I, manufacturing workers in most Western countries were working around 55-60 hours per week. However, immediately after the war, legislation introduced the 8-hour day across most of the Western world, leading to a sudden reduction of working time to around 48 hours per week. The 48-hour week persisted in most countries until World War II, with a few exceptions.[9] By the early 1970s, most developed economies had shifted from the 6-day week to the 5-day week, resulting in a 40-hour week. Simultaneously, there was an increase in paid holidays (a pattern described in the section below on Holidays and annual working time). Since the early 1970s, the decline in hours has somewhat stalled, with full-time manufacturing workers generally continuing to work roughly around a 40-hour week, although the average working week for manufacturing workers for some countries in Western Europe declined by 1-2 hours. However, the growing numbers of workers working part-time or flexi-time makes the analysis of this decline hard to interpret. Uniquely in France, there has been a greater reduction in hours with the shift to a 35-hour week.

Outside of Western Europe and the Western Offshoots, the shift to a 48- or 45-hour week in the post-war period has also stalled, albeit at a higher level. The transition to a 40-hour week has been achieved in relatively few developing countries, and where it was achieved it did not persist for long. Indeed, in some developing countries not only has the reduction in hours stalled but over the last 20-30 years the average working week in manufacturing has increased in length. Conversely, other countries such as Ukraine and Moldova saw substantial reductions in hours of work in the late 1990s. This decline was presumably driven by the economic distress and poor economic performance experienced in those countries. The low value of 36 hours per week as the average working week for Eastern Europe and the former USSR for the 1990s in Table 3.4 below is partially explained by these developments.

Generally, the population-weighted regional averages shown in Table 3.4 tell the same story as above. For Western Europe, Eastern Europe and the former USSR and the Western Offshoots: a gradual decline until the introduction of the eight-hour day after World War I, the shift to a five-day week over the subsequent 40 years, followed by a lengthy period since the 1970s with little if any reduction in the average working week. In Latin America and the Caribbean, the average week seems to have stabilised earlier and at a slightly higher average level. In the rest of the developing world, average working hours in manufacturing appear to have persistently remained at higher levels than in the West.

Table 3.4. Average working week in manufacturing by region and decade, 1820-2010

Hours per week

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	South and Southeast Asia	Sub-Saharan Africa	Middle East and North Africa	East Asia	World
1820s	[65]	..	..	..	..	..	..	..	[65]
1830s	[77]	..	65	..	..	..	..	..	[76]
1840s	77	..	68	..	..	..	..	..	[76]
1850s	[77]	..	67	..	..	..	..	..	[76]
1860s	[73]	..	64	..	..	..	..	..	[71]
1870s	63	[66]	62	[55]	..	[52]	..	..	[63]
1880s	62	[65]	61	[57]	..	[51]	..	..	[62]
1890s	61	[63]	59	[58]	..	[51]	..	..	[61]
1900s	59	[62]	57	[60]	..	[51]	..	..	[60]
1910s	56	[57]	55	[56]	..	[45]	..	..	[56]
1920s	48	[53]	49	[52]	..	[39]	..	[58]	[53]
1930s	45	[49]	40	[48]	..	[40]	..	[59]	[51]
1940s	44	[45]	42	[47]	..	[38]	[50]	58	50
1950s	45	41	40	46	46	[38]	49	[54]	47
1960s	44	42	41	46	[47]	[39]	[47]	[52]	[47]
1970s	41	41	40	46	[46]	[40]	[49]	[49]	[45]
1980s	39	40	40	45	46	[41]	48	[47]	45
1990s	39	36	41	44	46	[44]	50	44	44
2000s	38	37	40	44	46	[34]	[54]	48	45

Note: Square brackets indicate estimates based on a population coverage below 40%.

StatLink https://stat.link/36hz4u

Of course, caution should be exercised when comparing these regional averages. As described above, the influence of Ukraine and Moldova in skewing the regional value for "Eastern Europe" should caution against overinterpreting these averages when they rest on a small number of observations. Special caution should also be paid with the estimates for sub-Saharan Africa, the Middle East and North Africa, and East Asia. As Table 3.5 shows, the number of observations for these regions is rather low. Data coverage for sub-Saharan Africa is especially poor. The average for the East Asia region includes only four countries, although one of these countries is China; unfortunately, data coverage for China is very poor with almost no observations before 1990.

However, while the regional data might be based on a small number of observations, they are normally quite reliable. And when considered along with the entire dataset presented in this chapter, which includes around 4 300 observations on working in manufacturing across over 120 countries or political units and over a period of nearly 200 years, we can be reasonably confident about these estimates. Considered together, this dataset allows us to draw some broad insights into the history of working time in manufacturing.

Table 3.5. Number of observations in working-time dataset by region and decade, 1800-2010

	East Asia	Eastern Europe	Latin America and Caribbean	Middle East and North Africa	South and Southeast Asia	Sub-Saharan Africa	Western Europe	Western Offshoots
1800s	0	0	0	0	0	0	0	0
1810s	0	0	0	0	0	0	0	0
1820s	0	0	0	0	0	0	2	0
1830s	0	0	0	0	0	0	15	1
1840s	0	0	0	0	0	0	23	2
1850s	0	0	0	0	0	0	22	2
1860s	0	0	0	0	0	0	24	2
1870s	0	2	2	0	0	1	49	7
1880s	0	0	0	0	0	0	49	11
1890s	0	0	0	0	0	0	56	14
1900s	0	2	6	0	0	1	62	31
1910s	0	0	0	0	0	0	64	30
1920s	3	7	0	0	0	2	82	30
1930s	10	26	1	0	0	10	106	29
1940s	11	11	17	13	0	10	88	39
1950s	20	18	45	29	33	7	121	27
1960s	26	49	83	30	46	19	162	36
1970s	20	61	112	43	47	24	190	40
1980s	33	63	157	38	90	16	185	40
1990s	41	124	174	50	70	16	189	39
2000s	45	143	170	49	61	2	188	36

Holidays and annual working time

There are several reasons to be interested in extending the analysis of weekly working hours to annual working hours. First amongst these is that any consideration of the welfare implications of a reduction in working time would need to take into account the increase in the numbers of days of annual leave and holidays enjoyed by workers. Further, the levelling off in the reduction of weekly working time since the 1970s described above has been partially compensated for, at least in Western Europe, by a substantial increase in the number of days of leave and holidays over the same period.

Unfortunately, however, data on the number of days of leave and holidays enjoyed by workers is not widely available, especially for the entire panel of countries reviewed above. Therefore, this section relies on the data in Huberman and Minns (2007[1]) to describe the experience of a restricted panel of exclusively Western countries.[10]

Table 3.6 presents the number of days of leave and holidays enjoyed by workers in ten Western countries between 1870 and 2000. As can be seen, between 1870 and 1940 the number of days of leave and holidays increased in every country, with the 1920s seeing the emergence of paid leave. In the mid-20th century there was little change. Then in the later 20th century the number of days of leave and holidays enjoyed by workers in Western Europe increases substantially, while in the Western Offshoots there is almost no change in the number of days of leave and holidays, except for a small reduction for workers in Canada and the United States.

Table 3.6. Annual days of leave and holidays per year for manufacturing workers in selected countries 1870-2000

	Western Europe							Western Offshoots		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	AUS	CAN	USA
1870	14	4	19	13	23	31	11	8	8	4
1910	20	5	23	18	24	31	13	9	9	5
1938	30	21	33	31	37	44	28	22	22	17
1950	24	24	28	29	24		29	22	22	18
1980	28	33	30	29	35	30	30	32	25	22
1990	30	35	36	35	40	35	37	32	25	23
2000	33	38	36	43	42	36	38	32	24	20

StatLink https://stat.link/gwo8j4

As described in Ward, Zinni and Marianna (2018[32]), the methods in use today to measure annual working time are not currently standardised. Our estimates of the number of annual working hours are based on the method currently used by Eurofound (2019[33]) and Eurostat (2018[34])[11]. These are only rough estimates of annual working time, as some significant components of annual working time are not considered, such as extra hours worked (i.e. overtime) and hours not worked due to sickness absences, maternity/paternity and parental leave, strikes and lock-outs, etc. These estimates are not comparable in any case with those made by the OECD for the entire economy. The method used here involves calculating the number of annual hours of work by multiplying the number of weekly hours times 52, and then subtracting the number of days of public holidays and paid annual leave, converted into hours under varying assumptions about the length of the working week across different periods.[12]

$$(\text{Weekly hours x } 52) - (\text{Annual leave} + \text{Annual Holiday}) = \text{Annual Hours}$$

The resulting estimates for annual working hours of manufacturing workers are given in Table 3.7. According to this estimate, the number of hours worked per year in Western Europe continued to decline from 1980 to 2000, rather than plateauing as in the series of hours worked per week. In the Western Offshoots, on the other hand, the end to the reduction in weekly working time observable in Table 3.5 remains visible when looking at annual working hours.

Table 3.7. Annual working hours for manufacturing workers in selected countries, 1870-2000

	Western Europe							Western Offshoots		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	AUS	CAN	USA
1870	2 875	3 316	3 224	3 040	3 049	3 007	3 462	2 809	2 923	3 203
1910	2 516	3 070	2 799	2 688	2 912	2 640	2 707	2 457	2 823	2 815
1940*	2 175	2 306	1 980	1 944	1 985		2 208	2 063	2 258	2 031
1950	2 186	2 325	2 101	2 215	2 110		1 883	1 918	1 972	1 961
1980	1 986	1 828	1 801	1 876	1 739	1 698	1 750	1 691	1 809	1 921
1990	1 919	1 755	1 728	1 708	1 783	1 650	1 651	1 758	1 811	1 964
2000	1 867	1 698	1 616	1 648	1 650	1 618	1 659	1 731	1 795	1 958

Note: The calculation of the annual working hours figure for 1940 is based on data for days of leave and holidays in 1938.

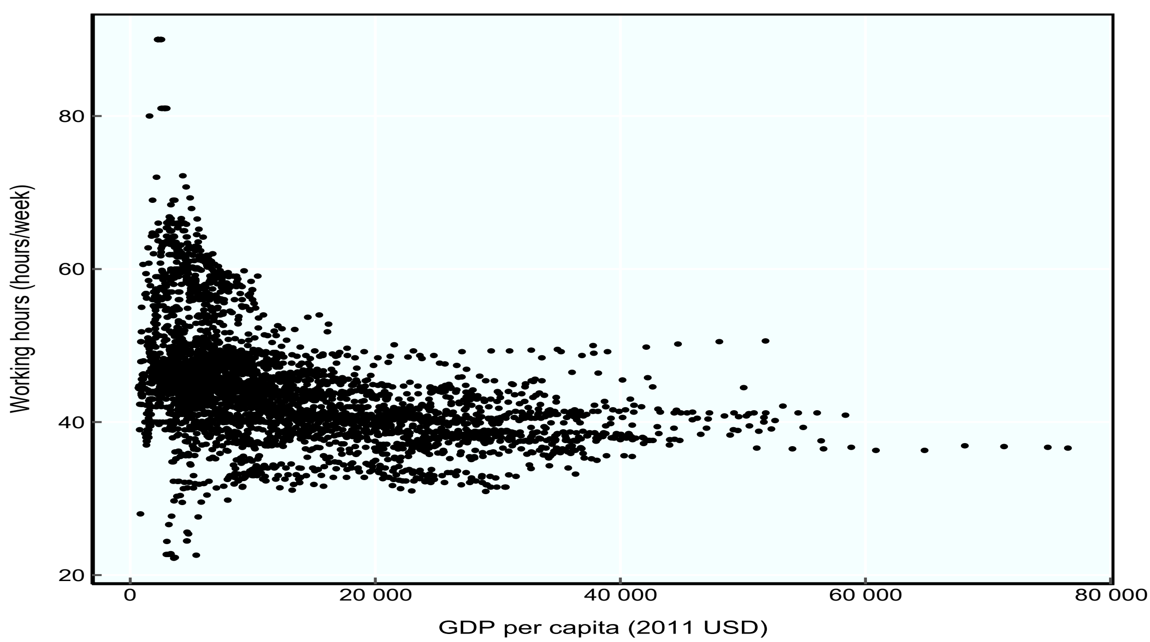
StatLink https://stat.link/c7q52w

Correlation with GDP per capita

We might expect that the relationship between GDP growth and working time would be a relatively simple one: as income increases, people will choose to work less, and therefore hours worked will decrease. This would imply a simple inverse linear relation between GDP and working time. However, the above description of developments of working time in manufacturing since 1820 does not support the view that working time has declined smoothly over time. Rather, there are both identifiable turning points and periods when, despite changes in GDP, there is relatively little change in the length of time that workers spend working.

Figure 3.2 presents the relation between working time and real GDP per capita across all countries included in our dataset. The data on real GDP per capita are taken from the Maddison Project (Bolt et al., 2018[35]). As can be seen, the relation is a quite tight non-linear relation.

Figure 3.2. Hours of work and GDP per capita since 1820



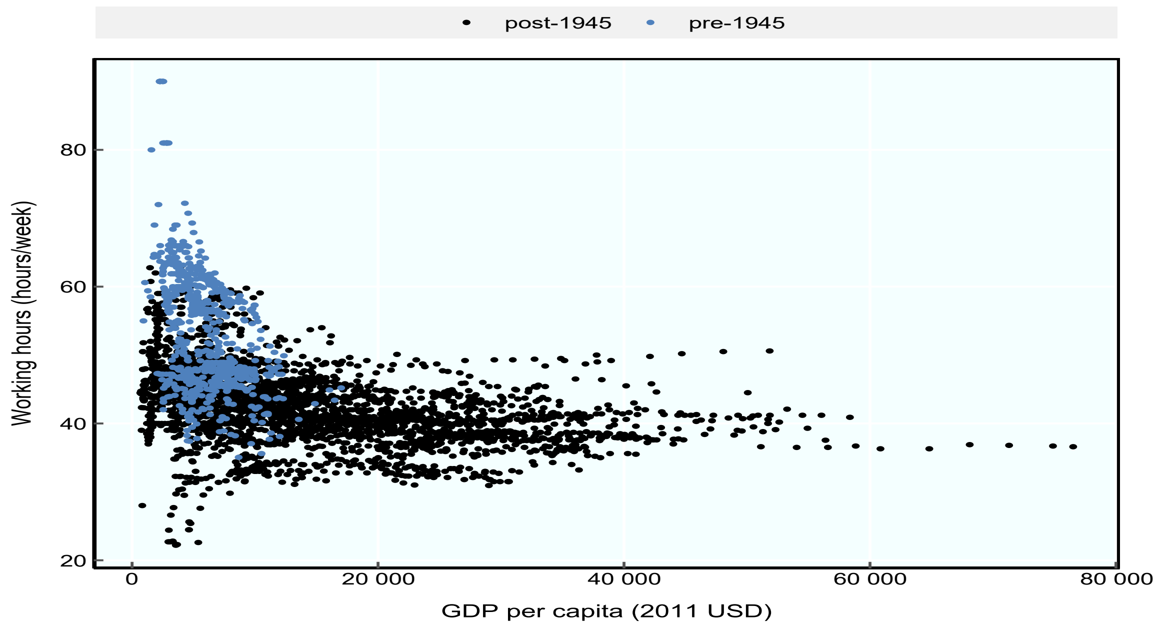
StatLink https://stat.link/xlt20b

The figure suggests that hours worked decline quite rapidly with increases in GDP when GDP per capita is below USD 20 000. When GDP per capita is above that point, the average working week clusters quite closely around 40 hours per week regardless of the country's level of income.

The concave positive relation between hours worked and GDP per capita visible in the bottom left corner of this figure reflects a small number of observations, all of which are from the period after 1970. These are primarily from countries such as Ethiopia, where the average working week in manufacturing in the 1990s and 2000s dropped to very low levels, which is presumably related to poor economic performance.

Figure 3.3 sheds further light on the historical dynamics of the relation between GDP per capita and weekly working hours in manufacturing. All blue dots are observations from before 1945, while all black dots are observations from after 1945. In both time periods, the relationship is almost linear, i.e. an increase in GDP per capita is associated with a reduction in working hours. But in the pre-1945 period, even a small increase in GDP per capita was associated with a substantial reduction in working time. After 1945, almost all the observations are clustered around 40 hours per week, implying that an increase in GDP is associated with only very small reductions in working hours. These figures hence provide some support to the idea that there are identifiable turning points in the relation between working time and GDP.

Figure 3.3. Hours of work and GDP per capita since 1820



Note: Data before 1945 are displayed in blue dots, those after 1945 in black dots.

StatLink https://stat.link/1ks453

Priorities for future research

While the data presented in the chapter represent a substantial improvement in our knowledge of long-run changes in working time, there are several ways in which this evidence could be improved in the future.

First, the data coverage for the early period of the dataset can be improved. It is highly likely that further national data series could be constructed for the early 20th century.

Second, this chapter has ignored both women's working time in manufacturing and working time outside manufacturing. It is highly likely that working time across the non-agricultural economy has tracked manufacturing closely, but that does not mean that the development of working time outside manufacturing is not worthy of further study. There are still many large unanswered questions here.

Third, while data on working time in the average working week are very useful for understanding the burden of work experienced by workers, they do not provide a complete picture. Historical research would gain from an improved understanding of holidays, breaks and leisure time, from data on annual hours of work and from knowledge of work patterns and work intensity during hours of work.

Finally, this chapter has been highly descriptive, with a focus on what has happened since 1820. Based on the data presented here, further analytic work on the causes and consequences of changes in working time can and should be pursued in the future.

References

Allen, R. and J. Weisdorf (2011), “Was there an ’industrious revolution’ before the industrial revolution? An empirical exercise for England, c. 1300-1830”, *The Economic History Review*, Vol. 64/3, pp. 715-729. [6]

Bienefeld, M. (1972), *Working Hours in British Industry: An Economic History, London School of Economics and Political Science*, Weidenfeld and Nicolson. [30]

Bolt, J. et al. (2018), “Rebasing ‘Maddison’: New income comparisons and the shape of long-run economic development”, *GGDC Research Memorandum*, Vol. 174. [35]

Clark, G. (2005), “The condition of the working class in England, 1209-2004”, *The Journal of Political Economy*, Vol. 113/6, pp. 1307-1340. [4]

Clark, G. (1994), “Factory discipline”, *The Journal of Economic History*, Vol. 54/1, pp. 128-163. [36]

Clark, G. and Y. Van der Werf (1998), “Work in progress? The industrious revolution”, *The Journal of Economic History*, Vol. 58/3, pp. 830-843. [31]

Crafts, N. (1997), “Economic growth in East Asia and Western Europe since 1950: Implications for living standards”, *National Institute Economic Review*, Vol. 4/97. [19]

de Jong, H. (2015), “Living standards in a modernizing world – A long-run perspective on material wellbeing and human development”, in *Global Handbook of Quality of Life*, Springer, Dordrecht. [2]

Eurofound (2019), *Working time in 2017-2018*, Publications Office of the European Union, https://op.europa.eu/en/publication-detail/-/publication/2d79f703-0a84-11ea-8c1f-01aa75ed71a1. [33]

Eurostat (2018), *Quality issues regarding the measurement of working time with the Labour Force Survey (LFS)*, Publications Office of the European Union, Luxembourg, https://ec.europa.eu/eurostat/documents/7870049/8734970/KS-FT-17-005-EN-N.pdf/0eb7518a-d0ca-45c4-942e-49107724e09a. [34]

Gershuny, J. and O. Sullivan (2019), *What We Really Do All Day: Insights from the Centre for Time Use Research.*, Pelican, UK. [11]

Gilmore (forthcoming), *Working Time in Industry: A New Dataset*. [24]

Hoffman, A. (1998), *Latin American Economic Development. A Causal Analysis in Historical Perspective*, GGDC monograph series, No.3, Groningen. [20]

Huberman, M. (2012), *Odd Couple: International Trade and Labor Standards in History*, Yale University Press, New Haven. [22]

Huberman, M. (2004), “Working Hours of the World Unite? New International Evidence of Work-time, 1870-1913”, *The Journal of Economic History*, Vol. 64/4, pp. 964-1001. [21]

Huberman, M. and C. Minns (2007), “The times they are not changin’: Days and hours of work in Old and New Worlds, 1870-2000”, *Explorations in Economic History*, Vol. 44/4, pp. 538-567. [1]

Humphries, J. and J. Weisdorf (2017), *Unreal Wages? Real Income and Economic Growth in England, 1260-1850*. [9]

International Labour Organization (2008), *Measurement of working time, Report II*, 18th International Conference of Labour Statisticians (Geneva), https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_099576.pdf (accessed on 11 January 2020). [12]

International Labour Organization (1962), *Resolution concerning statistics of hours of work*, 10th International Conference of Labour Statisticians (Geneva), https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_112455.pdf (accessed on 11 January 2020). [13]

Isett, C. (2016), *The social history of agriculture from the origins to the current crisis*, Rowman & Littlefield. [7]

Le Goff, J. (1982), *Time, work & culture in the Middle Ages*, Chicago: The University of Chicago Press. [5]

Maddison, A. (2001), *The World Economy: A Millennial Perspective*, Development Centre Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264189980-en. [17]

Maddison, A. (1995), *Monitoring the World Economy, 1820-1992*, OECD Development Centre, Paris, http://www.ggdc.net/maddison/Monitoring.shtml (accessed on 11 January 2020). [15]

Maddison, A. (1982), *Phases of Capitalist Development*, New York: Oxford University Press. [16]

Maddison, A. (1964), *Economic Growth in the West*, New York: Twentieth Century Fund. [14]

Mattesini, F. and B. Quintieri (2006), “Does a reduction in the length of the working week reduce unemployment? Some evidence from the Italian economy during the Great Depression”, *Explorations in Economic History*, Vol. 43/3, pp. 413-437. [37]

Mitchell, B. (2013), *International Historical Statistics.*, Palgrave Macmillan Publishers Limited. [26]

Neumann, T., J. Taylor and P. Fishback (2013), “Fluctuations in weekly hours and total hours worked over the past 90 years and the importance of changes in federal policy toward job sharing”, *National Bureau of Economic Research*, Vol. w18816. [38]

Pollard, S. (1965), *The Genesis of Modern Management: A Study of the Industrial Revolution in Great Britain*, Penguin Books. [39]

Reid, D. (1976), “The decline of Saint Monday 1766-1876”, *Past & Present*, Vol. 71, pp. 76-101. [29]

Sahlins, M. (1972), *Stone Age Economics*, Chicago: Aldine-Atherton. [3]

Spoerer, M. and J. Streb (2008), “Leben, um zu arbeiten, oder arbeiten, um zu leben? Warum uns der Rückgang der Jahresarbeitszeit in den letzten 125 Jahren nicht beunruhigen sollte”, *Geschichte Und Gesellschaft*, Vol. 34/1, pp. 116-128. [27]

Stephenson, J. (2018), “Real wages? Contractors, workers, and pay in London building trades, 1650-1800”, *The Economic History Review*, Vol. 71/1, pp. 106-132. [10]

The Conference Board (2019), *Total Economy Database™*, https://www.conference-board.org/data/economydatabase/TED1. [18]

Thompson, E. (1967), “Time, work-discipline, and industrial capitalism”, *Past & Present*, Vol. 38, pp. 56-97. [28]

U.S. Department of Labor (1990), *Fifteenth Annual Report of the Commissioner of Labor: Wages in Commercial Countries. 2 vols. Washington, DC: GPO.*. [23]

van Zanden, J. et al. (eds.) (2014), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-en. [25]

Voth, H. (1998), "Time and Work in Eighteenth-Century London", *The Journal of Economic History*, Vol. 58/1, pp. 29-58. [8]

Ward, A., M. Zinni and P. Marianna (2018), "International productivity gaps: Are labour input measures comparable?", *OECD Statistics Working Papers*, No. 2018/12, OECD Publishing, Paris, https://dx.doi.org/10.1787/5b43c728-en. [32]

Notes

[1] Oisín Gilmore would like to thank Herman de Jong, Rick Veldkamp, Pedro Miguel, the editors of this volume and other participants in the OECD workshops on historical well-being held in Utrecht in September 2018 and in Paris in June 2019. This chapter was supported by the Institut für die Geschichte und Zukunft der Arbeit (IGZA) and the Netherlands Organization for Scientific Research (NWO).

[2] Almost all the 120 political units are sovereign nations. However, a small number of other political units (such as the Saar Protectorate, West Berlin, the Palestinian State, etc.) are also covered by the database.

[3] See Isett and Miller (2016, p. 70[7]), for a description of some of these differences.

[4] It should be noted that while there is a broad consensus on this point, a number of researchers (Bienefeld, 1972[30]; Voth, 1998[8]; Allen and Weisdorf, 2011[6]; Humphries and Weisdorf, 2017[9]; Stephenson, 2018[10]; Clark and Van der Werf, 1998[31]) disagree.

[5] All the figures in this chapter as well as in Table 3.4 and Table 3.5 are based on data for all 120 political units included in our database.

[6] A very small number of observations were adjusted in other manners, as described in Gilmore (forthcoming[24]).

[7] For a discussion of this transformation see Clark (1994[36]) and Pollard (1965[39]).

[8] Parish apprenticeships were a means through which poor, illegitimate and orphaned children could be put to work and provided with some training in how to work. Generally, these "apprenticeships" were for very low status work.

[9] The 40-hour work week was introduced in France in 1936. Around the same period, significant reductions of working time also occurred in Italy and the United States. See Mattesini and Quintieri (2006[37]); Neumann, Taylor and Fishback (2013[38]).

[10] These figures are for the entire economy, not just manufacturing.

[11] In Eurofound (2019[33]) the figure for weekly hours is normal hours not actual hours as is the case in the estimates presented in this chapter. Eurostat (2018[34]) discuss some aspects of the difference between these two figures.

[12] For this calculation, days of leave have to be converted into annual hours of leave. This can be done using the following equation:

$$Total\ days\ of\ leave\ x\ \frac{Actual\ weekly\ working\ hours}{Working\ days\ per\ week} = Annual\ hours\ of\ leave$$

The Eurofound (2019[33]) and Eurostat (2018[34]) studies assume 5 working days per week. However, prior to World War I, manufacturing workers typically worked 6 working days per week, with the 5-day week introduced in most Western countries only after 1960. By 1980, the 5-day week was a well-established norm across Western Europe and the Western Offshoots (Gilmore, forthcoming[24]).Therefore, for the above calculation, I assume a working week of 6 days in 1870 and 1910, and of 5.5 days in 1940 and 1950.

4 Social spending and the welfare state

Peter Lindert,
Distinguished research Professor of Economics, University of California – Davis

The use of social spending to provide safety nets barely existed before 1820. In the next two centuries, it spread around the world. Countries now differ greatly in their commitments to social spending, which continue to take a larger share of national product in richer countries toward the north and west, and lower shares in poorer countries to the south and east. The most striking trend in the make-up of government social spending is the long drift from public investments in the young towards public subsidies to the elderly.

Introduction

Social assistance and social insurance are important indirect correlates of a population's well-being. Spending on such programmes was tiny or non-existent until political developments assigned more and more of these tasks to government within the last hundred years.

Today's political-historical landscape receives abstract welfare-economic support from familiar textbook reasons for tax-based social spending. There is a case for social spending based on capital-market imperfections whenever private loans could not have done the job of repairing the insufficiency of insurance, thereby smoothing consumption. Even where loans could have smoothed private consumption, externalities could justify government subsidies to income equalisation and social investments such as public health or public schooling. In addition, paternalism may justify forcing people to save more for old age, through mandatory contributions to their own pensions.

These justifications, while plausible, are static and ahistorical. Their call for social spending does not match the types, timing or geography of the social spending observed in world history. That is, the times and places with the greatest social spending are not those for which the capital imperfections and externalities should have called for it the most. On the other side of the same coin, the world's failure to engage in any sizeable social spending until the late 18th century clashes with the prevalence of the capital-market imperfections and the externalities that should have brought massive social assistance and social insurance in earlier centuries.

This chapter fills in the history of how government expenditures on social assistance and social insurance were first introduced and then expanded. These programmes seem to be attributes of richer and earlier-developing countries. The use of social spending as safety nets continues to spread. Today at least 164 countries around the world have such programmes, delivering about 10% of world product to targeted beneficiaries as of 2017.[1]

The evidence reveals a striking trend in what kinds of social supports are offered. Before World War I, social spending, while meagre, focused mainly on providing mass education, a strongly pro-growth policy. Yet in today's social budgets, the emphasis is largely, probably too largely, on non-contributory transfers to the elderly, implicitly at the expense of investments in the young.

Description of the concepts used

The central measure of social expenditures used in this chapter is the share of *gross public social spending* in GDP. These expenditures are "gross" in that they do not deduct any taxation, nor any present or past contributions paid by the beneficiaries. They are "public" in the sense that the government collects and distributes the funds, thus excluding private expenditures, whether or not they are mandated or controlled by government or whether or not they are stimulated by governments through tax breaks. That is, this definition is linked to the source of finance, not to the locus of control over how the funds are spent.

While it is sometimes convenient to call social expenditures "transfers", they are not necessarily transfer payments from one part of society to another. Some of them are funded partly by the prior contributions from the recipients themselves. And some are not strictly transfer payments in the national accounting sense either. Some of these expenditures, most notably health care expenditures, are payments for final services and are, unlike transfer payments, counted in GNP or GDP.[2] At the same time, public education expenditures, which should have been conceptually comparable to health expenditures, are excluded from social spending data reported here.

These arbitrary choices were made for practical reasons. To refer to all "social expenditures", one should have included public education subsidies.[3] Yet the OECD's social expenditure series does not do so. To present a concept matching the OECD's gold-standard practice, most of this chapter restricts its measures to social spending excluding public education – as in an earlier treatise (Lindert, 2014[1]). Still, although public education expenditures are excluded from the main aggregates presented here, their magnitudes and patterns will be noted at times, to help reveal patterns in political priorities.

Measures of social spending sometimes exclude some budgetary flows serving the same objectives as social assistance and social insurance. They omit most subsidies based on sectoral output, even though such product subsidies sometimes perform safety-net and income-equalising functions. Examples include farm subsidies and price supports, as well as Egypt's energy subsidies, *baladi* bread and food ration cards, which are not transfers between groups based on their income or wealth status. Similarly, the social expenditure measures exclude payments of wages under such public works programmes as India's National Rural Employment Guarantee. As for tax-break "revenue expenditures" such as earned-income tax credits (EITC), some partial progress has been made to include them as social expenditures. While other sources seem to omit these, the OECD's concept of net public social expenditures now includes revenue expenditures in its updates wherever they are reported, e.g. the EITC for the United States back to 1980, and back to the first positive EITC flows in Israel and Korea.[4] These deserve a place in the telling of the history of social spending, and the OECD measures including them are used in this chapter's data for 1980 and later.

One other exclusion should be made whenever possible – yet most data sources do not allow this exclusion. In principle, our measure of public benefits should include only their "non-contributory" component, i. e. the part transferred from the rest of society. Consider public pensions, a classic social insurance programme. How should the contributory part, paid for by the beneficiaries themselves, be excluded? One way has been to subtract the targeted group's contributions into public pension funds in the current year from the gross public pension benefits they receive in this same year. Such convenient period measures, unfortunately, compare contributions into *future* retirement by today's *workers* with benefits reflecting, in part, the *past* contributions by today's *retirees*. These same-year measures might work well enough in a steady state but are more problematic in the real world's growing pension budgets for growing elderly populations. It would be better to use cohort measures, following each birth cohort over its lifetime.

Furthermore, some of the convenient same-year measures often mix the contributory and non-contributory pension revenues together. Where the two kinds of programmes are separated in the underlying data, the Commitment to Equity (CEQ) Institute offers a "mid-range" compromise: call the strictly non-contributory programmes (e.g. universal minimum pensions) "subsidies" and add to this a value equal to half the benefits of the programmes for which we cannot separate the contributory and non-contributory components.

Until the start of this century, the OECD had lumped contributory and non-contributory sources of revenue together into the same aggregate. In recent years, however, the OECD has been able to quantify contributions to pension funds for a growing number of countries, and it is identifying them separately in the SOCX series. See Adema, Fron and Ladaique (2011[2]) and OECD (2019[3]).

This chapter presents measures of social spending by major programme objective – health care, aid to the elderly, labour market subsidies, family/welfare assistance and housing subsidies – for all available benchmark years.

Whenever possible, the chapter's measures of social spending draw on consolidated accounts for all levels of government – national, provincial and local.

Historical sources and data quality

This chapter's tables and figures map the same historical geography offered in the other chapters of this book. The chapter covers the 20 decadal benchmarks from 1820 through 2010 for 25 countries. There are, however, some data gaps dictated by underdevelopment and by historical turbulence. Table 4.1 inventories the available direct estimates of social spending. Lack of data is most serious for the intermediate period from 1950 through 1980, when many of these 25 countries with positive social spending did not publish the relevant figures. It seems likely, in any case, that countries with the largest shares of social spending in GDP were those reporting it. Beyond our 25 countries, there was a clear rise in the numbers of countries publishing direct estimates of social expenditures. For the world as a whole, the number of published estimates rose from two countries in 1820 to at least 164 countries reporting in or around 2010 (Table 4.1).[5,6] Social safety-net programmes continue to spread around the globe.

For some country/years, more than one source offers data. The international compilations can be ranked by their cross-country and inter-temporal consistency of accounting concepts, even though all of them draw on the same national sources. The source rankings can be arranged separately for the periods before and after 1945. For the latter period, the international-source rankings and "data quality" classes are as follows:

1. The *OECD's Social Expenditure database*, annually from 1980 on, receives the top ranking. As noted, the present chapter follows the OECD accounting concepts, based on its editing procedures and its better coverage of public expenditures, especially those by non-central governments. The number of countries covered by the OECD has grown from 21 to 35. Data class = 1 (best).
2. The *OECD's earlier (1985) social expenditure database* covered 19 countries from 1960 to 1981, though with less detail and less explanation than are offered for the new series since 1980.[7] Data class = 1.
3. Lustig (2017[4]) and Lustig (2018[5]), based on the Commitment to Equity Project, covers social expenditures in 30 low- and middle-income countries for 2008-2013. Data class = 1.
4. The Asian Development Bank (2013[6]) carefully detailed social spending in 32 countries as of 2009. The set of countries covered includes many in Oceania and Asia. Data class = 1.
5. Espuelas (2012[7]) and Espuelas (2013[8]) for Spain and Portugal in all years up to 2000 wove annual estimates from national sources. Data class = 2.
6. Hedberg, Karlsson and Häggqvist (2018[9]) offer annual series for the period 1920-1959 for Australia, Canada, France, Germany, Italy, the Netherlands, Sweden, the United Kingdom and the United States. This compilation draws from Flora (1983[10]) for continental Europe and from national sources for several other countries. Data class = 3.
7. For several other European countries in 1950 and 1960, we draw on Flora (1983[10]), which offers annual estimates, with gaps, up to 1978. Some categories are over-aggregated. Data class = 3.
8. For some non-OECD countries in the period 1990-1996, the ILO attempted to report "social security" expenditures, before yielding the estimation of social expenditures to other agencies, though it has recently shown signs of returning to this task.[8] Data class = 3.
9. *The World Bank's "ASPIRE" database* (The World Bank, n.d.[11]), concentrating on government spending targeted at poor populations in the 21st century, for developing countries all over the world. This omits social insurance expenditures and the often-large shares of benefits delivered to the non-poor. The low ranking of this series stems from our decision to target the broader social-expenditure concept now used by the OECD. Data class = 3, for present purposes.
10. Guesses based on other expenditure series, such as parts of social spending or over-aggregations such as "social and economic" spending. Data class = 4.

Table 4.1. Data quality of estimates of social expenditures, 1820-2010

	1820	1880	1910	1930	1950	1960	1980	1990	2000	2010
Western Europe										
GBR	3	2	2	2	3	1	1	1	1	1
NLD	3	2	2	2	3	1	1	1	1	1
FRA	3	2	2	2	3	1	1	1	1	1
DEU		3	4	2	3	1	1	1	1	1
ITA		2		2	3	1	1	1	1	1
ESP		2	2	2	2	1	1	1	1	1
SWE		2	2	2	3	1	1	1	1	1
Eastern Europe										
POL				2				1	1	1
RUS		2		3	4	3				1
Western Offshoots										
AUS			2	2	3	1	1	1	1	1
CAN			4	2	1	1	1	1	1	1
USA		2	2	2	1	1	1	1	1	1
Latin America and Caribbean										
MEX								1	1	1
BRA								3	4	4
ARG		3	3	3	2	2	1	1	1	1
Middle East and North Africa										
EGY										3
TUR							1	1	1	1
Sub-Saharan Africa										
KEN								3		3
NGA								3		4
ZAF				3						1
East Asia										
CHN				3	3	3	3	2	2	2
JPN		2	2	2	3	1	1	1	1	1
South and Southeast Asia										
IND								3		1
IDN								3		1
THA								3		1

Note: Data categories: High quality: the product of official statistical agency (national or international); 2 = Medium quality: the product of economic-historical research using the same sources and methods as applied by official statistical agencies; 3 = Moderate quality: economic historical research, but making use of indirect data and estimates; and 4 = Low quality: estimates based on a range of proxy information. In case of multiple sources, the lowest quality source is given.

Source: Aside from the country-specific sources identified in Annex 4.A, the sources are international data compilations. The preferred international sources are as follows: For 1880-1930, Lindert (1994[12]), Table 1A; For 1940 and 1950, Hedberg, Karlsson and Häggqvist (2018[9]); For 1960 and 1970, OECD (1985[13]); For 1980-2010, OECD SOCX database (n.d.[14]), last accessed on 12 May 2017. The 2010 estimates shown in this chapter do not reflect the revisions included in OECD (2016[15]), Table 5.9. From around 2010, a second preferred (class 1) database is the set of Asian Development Bank (2013[6]), covering the year 2009. For 2016, data for OECD countries are from SOCX, while those for Japan refer to 2015. See Annex 4.A for detailed country sources. The data in OECD (1985[13]) seem to exclude government spending on housing, an exclusion that is sizeable for some countries such as New Zealand and United Kingdom. For the latter country, the older OECD estimates for the period 1950-70 have been replaced with those from Thomson (1996[16]), p. 50, which include housing. No splicing is applied to link the older OECD's series for the period 1960 to 1981 with the new ones, following Ortiz-Ospina and Roser (2019[17]).

Each country's official sources are potentially the most reliable, as international compilations depend on these national sources. It may be the case, however, that a government's official data *publications* can be less reliable than the OECD or other publications, which have edited the nation's data for accounting consistency.

Reaching back to the pre-1945 era, the international compilations play out before 1880, and even the national statistical series fade away, as sketched globally in Table 4.1. For this earlier era the international-source preference rankings are as follows:

1. A peculiarity of the social spending measure for the pre-WW1 and interwar eras is that we often know from institutional history that the value must have been zero, either since the later-independent country did not exist yet or because there were no social programmes that would have generated such expenditures: so value = 0 and data class = 1. Care must be taken, of course, not to infer the lack of social spending from the lack of data, as signalled by Table 4.1's cautionary use of the probably zero symbol ("0?"). Searching through national legislative histories clarifies the dawn of spending when the compilations and natural statistical annuals fail to do so.
2. Lindert (1994[12]) unveiled social spending for 30 countries at the six decadal benchmarks 1880, 1890, … , 1930, starting from ILO studies in the 1930s (International Labour Office, 1933[18]; International Labour Office, 1936[19]), and mining national government sources back to 1880. These are now supplemented by his coverage of poor relief expenditures back to 1820. Data class = 2, except when this study confirmed true zeroes (so that class = 1).
3. Hedberg, Karlsson and Häggqvist (2018[9]) have pushed 18 countries' annual series on social expenditures, with and without education, back to 1920. For their 13 countries, they drew on Flora (1983[10]), supplemented by Espuelas (2012[7]) and Espuelas (2013[8]) for Spain, a study that extends back to 1850. For five non-European countries (Australia, Canada, Japan, New Zealand and the United States) they used national sources. Data class = 3.
4. Flora (1983[10]) sketched aggregate social expenditures for about 20 European countries back to the earliest statistical annuals in the 19th century, and Flora and Heidenheimer (1981[20]) gave starting-legislation years for a similar set of countries. Data class = 3, again excepting cases in which this study confirmed true zeroes (so that class = 1).
5. For six Latin American countries – Argentina, Chile, Costa Rica, Colombia, Peru and Uruguay – Arroyo, Lindert and Lindert (2017[21]) exploited the available statistical yearbooks back to their earliest publications in the 19th century. Data class = 3 (except = 1 for confirmed zeroes).
6. Van Bavel and Rijpma (2016[22]) have ventured earlier benchmark estimates for England-Wales, Italy and the Netherlands from 1850 all the way back to 1427, weaving together clues from both the revenue and the expenditure sides of government budgets. Data class = 3.

Supplementing these compilations for the pre-war and interwar periods will require applying the same procedures to individual-country sources that their compilations omitted.

Main highlights of social spending trends

Only very recently in the long sweep of world history has social spending, as defined here, absorbed a large share of the economy. So say Table 4.2 and Table 4.3 and Figure 4.1 for our 25 selected economies. If we were to take 20% of GDP as an arbitrary threshold defining the large-budget "welfare state," then no country had become a welfare state any time in world history before the 1960s. The main ascent took place between the 1930s and 1980s. By 1940 the United States had begun spending over 5% of GDP on New Deal Programmes, and a decade later Britain's Labour government was spending over 10% on public health, pensions and other social spending. On the 20% definition, the welfare state was ushered in first in Germany and the Netherlands around 1967, with the lead in spending share passing to Belgium in the 1970s, to Sweden in the 1980s and to France in the early 21st century.

While the rise in social spending shares has been accelerating over the long run, there were reversals for some of the top-spending countries. In fact, the 200-year span used in this volume happens to begin with a decline from a local peak in public aid to the poor, in England and Wales around 1820, after which poor relief was slashed, especially by the Poor Law Reform of 1834. Throughout northwest Europe, poor relief, the main form of social spending, remained low.

The late 20th century also saw some national retreats from generous social spending. Sweden's sweeping economic reforms launched by the Economics Commission's report (Lindbeck et al., 1994[23]) in 1993 included cutbacks in social spending, and a 1998 pension reform also introduced a systemic restraint on pensions in an aging world. In the mid-1990s, the Netherlands made deep cuts in certain kinds of spending, especially overly generous disability payments. Neither Sweden nor the Netherlands spends as great a share today as in 1990. Nonetheless, the overall international trend has been toward channelling higher shares of GDP into social spending.

Within the overall rise in social spending as a share of GDP, which kinds of spending rose the most, and what are the trends in the composition of social budgets? Of course, if we wanted to know only about the composition of the increase in the entire social budgets over the last 200 years, starting from near-zero spending, then we need to look only at the latest breakdown by type of spending, to see the net changes. Table 4.4 does this, by showing the composition of social spending as of 2010, a composition that remains roughly the same today.

Could the rise in gross public social spending have been offset by a decline in private charity, so that there was no true rise in overall social spending? Historical evidence, and simple arithmetic, reject this suspicion of a full "crowding out". Back in 1820, northern Europeans and North Americans surely did not privately transfer more than 20% of national product to vulnerable populations, as their welfare states are doing today. The best spotty evidence available for dates between 1500 and 1820 suggests that private and religious charities, aggregated to the national level, always gave less than 1% of national income, with the brief exception of the Netherlands in the mid-18th century. Furthermore, US data show that before 1927 poor and disabled people received little in private transfers, as in public transfers, and that since 1927 both private and public transfers have risen together, contrary to the crowding-out hypothesis.[9]

Another suspicion that can be set aside is that the dramatic rise in gross social spending might have been cancelled by an equal rise in the "clawback" of taxes paid by the recipients themselves. This may be partially true for social insurance programmes, given the difficulties of sorting out the prior contributions and insurance premiums paid by the current cohort of recipients. Still, the literature quantifying fiscal redistribution finds that most of the gross transfers between income ranks are also net transfers after taxes.[10]

Table 4.2. Total social spending since 1820: The first 100 years, 1820-1910

Percentage shares of GDP, excluding public education spending

	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910
Western Europe										
GBR	2.66	2.00	1.12	1.07	0.86	0.85	0.86	0.83	1.00	1.39
NLD	1.30	?	?	1.38	1.24	1.18	0.29	0.30	0.39	0.39
FRA	?	0.63	0.46	?	0.49	0.50	0.46	0.54	0.57	0.81
DEU	?	?	?	?	?	?	0.50	0.53	0.59	?
ITA	0?	0?	0?	0?	?	0.80	?	?	?	?
ESP	?	?	?	0.84	0.98	1.17	0.87	1.02	1.06	1.00
SWE	?	0.02	0.20	0.40	0.60	0.66	0.72	0.85	0.85	1.03
Eastern Europe										
POL										
RUS	0?	0?	0?	0?	?	0.59	0.59	0.59	0.58	?
Western Offshoots										
AUS										1.12
CAN						0	0	0	0	0.0046
USA	0?	0?	0?	0.12	0.12	0.13	0.29	0.45	0.55	0.56
Latin America and Caribbean										
MEX	0	0	0	0	0	0	0	0	0	0
BRA	0	0	0	0	0	0	0	0	0	0
ARG	0	0?	0?	0?	0?	0.05	0.08	0.02	0.05	0.04
Middle East and North Africa										
EGY										
TUR	0	0	0	0	0	0	0	0	0?	0?
Sub-Saharan Africa										
KEN										
NGA										
ZAF										
East Asia										
CHN	0	0	0	0	0	0	0	0	0	0
JPN	0	0	0	0	0	0?	0.05	0.11	0.17	0.18
South and Southeast Asia										
IND										
IDN										
THA	0	0	0	0	0	0	0	0	0	0

Note: Blank = not a sovereign state at that time (e.g. colonies); "0" = known to have none; "0?" = probably zero or below 0.1%; and "?" = no estimate yet, though true value > 0.

Source: Aside from the country-specific sources identified in Annex 4.A, the sources are international data compilations. The preferred international sources are as follows: For 1880-1930, Lindert (1994[12]), Table 1A; For 1940 and 1950, Hedberg, Karlsson and Häggqvist (2018[9]); For 1960 and 1970, OECD (1985[13]); For 1980-2010, OECD SOCX database (n.d.[14]), last accessed on 12 May 2017. The 2010 estimates shown in this chapter do not reflect the revisions included in OECD (2016[15]), Table 5.9. From around 2010, a second preferred (class 1) database is the set of Asian Development Bank (2013[6]), covering the year 2009. For 2016, data for OECD countries are from SOCX, while those for Japan refer to 2015. See Annex 4.A for detailed country sources. The data in OECD (1985[13]) seem to exclude government spending on housing, an exclusion that is sizeable for some countries such as New Zealand and United Kingdom. For the latter country, the older OECD estimates for the period 1950-70 have been replaced with those from Thomson (1996[16]), p. 50, which include housing. No splicing is applied to link the older OECD's series for the period 1960 to 1981 with the new ones, following Ortiz-Ospina and Roser (2019[17]).

StatLink https://stat.link/3tpi21

Table 4.3. Total social spending since 1820: Developments since 1920

Percentage shares of GDP, excluding public education spending

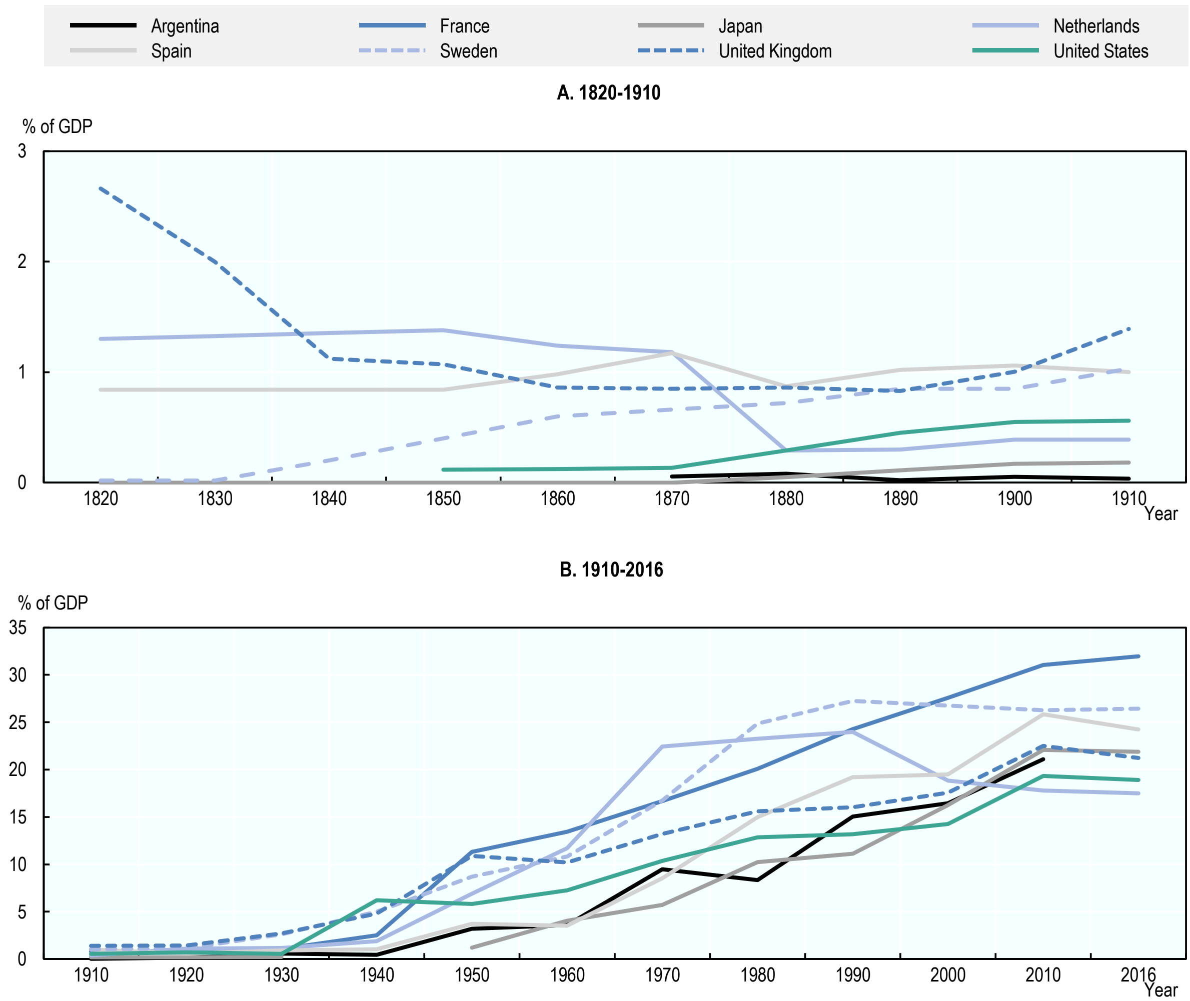
	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
					Western Europe					
GBR	1.42	2.69	4.80	10.90	10.21	13.20	15.59	14.90	16.19	22.42
NLD	1.10	1.15	1.90	6.90	11.70	22.45	23.26	23.99	18.85	17.78
FRA	0.64	1.08	2.50	11.30	13.42	16.68	20.07	24.28	27.58	31.04
DEU	?	4.96	?	14.80	18.10	19.53	21.79	21.35	25.39	25.90
ITA	?	0.10	4.20	8.70	13.10	16.94	17.38	20.70	22.68	27.12
ESP	0.57	0.91	1.05	3.73	3.52	8.53	14.98	19.20	19.48	24.72
SWE	1.14	2.60	5.00	8.70	10.83	16.76	24.84	27.24	26.77	26.26
					Eastern Europe					
POL	0.16	0.65	?	?	?	≥ 9.4	≥ 4.9	14.21	20.22	20.63
RUS	0?	4.31	?	?	18.38	22.03	?	?	?	13.09
					Western Offshoots					
AUS	1.68	2.11	4.30	5.60	7.39	7.37	10.26	13.14	18.25	16.59
CAN	0.06	0.31	3.20	6.10	9.12	11.80	13.31	17.55	15.76	17.53
USA	0.70	0.56	6.20	5.80	7.26	10.38	12.84	13.16	14.25	19.37
					Latin America and Caribbean					
MEX	0	0	?	?	?	?	?	3.14	4.39	7.37
BRA	0	0	?	?	?	?	?	4.49	12.80	14.90
ARG	0.17	0.58	0.45	3.20	3.60	9.47	8.34	15.05	16.45	21.11
					Middle East and North Africa					
EGY	0?	0?	0?	0?	0?	?	?	?	?	< 3.70
TUR	0?	0?	?	?	?	?	2.25	3.80	7.55	12.34
					Sub-Saharan Africa					
KEN						0?	0?	0.01	?	2.81
NGA						0?	0?	0.002	?	0.60
ZAF		0.26	?	?	?	?	?	?	?	11.05
					East Asia					
CHN	0	0.91	?	2.94	5.51	?	7.84	8.50	6.77	8.01
JPN	0.18	0.22	?	1.20	4.05	5.72	9.99	10.93	15.43	21.26
					South and Southeast Asia					
IND				0?	0?	0?	0?	0.32	?	2.29
IDN				0?	0?	0?	0?	0.17	?	2.63
THA	0?	0?	0?	0?	0?	0?	0?	0.36	?	3.60

Source: Aside from the country-specific sources identified in Annex 4.A, the sources are international data compilations. The preferred international sources are as follows: For 1880-1930, Lindert (1994[12]), Table 1A; For 1940 and 1950, Hedberg, Karlsson and Häggqvist (2018[9]); For 1960 and 1970, OECD (1985[13]); For 1980-2010, OECD SOCX database (n.d.[14]), last accessed on 12 May 2017. The 2010 estimates shown in this chapter do not reflect the revisions included in OECD (2016[15]), Table 5.9. From around 2010, a second preferred (class 1) database is the set of Asian Development Bank (2013[6]), covering the year 2009. For 2016, data for OECD countries are from SOCX, while those for Japan refer to 2015. See Annex 4.A for detailed country sources. The data in OECD (1985[13]) seem to exclude government spending on housing, an exclusion that is sizeable for some countries such as New Zealand and United Kingdom. For the latter country, the older OECD estimates for the period 1950-70 have been replaced with those from Thomson (1996[16]), p. 50, which include housing. No splicing is applied to link the older OECD's series for the period 1960 to 1981 with the new ones, following Ortiz-Ospina and Roser (2019[17]).

StatLink https://stat.link/t1qypc

Figure 4.1. Social transfer in seven countries, 1820-2016

Percentages of GDP



Note: For the assessment of data quality see Table 4.1.
Source: See Table 4.1 and Table 4.2.

StatLink https://stat.link/4wgiyb

Within the overall rise in social spending as a share of GDP, which kinds of spending rose the most, and what are the trends in the composition of social budgets? Of course, if we wanted to know only about the composition of the increase in the entire social budgets over the last 200 years, starting from near-zero spending, then we need to look only at the latest breakdown by type of spending, to see the net changes. Table 4.4 does this, by showing the composition of social spending as of 2010, a composition that remains roughly the same today.

Over all recent centuries, the kinds of government social spending that have grown the most as shares of GDP in Europe, Japan and the Americas are pension-and-survivor transfers to the elderly, followed by health subsidies and public education. Means-tested "welfare" payments and family benefits are a very small share of any welfare-state social budget, contrary to their dominance in media discussions.

Table 4.4. Government social spending in 25 countries in 2010, by type

Percentages of GDP

	All social spending (exc. educ.)	*of which* Health subsidies	Old-age benefits	Labour market	Welfare/ Family	Housing	*Memorandum Item:* Public education
Western Europe							
GBR	22.5	7.4	6.5	0.8	3.9	1.4	5.6
NLD	22.1	7.6	5.9	2.5	1.5	0.4	5.9
FRA	30.7	8.4	13.6	2.7	2.9	0.8	5.9
DEU	25.9	8.0	10.6	2.4	2.2	0.7	5.1
ITA	27.6	7.0	15.4	1.9	1.3	0.03	4.7
ESP	25.8	6.7	10.5	4.1	1.5	0.2	5.0
SWE	26.3	6.3	9.5	1.7	3.4	0.4	7.3
Eastern Europe							
POL	20.6	4.6	11.1	1.0	1.3	0.1	5.1
RUS	17.2	3.7	8.7	0.4	0.5	0.1	4.1
Western Offshoots							
AUS	17.5	5.8	4.5	0.8	2.6	0.4	5.1
CAN	17.6	7.3	4.3	1.1	1.3	0.4	5.0
USA	19.3	8.0	6.7	1.2	0.7	0.4	5.4
Latin America and Caribbean							
MEX	7.5	2.8	1.8	0.01	1.1	1.2	5.3
BRA	20.2	5.3	11.9	0.8	0.4	0	5.6
ARG	21.1	6.2	9.5				5.8
Middle East and North Africa							
EGY	(≥ 0.2)						3.8
TUR	12.8	4.2	7.8	0.1	0.3	0	2.9
Sub-Saharan Africa							
KEN	(≥ 2.5)						6.7
NGA	(≥ 0.2)						..
ZAF	10.6	4.1	1.3	0	1.8	1.5	6.0
East Asia							
CHN	8.0						3.6
JPN	22.1	7.3	11.6	0.6	1.3	0.1	3.8
South and Southeast Asia							
IND	1.7	0.05	0.2	0.7	0.1	0	3.3
IDN	1.2	0.1	0.1	0.04	0.4	0	3.0
THA	3.5	1.0	1.7	0.1	0.2	0	3.8

Source: For 2010 data for OECD countries, including Mexico, Poland and Turkey, the source is the OECD's Social Expenditure series, available through the OECD iLibrary as accessed in 2017. The 2019 version of the same dataset from the same source, used in updating Table 4.2 and Table 4.3, gives some different GDP shares for total social spending from those shown here. As of this writing, the breakdown in Table 4.4 cannot be updated yet from 2017 estimates to 2019 estimates. For the 5 Asian countries, the source is the set of Asian Development Bank (2013[6]) for 2009. For Argentina in 2009, the values refer to all levels of government consolidated; see Lustig (2017[4]) and the time series in Arroyo, Lindert and Lindert (2017[21]). For Brazil in 2009, the source is Higgins and Pereira (2014[24]), Table 1, which "includes spending at the federal, state and municipal levels" (p. 3). For South Africa in 2010-2011, the source is Inchauste and Lustig (2017[25]), Table 8.2, drawing on (Statistics South Africa, n.d.[26]), 2012a. "Financial Statistics of Consolidated General Government." Annual statistical release, Stats SA, Pretoria; the title of this publication implies that the numbers refer to consolidated general government, not just central government; contributory pensions are not included, nor are public sector pensions. The "(≥)" for China 2009, Egypt 2010, Kenya 2010 and Nigeria 2012 signifies that the social spending is likely to have been understated by these "All Social Assistance" numbers from the *World Bank's ASPIRE database* (The World Bank, n.d.[11]), since they exclude social spending of the "social insurance" variety.

StatLink https://stat.link/x3af7e

Table 4.5. Government social spending in 13 OECD countries in 1910, by type

Percentages of GDP

	All social spending (exc. educ.)	of which					*Memorandum item:* Public education
		Health subsidies	Old-age pensions	Labour market	Welfare / Family	Housing	
Western Europe							
GBR	1.39	0.32	0.38	0	0.68	0.01	0.74
NLD	0.39	0	0	0	0.39	0	1.77
FRA	0.81	0.26	0.26	0	0.29	0	0.99
ITA	0	0	0	0	0	0	0.52
DEU	0.60	0	0.12	0	0.42	0.06	2.72
NOR	1.18	0.27	0.05	0	0.86	0	2.80
ESP	0.02	0	0.02	0	0	0	0.41
SWE	1.03	0.31	0	0	0.72	0	1.26
Western Offshoots							
AUS	1.12	0.38	0.60	0	0.14	0	0.95
CAN	0	0	0	0	0	0	1.72
NZL	1.35	0.72	0.63	0	0	0	1.58
USA	0.56	0.26	0	0	0.30	0	1.42
East Asia							
JPN	0.18	0.10	0	0	0.08	0	0.29

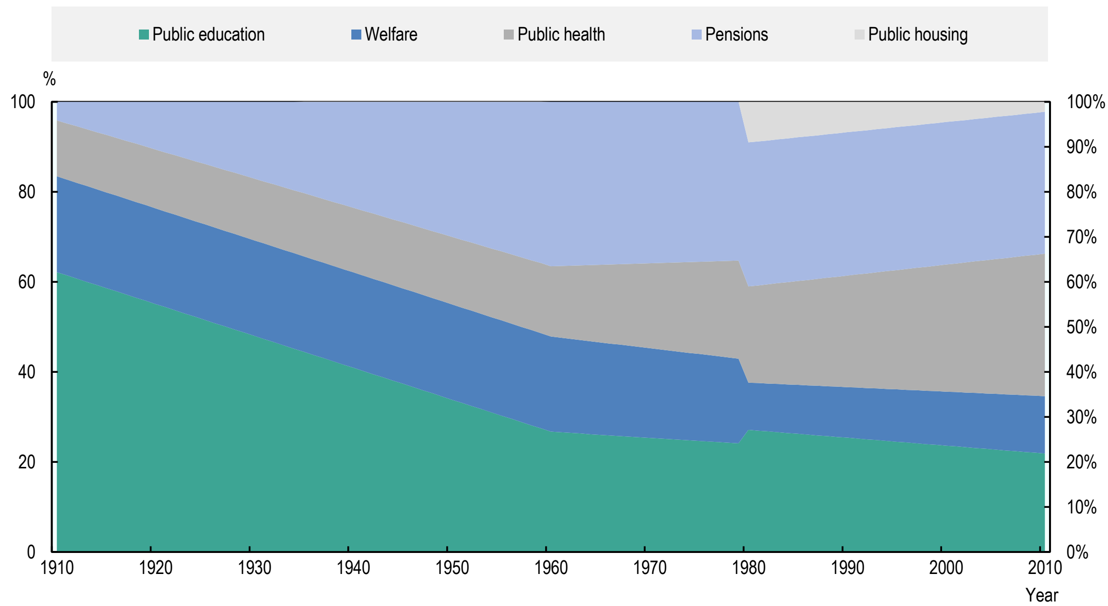
Source: For education, Lindert (2004[27]) see App. C, except for Japan, New Zealand and Spain, as noted below. For all other social expenditures, Lindert (1994[12]), Tables 1A-1E and its working-paper appendices. Data on public education expenditures for Japan in 1910 were kindly provided by Yuzuru Kumon; those for New Zealand in 1910 are from New Zealand (1912); those for Spain in 1910 are from Espuelas (2013[8]).

StatLink https://stat.link/9abj67

Yet within the last 200 years, there has been an important shift in the composition of social expenditures. To highlight just the largest of several shifts, we can compare the 2010 expenditures, just summarised in Table 4.4, with those of 100 years earlier, in Table 4.5. For a better sense of the timing of these changes, Figure 4.2 adds two benchmarks between 1910 and 2010, the first in 1960 (to highlight the progressive decline in the share of public education) and the second in 1980 (in correspondence to a break in the series). Before World War I, social spending was directed primarily toward education. Most advanced countries had public subsidies for mass schooling. Investing in schooling other people's children through taxes was a sign of an advanced country. Even Adam Smith and Thomas Jefferson, and later Milton Friedman, approved of subsidising mass primary education. Granted, the amounts of taxpayer money spent on public education were still below 2% of GDP as late as 1910 (Table 4.5). Yet advanced countries clearly invested in the young in the pre-war era. On the other hand, most of them had no unemployment compensation nor housing subsidies, nor even public pensions. Public education led the way.

Figure 4.2. Types of social expenditures, average of 13 OECD countries, 1910-2010

Percentage of the total



Note: The vertical line indicates a break in series. The 13 countries covered are Australia, Canada, France, Germany, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, the United Kingdom and the United States. The weight given to each country is its real gross national expenditures at international prices (rgdpe = real GDP in the Penn World Tables) for the year 1960. The data for public housing before the 1980 in the new series were too close to zero to show up on the figure. "Welfare" is the sum of family assistance, labour-market assistance and public housing subsidies. In the new OECD series for 1980 and 2010, pensions are the sum of all "elderly" benefits plus survivor benefits.
Source: See Table 4.1 and Table 4.2.

StatLink https://stat.link/lqwrf3

Over the hundred years between 1910 and 2010, and especially in the half-century before 1960, what grew most were public payments to the elderly. On first thought, this might seem to be an automatic by-product of population aging – surely society pays a greater share of GDP to the elderly as their share of the population expands. Yet the shift toward pensions has gone beyond that. Indeed, what have expanded fastest, even faster than public education, have been pension subsidies *per person of the targeted age group*, as shown in Table 4.6 and Figure 4.3. That is, social expenditures have shifted toward the elderly and (relatively) away from the young or the poor.

Table 4.6. Public pension support ratios and public education support ratios, 2010 and 1910

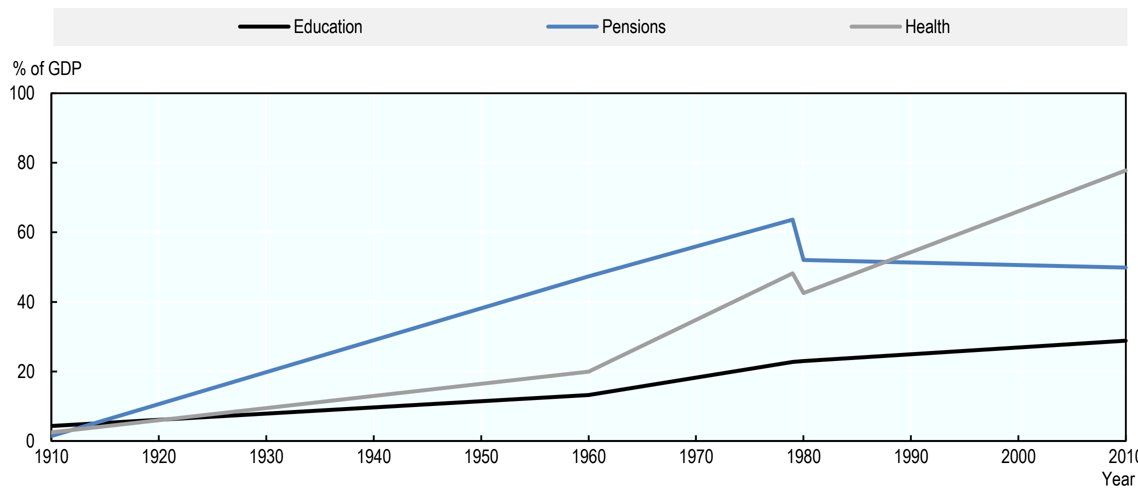
	Share of total population aged 5-19 (percentage)	Share of total population aged 65 and over (percentage)	Support ratios for school-age population (5-19)	Support ratios for the elderly (aged 65 and over)
In 1910				
Western Europe				
GBR	19.1	5.7	0.04	0.1
NLD	31.4	6.1	0.1	0
FRA	25.0	8.4	0.04	0.03
DEU	31.7	5.0	0.07	0.02
ITA	30.7	9.8	0.02	0
NOR	19.1	15.0	0.02	0.01
ESP	30.5	5.6	0.01	0.004
SWE	19.2	8.4	0.1	0
Western Offshoots				
AUS	29.9	4.3	0.03	0.14
CAN	30.2	4.7	0.1	0
NZL	20.7	13.0	0.07	0
USA	30.4	4.3	0.05	0
East Asia				
JPN	31.1	5.4	0.01	0
In 2010				
Western Europe				
GBR	17.9	16.2	0.35	0.40
NLD	17.9	15.6	0.33	0.38
FRA	18.2	17.0	0.32	0.80
DEU	14.6	20.6	0.35	0.51
ITA	14.2	20.4	0.32	0.75
NOR	19.1	15.0	0.38	0.75
ESP	14.1	17.2	0.36	0.61
SWE	17.5	18.2	0.54	0.52
Eastern Europe				
POL	16.2	13.5	0.32	0.82
RUS	15.5	13.1	0.26	0.66
Western Offshoots				
AUS	19.1	13.5	0.30	0.33
CAN	17.5	14.2	0.31	0.30
NZL	20.7	13.0	0.35	0.35
USA	20.3	13.0	0.27	0.51
Other countries				
MEX	29.8	5.9	0.17	0.31
BRA	26.0	6.7	0.22	1.78
ARG	25.4	10.4	0.23	0.91
TUR	26.7	7.0	0.11	1.11
ZAF	30.2	5.1	0.20	0.26
JPN	13.8	22.9	0.28	0.51
IND	30.2	5.1	0.11	0.04
IDN	28.2	4.9	0.11	0.02
THA	20.3	8.9	0.19	0.19

Source: Support ratios are defined as in Figure 4.3. Public expenditures as percentages of GDP are from Table 4.4 and Table 4.5; the age shares of total population are from (United Nations, Department of Economic and Social Affairs and Population Division, 2015[28]). Also see sources and notes to Table 4.1, Table 4.2 and Table 4.3.

StatLink https://stat.link/3tl8w1

Figure 4.3. Support ratios for public education, pensions and health, average of 13 OECD countries, 1910-2010

Public benefits per member of targeted age group as a percentage of GDP per capita

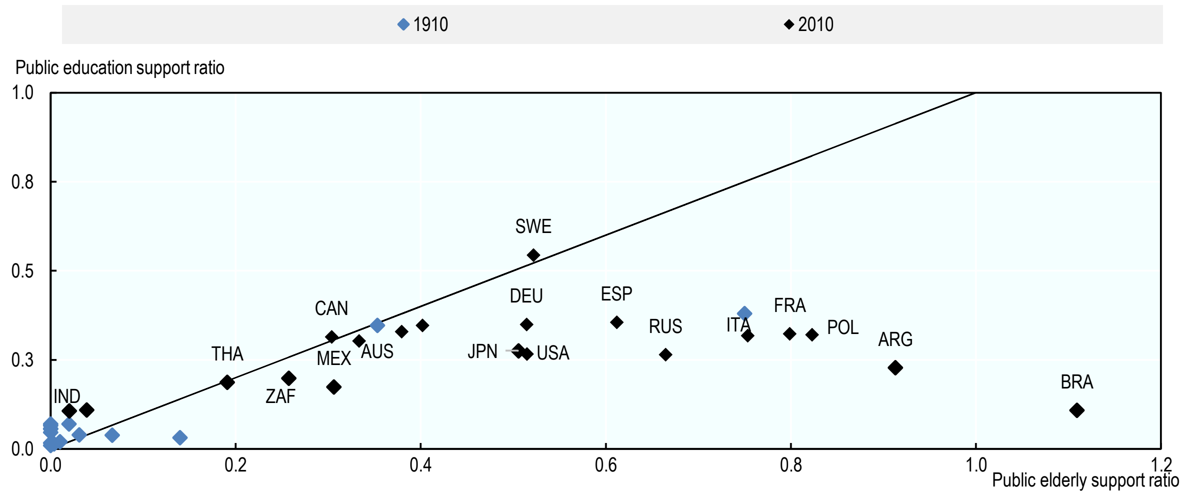


Note: The 13 countries included in the OECD average are Australia, Canada, France, Germany, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, the United Kingdom and the United States. The weight given to each country is its real gross national expenditures at international prices (rgdpe = real GDP in the Penn World Tables) for the year 1960. The data for public housing before 1980 in the new series were too close to zero to show up on the graph. "Welfare" is the sum of family assistance, labour-market assistance and public housing subsidies. In the new OECD series for 1980 and 2010, pensions are the sum of all "elderly" benefits plus survivor benefits.
Source: The sources are the same as those for Table 4.1, Table 4.2 and Table 4.3.

StatLink https://stat.link/l0e9ky

There is no obvious explanation for such a shift in the make-up of social spending on either efficiency or paternalistic grounds. As these countries prospered, as poverty declined, and as capital markets and public information improved, individuals should have been increasingly able to save for their own retirement, without government subsidy or compulsion. Note in Table 4.6 and Figure 4.4 those countries in which the shift toward pensions rose most over the last century. The generosity of support for the elderly increased the most in South America, Mediterranean countries and Eastern Europe. The literature on these issues shows that these three regions, and developing countries around the globe, share a telling bias in their policy preferences. Their pension generosity is biased toward those who served in higher-paying occupations in the formal sector, including the civil service and the military. Combining their bias in favour of higher-paid occupations with an overlapping bias against targeting social spending at the young means that transfers have reduced poverty more successfully for the elderly than for children or for those of working age, as the OECD has recently shown.[11]

Figure 4.4. Public pension support and public education support, 2010 and 1910



Note: Data for 1910, limited to 9 countries, are shown as blue diamonds (see Table 4.6). Public elderly-support ratio is the ratio between public pension and survivorship benefit payments per person aged 65 or over divided by GDP per capita. Public education-support ratio is the ratio between public education spending per person aged 5-19 divided by GDP per capita. No data are available for China, Egypt, Kenya and Nigeria. Regional markers are defined as in Figure 4.7.
Source: See Table 4.1, Table 4.2 and Table 4.3.

StatLink https://stat.link/i3xb7w

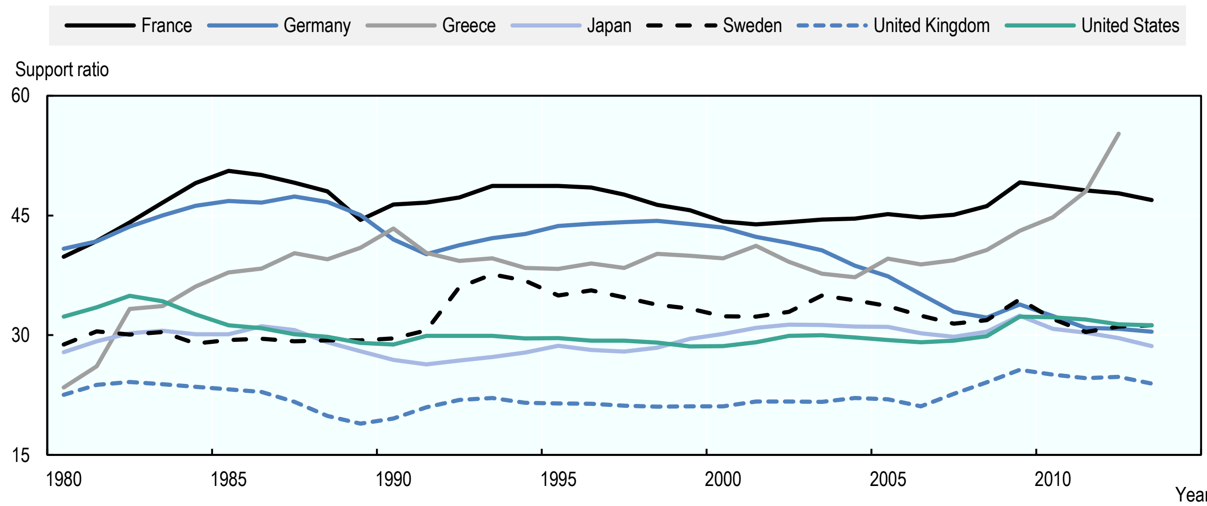
The combination of population aging and the rising generosity of annual support per elderly person over the last hundred years implies an ominous rise in the tax effort that younger age groups are paying to support the current elderly. So says this simple accounting identity:

Implied tax effort (transfers to elderly / GDP) =
(elderly share of population) times (public elderly-support ratio)

where the public elderly-support ratio is as defined in Table 4.6 and Figure 4.3 and Figure 4.4. To keep the tax effort from rising indefinitely in the presence of a higher elderly share of the population requires an offsetting drop in the support ratio. Over the last hundred years every term in this equation has been rising, which seems unsustainable. What does this portend for the future of pension finance?

Fortunately, in at least some countries, the 100-year rise in the annual-support ratio has been reversed since 1980, without actually reducing the real absolute value of the transfers per elderly person. Figure 4.5 shows that the support ratio has slowly declined since 1987 in Germany and since 1992 in Sweden.[12] Additional countries, not shown in Figure 4.5, have succeeded in slowly suppressing the same ratio in the face of population aging; Australia, Canada, the Netherlands and New Zealand have succeeded in doing so since 1980, and Estonia and Poland since 1995. On the other side of the coin, represented in Figure 4.5 by the extreme case of Greece, several countries have increased the generosity of support per elderly person, relative to average earnings. This has been the case since 1980 not only for Greece but also for Finland, France, Israel, Italy and Portugal, and slightly so for Denmark and the United Kingdom. These trends suggest an increasing risk for the latter group of countries, but not for the former.

Figure 4.5. The generosity of elderly support in six countries, 1980-2013



Note: The elderly support ratio is defined as old-age and survivor benefits per person aged 65 or older as a percentage of GDP per person aged 18 to 64.
Source: See Table 4.6 and Figure 4.2.

StatLink https://stat.link/cwhtai

Social spending under, and since, communism

How does one devise a historical measure for "social spending" during the communist command economy – i.e. Poland 1950-1990, Russia 1920-1990 and China 1950-1970? The institutions were very different from those in the other 22 countries covered in this chapter. Yet, on reflection, one can see that the complexity of the differences does not greatly affect our choice of social spending measures.

Consider, for example, a case in which the government confiscated a private house, and then let a government loyalist occupy it rent-free. This is a stark redistribution, one that does not easily fit the social insurance motive inspiring most social spending programmes. Redistributing that house also does not fit the notion of egalitarian social assistance, unless the recipient happens to be poorer than the previous private owner. Yet on both sides of the institutional chasm, the logical choice of a measure of *social spending* is the same. The fact that the government is providing the home free of charge means that the rental value of that home contributes the same value to gross social spending as in a market economy. In both settings, the government's transfers in kind (this house, in our example) or its transfers in cash are gross flows of social spending as long as they go to health, old age, disability, family needs, unemployment or housing.

How large were those expenditures as shares of GDP during the communist era? We now have at least partial benchmarks for the Soviet Union, Poland and China (Adema, Fron and Ladaique, 2014[29]).

Soviet social spending is at least partially illuminated by official numbers for 1930 and by the study (McAuley, 1979[30]) of Soviet living standards in the 1960s and 1970s (Table 4.2, Table 4.3). As of the early 1930s, during Stalin's First Five-Year Plan, official numbers say that the Soviet Union spent what amounted to 4.31% of estimated GDP on housing and other social programmes, perhaps the second-highest share in the world, behind Weimar Germany's 4.96%. The Soviet numbers do not, of course, reflect the unmeasured negative transfers to the peasantry during Stalin's collectivisation campaign. By the 1960s and 1970s, under Khrushchev and Brezhnev, Soviet social spending had risen to 15-17% of GDP, roughly comparable to that in the Western European countries that were about to become welfare states on the 20%-of-GDP criterion. Its mix of social expenditures was also similar to Western European practice, if we follow the usual OECD practice of excluding public education subsidies. However, the Soviets spent a larger share of GDP on public education – 6.7% in 1960 and 7.7% in 1970 – than any OECD country other than the United States, Canada and Sweden. The commitments to social spending and to education retreated somewhat, however, in the post-Soviet republics, including the Russian Federation. Broadly speaking, the shares of GDP they devoted to social spending plus public education settled back to the average practice of Western Europe.

For Poland in the communist era, the available estimates in Table 4.2 and Table 4.3 imply less social support than those in the Soviet Union, and much less than the generous supports offered by the Polish government since 1990. However, these numbers may cover social spending incompletely, as explained in Annex 4.A.

China under Mao (1949-1976) lagged behind the Soviet Union in social spending. In 1960, during the heavily communal living of the Great Leap Forward, it still spent only 5.51% of a slumping GDP, slightly above the 4.31% under Stalin during the first Soviet Five-Year Plan and well below the 1960 Soviet social expenditures (18.38% of GDP) under Khrushchev. During the post-Mao reforms, China expanded its social spending to 6.8-8.5% of GDP, though these fluctuating shares are well below those of post-communist Russia and Eastern Europe.[13]

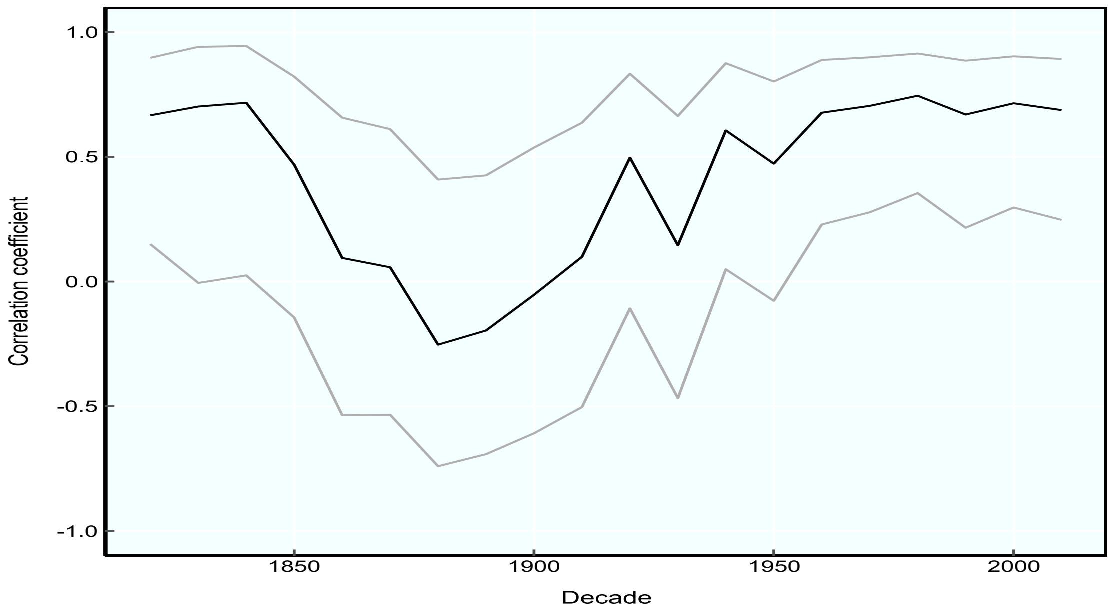
Looking at all post-communist or still-officially-communist Eurasian countries as a group since the 1990s, one can see a rough geographic pattern in their government commitments to social spending. Among these countries, social spending tends to rise from east to west, and sometimes also from south to north – just as it does in their never-communist near neighbours. In East Asia in 2009, Vietnam's central government spent less than 5% of GDP on social programmes, slightly less than in China (5.4% for central government only), whereas Mongolia's central government spent 9.6% (Asian Development Bank, 2013[6]). Among the central governments of formerly Soviet Central Asia and Transcaucasia, Tajikistan and Armenia spent less (1.2% and 2.2%), whereas the central governments of Kirghizia, Uzbekistan, Azerbaijan and Georgia spent somewhat more (6.1-10.2%), still less than the Russian Federation. Moving further west, Russia's spending was exceeded by its nearest neighbours to the west, and the highest spending shares in the former Soviet bloc are those of its western-tier states of Poland, the Czech Republic, Hungary and Slovenia, all of them around the arbitrary 20% threshold used in this chapter to define a welfare state.[14]

Correlation with GDP per capita and related social indicators

The fact that social spending has claimed its highest share of GDP only in recent times means that social spending must have been positively correlated with GDP per capita *over time*. Is it also positively correlated with GDP per capita *across countries* in a given year, like some other indicators surveyed in this volume?

Using just the 25 countries that are the focus of this book yields the correlation trends shown in Figure 4.6. The correlation between GDP per capita and the share of social spending in GDP was generally positive, even though errors in GDP measurement should have imparted a negative bias. Both at the start and at the end of our 200-year span, social spending tended to be significantly greater among the richer countries. This is what Figure 4.6 shows us for the early 1820-40 benchmarks and for the years since 1960. In between, however, the positive correlation vanished. This appears to have occurred because of the relative rise in GDP per capita of four rich frontier economies (Argentina, Australia, Canada and the United States) that were low social spenders in that interim period and began to catch up after 1890.

Figure 4.6. The correlation between social spending as a share of GDP and real GDP per capita among several countries, 1820-2010



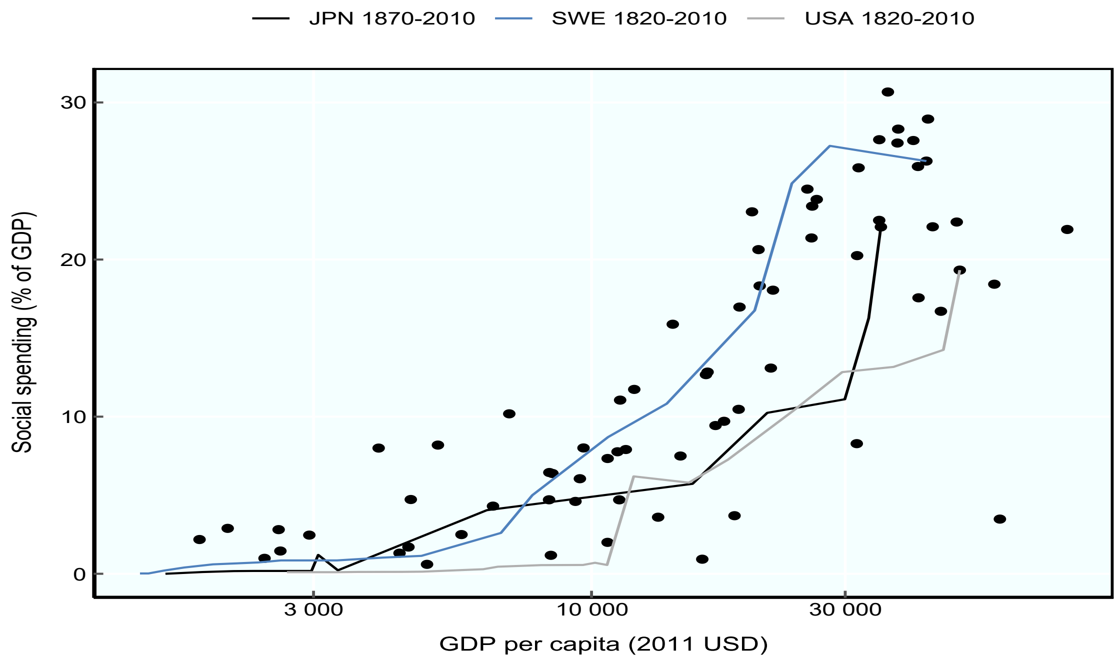
Note: The correlation is based on 8 countries for the period 1820-40, 12 countries for the period 1850-1940, and 14 countries for the period 1950-2010.
Source: For shares of social spending in nominal GDP, see Table 4.2 and Table 4.3. For real GDP per capita, see Chapter 2.

StatLink https://stat.link/0clwy6

Today, richer countries still commit a larger share of their resources to social spending, even if we look beyond our set of 25 countries to a larger group of countries supplying good data. Among the 70 countries shown in Figure 4.7, the correlation between the share of social spending in GDP and GDP per capita is 0.77.

In the early 21st century, the share of social spending in GDP is correlated not only with GDP per capita but also with several other societal outcomes with direct relevance to aggregate well-being. We would expect social spending to reduce income inequality, and it clearly does so, according to recent international compilations. Among 53 countries around the year 2013, social spending is highly correlated with the progressive fiscal redistribution of income (0.63), reducing the inequality of final incomes after taxes and transfers enough to reverse the fact that social spending tends to be lower in countries where pre-tax inequality is greater.[15] Spending a larger share of GDP on government social programmes is also associated with lower gender pay gaps.[16] Furthermore, large welfare states, particularly in northern Europe, have some of the world's cleanest and least corrupt governments, as surveyed by Transparency International, with lower budget deficits than the United States, Japan and other rich countries. And, for what it is worth, their populations express greater happiness in international surveys of public opinion.

Figure 4.7. Social spending as a share of GDP and GDP per capita in 2010, and paths followed by three leading countries since the 19th century



Source and note: See Table 4.4 and Table 4.5.

StatLink https://stat.link/o048k1

Priorities for future research

The available evidence invites a wide range of research projects that should help to understand trends in social spending. In particular, one cannot help noticing in Figure 4.7 that the levels of social spending can be much higher in one country than in another with a similar average income. Why is it that across Eurasia, social spending looms larger toward the northwest and much smaller toward the southeast even if we compare countries with about the same level of GDP per capita? The differences in behaviour are also great even among those countries that experienced communism and central planning between World War II and the start of the 1990s. For example, contrast in Figure 4.7 the generous social supports in Hungary with the near-zero supports in Azerbaijan. Why?

References

Adema, W. et al. (1996), “Net Public Social Expenditure”*, OECD Labour Market and Social Policy Occasional Papers*, No. 19, OECD Publishing, Paris, https://dx.doi.org/10.1787/122001486025. [66]

Adema, W., P. Fron and M. Ladaique (2014), “How much do OECD countries spend on social protection and how redistributive are their tax/benefit systems?”, *International Social Security Review*, Vol. 67/1. [29]

Adema, W., P. Fron and M. Ladaique (2011), “Is the European Welfare State Really More Expensive?: Indicators on Social Spending, 1980-2012; and a Manual to the OECD Social Expenditure Database (SOCX)”*, OECD Social, Employment and Migration Working Papers*, No. 124, OECD Publishing, Paris, https://dx.doi.org/10.1787/5kg2d2d4pbf0-en. [2]

Alexander Eckstein, W. (ed.) (1968), *Quantitative Trends in the Economy*, Social Science Research Council (U.S.), Committee on the Economy of China. [44]

Ansell, B. and D. Samuels (2014), *Inequality and Democratization: An Elite-Competition Approach*, Cambridge UK: Cambridge University Press. [57]

Arroyo, A., L. Lindert and P. Lindert (2017), “Fiscal Redistribution in Latin America since the Nineteenth Century”, *Has American Inequality Changed Direction? Looking over the Long Run*. [21]

Asian Development Bank (2013), *The Social Protection Index: Assessing Results for Asia and the Pacific*, Asian Development Bank, Mandaluyong City, Philippines, https://www.adb.org/sites/default/files/publication/30293/social-protection-index.pdf. [6]

CAPMAS, A. (2012), *The Statistical Yearbook. September*. [48]

Causa, O. and M. Hermansen (2017), “Income redistribution through taxes and transfers across OECD countries”*, OECD Economics Department Working Papers*, No. 1453, OECD Publishing, Paris, https://dx.doi.org/10.1787/bc7569c6-en. [51]

Cook, L. (2007), *Postcommunist Welfare States: Reform Politics in Russia and Eastern Europe*, Cornell University Press, http://www.jstor.org/stable/10.7591/j.ctt7zdbd (accessed on 12 January 2020). [67]

De Oliveira e Silva, P. (2017), “Social policy in Brazil (2004-2014): An overview”*, International Policy Centre for Inclusive Growth Working Paper*, No. 155, http://ipcig.org/pub/eng/WP155_Social_policy_in_Brazil.pdf. [42]

Dieguiez, H., J. Llach and A. Petrecolla (1990), *El Gasto Público Social*, Includes central administration, provinces, municipalities, state-owned enterprises, and social security. Ministerio de Economia, Buenos Aires. [32]

Dirección General de Estadística, Argentina (1915), *Extracto estadístico de la República Argentina*, Compañia Sud-Americana de Billetes de Banco. [34]

Dirección Nacional de Política Económica (n.d.), , http://www.mecon.gov.ar/peconomica/basehome/serie_gasto.html. [33]

Dostal, J. and G. Nakoshi (2017), “The Indonesian welfare state system: with special reference to social security extension in the development context”, in Aspalter, C. (ed.), *The Routledge International Handbook to Welfare State Systems*, London, New York: Routledge Taylor & Francis Group. 2017. [52]

Egypt Network for Integrated Development (2014), “Subsidies and the social safety net in Egypt”, *Policy Brief*, No. 012, https://socialprotection.org/discover/publications/subsidies-and-social-safety-net-egypt. [49]

Elveren, A. and T. Agartan (2017), *The Turkish Welfare State System: With Special Reference to Human Capital Development*, London: Routledge. [62]

Espuelas, S. (2013), “La evolución del gasto social público en España, 1850-2005”, *Estudios de Historia Economica*, Vol. 63. [8]

Espuelas, S. (2012), “Are dictatorships less redistributive? A comparative analysis of social spending in Europe, 1950-1980”, *European Review of Economic History*, Vol. 16/2, pp. 211-232. [7]

Ferreres, O. (2005), *Dos Siglos de Economía Argentina, 1810-2004: Historia Argentina En Cifras*, Editorial El Ateneo, Buenos Aires, Argentina. [40]

Flora, P. (1983), *State, Economy and Society in Western Europe, 1815-1975*, Frankfurt, Campus Verlag. [10]

Flora, P. and A. Heidenheimer (1981), *The Development of Welfare States in Europe and America*, London, Transaction Books. [20]

Garfinkel, I., L. Rainwater and T. Smeeding (2010), *Wealth and Welfare States. Is America a Laggard or Leader?*, Oxford University Press. [65]

Group, G. (ongoing), *Global Price and Income History Group: A Research Team Measuring Prices, Incomes, and Economic Wellbeing around the World Before 1950*, http://gpih.ucdavis.edu. (accessed on 12 January 2020). [60]

Gu, E. (2001), “Dismantling the Chinese mini-welfare state? Marketization and the politics of institutional transformation, 1979-1999”, *Communist and Post-Communist Studies*, Vol. 34, pp. 91-111. [47]

Hagen-Zanker, J. and H. Tavakoli (2012), *An analysis of fiscal space for social protection in Nigeria*, Overseas Development Institute Research Reports and Studies, https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7580.pdf (accessed on 12 January 2020). [55]

Hedberg, P., L. Karlsson and H. Häggqvist (2018), *How does openness to trade really affect the welfare state? Evidence from social spending in 21 countries, 1920-2000*, Conference Paper, XVIII World Economic History Congress, Boston, Massachusetts. [9]

Higgins, S. and C. Pereira (2014), “The Effects of Brazil’s High Taxation and Social Spending on the Distribution of Household Income”, *Public Finance Review*, Vol. 42/3, pp. 346-367. [24]

IBP Kenya (2015), *The Right Priorities? What Kenya’s National Government Spends Money On*, https://www.internationalbudget.org/wp-content/uploads/What-Does-the-Kenyan-Government-Spend-Money-On.pdf (accessed on 12 January 2020). [53]

IEERAL, I. (1986), “Estadísticas de la evolución económica de Argentina, 1913-1984”, *Estudios*, Vol. 9/36. [37]

ILO (2020), *Cost of Social Security 1990-96*, https://www.ilo.org/secsoc/areas-of-work/statistical-knowledge-base/cost-of-social-security-1990-96/lang--en/index.htm. [41]

IMF (1996, 2006), *International Financial Statistics*, International Monetary Fund, Washington DC. [59]

Inchauste, G. and N. Lustig (2017), *The Distributional Impact of Taxes and Transfers: Evidence from Eight Low- and Middle-Income Countries*, Washington: World Bank. [25]

InfoLEG (2014), *Presupuestos Nacionales*, http://www.infoleg.gob.ar/?page_id=837 (accessed on 12 January 2020). [36]

International Labour Office (1936), *International survey of social services*, ILO Studies and Reports, Series M, No.13, Kraus Reprint 1975, ILO, Geneva. [19]

International Labour Office (1933), *International survey of social services*, ILO Studies and Reports, Series M, No.11, Kraus Reprint 1975. ILO, Geneva. [18]

Kenworthy, L. (2019), *Social Democratic Capitalism*, Oxford University Press. [63]

Kuklo, C., J. Łukasiewicz and C. Leszczyńska (2018), *Historia Polski w Liczbach (History of Poland in numbers)*, Warsaw, https://stat.gov.pl/en/topics/other-studies/other-aggregated-studies/history-of-poland-in-numbers,8,1.html. [58]

Ladaique, M. (2014), *Measurement of Social Protection in OECD Countries. Re-thinking Social Protection in a Changing World*, ILO Regional Seminar. [68]

Lindbeck, A. et al. (1994), *Turning Sweden Around*, Cambridge: MIT Press, https://mitpress.mit.edu/books/turning-sweden-around. [23]

Lindert, K., E. Skoufias and J. Shapiro (2006), *Redistributing Income to the Poor and the Rich: Public Transfers in Latin America and the Caribbean*, The World Bank, Washington DC. [69]

Lindert, P. (2017), “The rise and future of progressive redistribution”*, Commitment to Equity (CEQ) Institute Working Paper, No. 73.* [70]

Lindert, P. (2014), *Private Welfare and the Welfare State*, Cambridge University Press. [1]

Lindert, P. (2004), *Growing Public. Social Spending and Economic Growth since the Eighteenth Century*, Cambridge University Press. [27]

Lindert, P. (1998), “Poor relief before the Welfare State: Britain versus the continent, 1780-1880”, *European Review of Economic History*, Vol. 2/2, pp. 101-140. [50]

Lindert, P. (1994), “The rise of social spending, 1880-1930”, *Explorations in Economic History*, Vol. 31/1, pp. 1-36. [12]

Liu, D. and K. Yeh (1965), *The Economy of the Chinese Mainland*, Princeton: Princeton University Press. [43]

Lustig, N. (2018), *Commitment to Equity Handbook. Estimating the Impact of Fiscal Policy on Inequality and Poverty*, Brooking Institution Press. [5]

Lustig, N. (2017), “Fiscal policy, income redistribution and poverty reduction in low and middle income countries”*, Commitment to Equity Institute Working Paper*, No. 54. [4]

Macon, J. (1985), *Las Finanzas Públicas Argentinas*, Includes central administration, provinces, and municipalities, Ediciones Macchi, Buenos Aires. [31]

McAuley, A. (1979), *Economic Welfare in the Soviet Union: Poverty, Living Standards, and Inequality*, University of Wisconsin Press, Madison. [30]

Mesa-Lago, C. (2017), “The Cuban welfare system: With special reference to universalism”, in *The Routledge International Handbook to Welfare Systems*, London, Routledge. [71]

Ministerio de Economía y Producción, A. (2006), *Presupuesto de la Administración Nacional: Gastos por Finalidad – Función y Naturaleza Económica, 1965-2006*, Ministerio de Economía y Producción. [38]

OECD (2019), *The OECD SOCX Manual, 2019 Edition. A guide to the OECD Social Expenditure Database*, https://www.oecd.org/social/soc/SOCX_Manuel_2019.pdf (accessed on 12 January 2020). [3]

OECD (2016), *Society at a Glance 2016: OECD Social Indicators*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264261488-en. [15]

OECD (2008), *Growing Unequal? Income Distribution and Poverty in OECD Countries*. [64]

OECD (2001), *1980-1998: 20 Years of Social Expenditure - The OECD Database*, OECD, Paris. [73]

OECD (1985), “Social expenditure 1960-1990. Problems of growth and control”*, OECD Social Policy Studies*, OECD, Paris, http://www.oecd.org/social/soc/40836112.pdf (accessed on 12 January 2020). [13]

OECD (n.d.), *Social Expenditure Database (SOCX)*, Covers social spending for the expanded OECD, 1980-2015, https://www.oecd.org/social/expenditure.htm. [14]

Oficina Nacional del Presupuesto, A. (2014), *Series estadísticas, Administración Nacional por Finalidad-Función y Naturaleza Económica, 2007 … 2013*, http://www.mecon.gov.ar/onp/html/#. [39]

Ortiz-Ospina, E. and M. Roser (2019), *Government Spending*, Published online at OurWorldInData.org, https://ourworldindata.org/government-spending. [17]

Statistics South Africa (n.d.), *Stats SA: Improving lives through data ecosystems*, http://www.statssa.gov.za/?s=database. [26]

The Global Economy.com (n.d.), *Kenya: Government spending, percent of GDP*, https://www.theglobaleconomy.com/Kenya/Government_size/ (accessed on 12 January 2020). [54]

The World Bank (n.d.), *ASPIRE Database*, http://datatopics.worldbank.org/aspire/indicator/social-expenditure. [11]

Thomson, D. (1996), *Selfish Generations? How Welfare States Grow Old*, White Horse Press, Cambridge, UK. [16]

Umukoro, N. (2013), “Poverty and Social Protection in Nigeria”, *Journal of Developing Societies*, Vol. 29/3 (September), pp. 305-322. [56]

Union of Soviet Socialist Republics (1936), *U.S.S.R. Handbook 1936*, Victor Gollancz Ltd, London. [61]

United Nations, Department of Economic and Social Affairs and Population Division (2015), *World Population Prospects: The 2015 Revision*. [28]

Van Bavel, B. and A. Rijpma (2016), “How important were formalized charity and social spending before the rise of the Welfare state? A long-run analysis of the selected Western European cases, 1400-1850”, *Economic History Review*, Vol. 69/1, pp. 159-187. [22]

Vasquez-Presedo, V. (1971), *Estadísticas Históricas Argentinas (comparadas), Primera Parte 1875-1914*, Ediciones Macchi, Buenos Aires. [35]

Wang, C. and K. Caminada (2012), “The redistributive effect of social transfer programmes and taxes: A decomposition across countries”, *International Social Security Review*, Vol. 65/3, pp. 27-48. [72]

Wang, Y. and A. Murie (1999), *Housing Policy and Practice in China*, St. Martin’s Press, New York. [45]

Zhang, X. (1998), *Privatisation: A Study of Housing Policy in Urban China*, Nova Science Publishers, New York. [46]

Annex 4.A. Data sources for individual countries

Countries covered in the chapter

Argentina

The estimate shown for 2010 refers to 2009. Sources for social spending: (a) 1950-1970: Macon (1985[31]), (b) 1970-1979: Dieguiez, Llach and Petrecolla (1990[32]), (c) 1980-2009: Dirección Nacional de Política Económica (n.d.[33]). Data refer to spending by central administration, provinces, municipalities, state-owned enterprises and social security. The "social assistance transfers" are larger after 1980 as the new source includes labour transfers. Other secondary sources used are as follows: for social expenditure, 1900-1915: Dirección General de Estadística (1915[34]); 1929-1939: Vasquez-Presedo (1971[35]); 1940-1962: Based on national budgets, InfoLeg, (2014[36]) and IEERAL (1986[37]); 1965-2006: Ministerio de Economía y Producción (2006[38]) and Oficina Nacional del Presupuesto (2014[39]); for nominal GDP: 1900-1932: Ferreres (2005[40]); 1932-2013: Ministerio de Economía y Producción (2006[38]) various years. Comparison with the CEQ Institute estimates for Argentina in 2012 suggests that these numbers include some contributory components to pensions. The difficulty is, again, to extract the contributions and derive the non-contributory component. For 2012, the CEQ compromise ("mid-range") estimates suggest a share of social spending of only 17.0% of GDP, versus the 20.5% yielded by the official aggregates that include all benefits paid from past contributions. Still, for comparability with the earlier Argentine benchmark totals, the larger estimate is used here.

Australia

Available federal data show positive social spending as early as 1910, even though Australia was officially a dominion of the British Empire (1901-1939), and not officially independent until 1986. Public pensions began in 1908. Housing subsidies were minimal before 1919. Health and welfare had very small coverage around 1910.

Brazil

The estimate shown for 1990 refers to 1994; the ILO (2020[41]) refers to 2002 (De Oliveira e Silva, 2017[42]), Table 3, which is also the source for 2010.

Canada

For the period before Dominion status in 1867, there clearly was no independent national budget, hence the values of o shown in this chapter. However, provinces could have had their own social programmes.

China

For 1933 (here “1930”), 1952 (“1950) and 1957 (“1960”), we have measures of “communal services” expenditure shares, plus the public share of estimated expenditures on housing services. The shares are defined in current prices for 1933 and 1952, with indirect evidence on price changes between 1952 and 1957. The main sources are Liu and Yeh (1965[43]), p. 68 and Liu (1968[44]), p. 138. Indirect evidence suggests that the shares of all housing service expenditures paid by government were 1% for 1933, a very rough 10% for 1952 and 61% for 1957. On these shares, see Wang and Murie (1999[45]), pp. 46-99, esp.

p. 81 and Zhang (1998[46]), pp. 18-42. For 1980-1996 ("2000"), see Gu (2001[47]), p. 94, which includes welfare spending and public housing for both *danwei* (workplace) and larger government units. The public housing expenditures for 1980 were extrapolated backward to 10 billion yuan from numbers given for 1981 and beyond. Estimates for the year 2012 (here shown as "2010") are from OECD (2016[15]), Table 5.9, citing "Asian Development Bank's Social Protection Index (SPI), World Health Organization (WHO)". Alternatively, the Asian Development Bank (2013[6]) gives a 2009 share of 5.4% of GDP, for central government only.

Egypt

For 1820-1870, Egypt seems to have had no tax-based social spending, like the rest of the Ottoman Empire. For 1880-1910, Egypt was under receivership and its finances were run by Britain and France, with no known social spending. For circa 2010, an official publication (CAPMAS, 2012[48]), pp. 413, 451, gives the "value of pensions & compensations for pensioners & their beneficiaries by reward system in the social insurance sector" as a percentage (8.4%) of GDP. As noted in the main text, most of Egypt's social subsidy programmes, such as energy subsidies, baladi bread and food ration cards, are nearly universal and not transfers between groups based on their income or wealth status (Egypt Network for Integrated Development, 2014[49]). They are thus excluded by the OECD concept measured in this chapter.

France

Estimates for 1830 refer to 1833, those for 1840 to 1841, and those for 1870 to 1871 (Lindert, 1998[50]), p. 113, while those for 1940 refer to 1938. The Flora (1983[10]) estimates for 1940 and 1950, taken from Hedberg, Karlsson and Häggqvist (2018[9]), tend to be higher than those used here for other years (higher than Lindert series in 1930 and higher than OECD in 1959 versus 1960).

Germany

Estimates are from our preferred international sources: Lindert (1994[12]) for 1880-1930, (Hedberg, Karlsson and Häggqvist (2018[9]) for 1950, OECD (1985[13]) for 1960 and 1970 and OECD (2017[51]) for 1980-2016.

India

The estimate for 1990 refers to April 1991-March 1992 (ILO, 2020[41]). For 2010 and latest, they refer to 2012 (2016[15]), Table 5.9, citing "Asian Development Bank's Social Protection Index (SPI), World Health Organization (WHO)".

Indonesia

The zeroes for 1950 and 1960 are confirmed by Dostal and Nakoshi (2017[52]), p. 368. The estimate for 1990 refers to April 1991-March 1992, derived from the receipts side of social programmes (ILO, 2020[41]). 2010 is from OECD (2016[15]), Table 5.9, citing "ILO."

Japan

The 1950 estimate is from Hedberg, Karlsson and Häggqvist (2018[9]) citing Statistics Japan. The estimate for 2016 refers to 2015.

Italy

For 1790, before the period covered here, van Bavel and Rijpma (2016[22]) give 1.1% of GDP as the social spending of “formal institutions”, most of which were not tax-funded. The similarly based 1870 estimate refers to 1868, from van Bavel and Rijpma (2016[22]), p. 171. The alternative estimates from Flora (1983[10]) for 1940 and 1950, sourced from Hedberg, Karlsson and Häggqvist (2018[9]) tend to be higher than those reported in this chapter for other dates (higher than Lindert series in 1930 and higher than OECD in 1959 versus 1960).

Kenya

The estimate for 1990 refers to 1994 (ILO, 2020[41]). For 2013/14, expenditures on Security = 16.1% of government expenditures and Health = 4.0% (IBP Kenya, 2015[53]). This presumably includes benefits paid out to civil servants. In turn, government spending = 14.0% of GDP for 2012-2013 (The Global Economy.com, n.d.[54]). So Security plus Health = 2.81% of GDP for 2012-2013.

Netherlands

The estimate for 1820 is for poor relief, from van Bavel and Rijpma (2016[22]), p. 171. Estimates for 1850-1870 are from Lindert (1998[50]), p. 113. Those for 1880-1930 are from Lindert (1994[12]), and those for 1950 from Hedberg, Karlsson and Häggqvist (2018[9]).

Nigeria

The estimate for 1990 refers to 1991 (ILO, 2020[41]). For 2010, the estimate of 0.6% of GDP excludes schemes for civil servants (Hagen-Zanker and Tavakoli, 2012[55]), p. 9. The same study notes major data caveats: "This study is based mostly on budget data, not actual spending data, with the exception of data on total government expenditure… A related problem is that corruption and leakage are very high. Leakages are hard to quantify". Hagen-Zanker and Tavakoli (2012[55]), p. 7. The bleak picture of corruption and programme non-delivery is confirmed by Umukoro (2013[56]).

Poland

The 1930 estimate from Lindert (1994[12]) is used here. Ansell and Samuels (2014[57]) obtained 0.16% of GDP for 1922, and 0.65% of GDP for 1930. For 1951-1990, social spending measures are available in Kuklo, Łukasiewicz and Leszczyńska (2018[58]). I am indebted to Mikolaj Malinowsky and to Cecylia Leszczyńska for providing these source materials. However, this source lacks nominal (current-price) GDP estimates. I have found nominal GDP numbers only for the 1970 and 1980 benchmarks, in (International Monetary Fund (1996, 2006[59]). For 1970 and 1980, the available estimates may omit parts of social spending. For both years, we have only government expenditures on education (omitted here), health protection and social care (ochrona zdrowia i opieka społeczna), financing (subsidies) and social security (finanse i ubezpieczenia społeczne). For 1970, estimates identify a small amount of housing subsidy (gospodarka mieszkaniowa oraz niematerialne), while those for 1980 give communal services (gospodarka komunalna and usługi komunalne). No coverage of subsidies for disability, family aid and labour-market subsidies is made explicit.

South Africa

The 1930 estimate is from Ansell and Samuels (2014[57]), with underlying data kindly supplied by David Samuels. For 2010, the CEQ "mid-range" estimate was used. A lower number (8.7%) would obtain if we had used year 2012 per OECD (2016[15]), Table 5.9, citing "National Budget 2014, Estimates of National Expenditure, National Treasury and World Health Organization (WHO)".

Spain

Note that estimates cover all kinds of social spending, even for the 19th century – unlike for other European countries before 1880, that cover only poor relief. The source is Espuelas (2013[8]) and Espuelas (2012[7]). Espuelas's work was used even up through the 1970 benchmark, since OECD (1985[13]) did not cover Spain (or Portugal).

Russia

For 1870-1910, Stephen Nafziger has kindly supplied budget numbers for the Imperial (central) government. Local governments probably allocated only negligible amounts to social spending, though possibly noticeable amounts to local public education. The small amounts shown here were rather regressive expenditures on bits of health care for the Army, Navy and a few other ministries with seemingly elite mandates. The nominal nations income denominator is available annually for 1885-1913 at Global Price and Income History Group (ongoing[60]). The Imperial government also spent a bit on education within the Army, Navy and four other ministries. These amounted to 0.47% of GDP in 1890 and 0.42% in 1900. For 1900, Ansell and Samuels (2014[57]), data provided by David Samuels, came up with the same share by different means. The 1930 estimates refer to the Soviet Union in 1931, and are derived from the 1936 Union of Soviet Socialist Republics (1936[61]). Categorical percentage shares of GDP for 1931 were welfare 1.19, pensions 1.56, health care 1.03 and housing 0.53. The housing estimate may be for expenditures building housing, and thus a capital-account entry, as opposed to an imputed rental value of housing provided or subsidised by government in 1931. For 1960 and 1970, estimates refer to social consumption expenditures (obshchestvennye fondy potrebleniya) of the entire Soviet Union, minus education and holiday pay, divided by McAuley's estimate of Soviet GDP (McAuley, 1979[30]), p. 262. Dividing by Mitchell's higher estimate of Soviet nominal GDP would have lowered these shares by about 30%. For 2010, our estimates use the CEQ Institute's "mid-range" measure, including half the benefits from partly contributory programmes plus all the benefits from strictly non-contributory programmes.

Sweden

For 1940 and 1950, the source is Flora (1983[10]) via Hedberg, Karlsson and Häggqvist (2018[9]). For the earlier 1930 benchmark, this source gives 3.5% instead of the 2.5% shown in this chapter, based on Lindert (1994[12]).

Thailand

The estimate for 1990 refers to 1993, and is derived from the receipts side of social programmes (ILO, 2020[41]). The estimate for 2010 refers to 2011, from the Asian Development Bank (2013[6]), Technical Report 44152 (2012, p. 27).

Turkey

Estimates for 1980-2010 use the OECD, our preferred source. The zeroes before 1950 seem quite secure. Elveren and Agartan (2017[62]), esp. p. 318, imply that even the first privileged social insurance for public sector employees was not legislated until 1946. The estimate for "2016" refers to 2013.

United Kingdom

Estimates refer to England-Wales through 1870, and to the United Kingdom thereafter. Estimates for England and Wales in 1820-1870 are for poor relief only, from Lindert (1998[50]), p. 114. Those for 1880-1930 are for all social spending, from Lindert (1994[12]). The social-transfer estimates for 1940 and 1950 are from Hedberg, Karlsson and Häggqvist (2018[9]), borrowing from Flora (1983[10]); those from 1960 on are from OECD (1985[13]) for 1960-70 numbers, and from OECD (n.d.[14]) for 1980-2016.

United States

For 1850-1870, see the sources cited in Lindert (2004[27]), Table 3.4. For 1880-1910, see Lindert (1994[12]), Table 2. For 1940 and 1950, see Hedberg, Karlsson and Häggqvist (2018[9]).

Notes

[1] The 10% value lies between two averages computed over the 25 countries that are the focus of this book. Weighting by population, as in other chapters of this volume, yields a value of 9.2% for 2010. On the other hand, weighting social transfer shares of GDP (correctly) by each country's nominal GDP implies that transfers took 15.8% of the 25-country GDP.

[2] Differently from National Accounts, another type of productive service is excluded from the social expenditure measure used in this chapter. The OECD series on social expenditures, which is used as a reference in this chapter, exclude administrative costs – see page 13 of OECD (2001[73]).

[3] Garfinkel, Rainwater and Smeeding (2010[65]) rightly argue for inclusion of public education expenditures. This author's joining the OECD in excluding them was originally a debating tactic, as explained in Chapter 1 of Lindert (2004[27]).

[4] See Adema and Pearson (1996[66]), Adema, Fron and Ladaique (2014[29]) and OECD (2019[3]).

[5] Of the 164 countries yielding data on social spending in 2010, 91 gave estimates of the desired aggregate expenditures, whereas for the other 73 countries we have only the *World Bank's ASPIRE* (The World Bank, n.d.[11]) estimates of the more progressive part of social expenditures defined as "social assistance".

[6] Presumably the data collections underlying the IMF's series *Government Finance Statistics* would include details on total social expenditures for all countries. Their data on the Internet, however, fail to provide useful breakdowns.

[7] The discrepancies between (1) and (2) for the years 1980 and 1981 are sometimes large and are not fully explained. The 1998 OECD manual (OECD, 2001[73]) notes that the new series are 2-3% lower than the old series due to a shift in the OECD accounting for GDP. This falls short of explaining why the new series tends to be 12% below the old for the median country covered in 1980.

[8] See the PowerPoint tables Ladaique (2014[68]).

[9] For the best available estimates on Italy, Netherlands and England before 1800, see van Bavel and Rijpma (2016[22]), esp. Table 1. See also Lindert (2004[27]), Chapter 3 and Lindert (2014[1]).

[10] See the whole set of developing-country studies by the Commitment to Equity (CEQ) Institute (Lustig, 2017[4]; Lustig, 2018[5]), plus Wang and Caminada (2012[72]), Lindert (2017[70]) and OECD (2019[3]). See also Adema, Fron and Ladaique (2014[29]).

[11] The geographical patterns of the regressive bias and the bias in favour of the elderly overlap but are not the same. The two occur together in Latin America (Lindert, Skoufias and Shapiro, 2006[69]; Arroyo, Lindert and Lindert, 2017[21]; Lustig, 2017[4]; Lustig, 2018[5]). The OECD study featured here (OECD, 2008[64]) (see Figure 5.12) also showed that poverty has been reduced more for the elderly than for children and those of working age, confirming an anti-young bias in policy reduction through government taxes and transfers. It did not cover Latin America or Eastern Europe, but did get this same result for Portugal and Italy, along with the United States, Japan and a few other countries.

[12] Note that Figure 4.3 defines the support ratio as relative to GDP per working-age population, whereas Table 4.6 and Figure 4.2 defined it as relative to GDP per capita. The patterns are similar, however.

[13] Again, see Table 4.2 and **Error! Reference source not found.** and the sources cited there. Public education expenditures, still excluded here, are also lower in China than in Russia and Eastern Europe after communism. From about 1990 to about 2010, China's public education expenditure drifted from 2.5% of GDP down to 1.9%, while the public-education share varied between 2.9 and 4.3% in Russia, and between 4.4 and 5.5% in Poland, according to UNESCO. China's spending on higher education seems to have expanded considerably since 2010, but we lack reliable estimates for this last decade.

[14] Cook (2007[67]), **Error! Reference source not found.**, Lustig (2018[5]) and Asian Development Bank (2013[6]). Extending still further west in the set of countries that was communist before 1990, Cuba has at times recorded the highest shares of GDP devoted to social spending. Even if one excludes public education spending, its social-spending share reached over 22% of GDP in 2008. Its public-education shares were also perhaps the highest in the world at 11-14% of GDP (Mesa-Lago, 2017[71]). These ratios speak not only to the achievements of Cuban public health and education, but also to the depressed state of the country's GDP.

[15] Lindert (2017[70]). Similarly, social spending and overall GDP per capita tend to be greater in countries where the income shares going to the top 1% or to the top 10% tend to be lower (Kenworthy, 2019[63]).

[16] So say the underlying numbers behind the negative correlation between GDP per capita and the gender pay gap in Kenworthy (2019[63]) on page 19. Kenworthy (2019[63]) on page 55 also finds that patented innovations tend to be highly correlated, across countries, with his index of social-democratic capitalism, one component of which is social spending as a share of GDP.

5 Wealth inequality in the long run

Guido Alfani, Bocconi University, Dondena Centre and Stone Center on Socio-Economic Inequality
Sonia Schifano, University of Luxembourg

This chapter provides an overview of long-term changes in wealth inequality based on the Gini index and the wealth share of the richest 10%.[1] The chapter relies on current databases such as the *World Inequality Database (WID)* but expands the time series for a range of countries (back to 1820 in the case of France, Italy, Sweden, the United Kingdom and the United States) and by producing new estimates. Our data confirm the general picture of a creeping increase in inequality during the 19th century, followed by declines from the onset of World War I (1914) until the 1960s and, in many (but not all) countries, a new tendency towards higher inequality since the 1970s. The correlation of wealth inequality with per capita GDP is found to be quite weak and not always positive, implying that higher wealth inequality cannot be considered a simple side effect of economic growth.

Introduction

In recent years studies on historical developments in economic inequality have flourished. Much of this new research has focused on income inequality during the 19th and 20th centuries, continuing a trend that had been triggered by Simon Kuznets's seminal article (Kuznets, 1955[1]). This greater attention to distribution is surely connected to the tendency towards the increase in income inequality in past decades, as well as to the Great Recession that started in 2007, whose effects are still felt today – among them a much-heightened perception of inequality as a social and political issue. One factor characterising this renewed interest in inequality is that it has not been limited to income, as was almost invariably the case in the past, but has involved wealth, which in recent decades has shown a tendency towards becoming ever more concentrated (Atkinson, 1980[2]; Di Matteo, 2018[3]; Lindert and Williamson, 2016[4]; Piketty, Postel-Vinay and Rosenthal, 2014[5]; Piketty, 2014[6]; Roine and Waldenström, 2015[7]). An even greater attention to wealth inequality has been paid by studies focusing on preindustrial times (as a rule of thumb, pre-1800), which – for a long period – had remained uncharted territory for inequality studies (Alfani, 2015[8]; Alfani, 2017[9]; Alfani and Ammannati, 2017[10]; Alfani and Di Tullio, 2019[11]; Bengtsson et al., 2018[12]; Canbakal, Filiztekin and Pamuk, 2018[13]; Nicolini and Ramos Palencia, 2016[14]). Indeed, for some areas of the world (especially Europe) we now probably know more about wealth inequality in preindustrial times than during the 19th or even most of the 20th century.

Overall, recent findings about long-run inequality trends, of both income and wealth, have significantly changed how we look not only at historical distributive dynamics, but also at current inequality levels and tendencies. For example, they have led many scholars to explicitly reject Kuznets's hypothesis about the existence of an innate tendency for economic inequality to decline after the achievement of a certain level of development (Lindert, 2000[15]), as well as to question all sorts of mono-causal explanations of changes in inequality (Alfani and Ryckbosch, 2016[16]; Alfani and Di Tullio, 2019[11]; Lindert and Williamson, 2016[4]). In this chapter, some of the insights from these studies will be used to better understand the nature of the inequality trends we describe.

One important reason why tendencies in wealth inequality have become crucial to current studies on distributional dynamics is the awareness that a high concentration of material and financial assets could hinder the way in which modern societies are supposed to function. No one has been more influential in making this point than Thomas Piketty (Piketty, 2014[6]; 2015[17]). In Piketty's view, as long as the rate of return on capital is higher than the growth rate of national income, and as long as wealth is highly inheritable, inequality (of both income and wealth) will continue to increase. But this might prove incompatible with the principles of social justice and equal access to public institutions, which are the basis of contemporary democracies. More generally, there is a growing concern that excessively high wealth inequality may compromise the openness of societies and have negative consequences for long-term social and economic development, for example by making upward mobility more difficult (OECD, 2015[18]).

This chapter presents the available information about wealth inequality in the last two centuries. Unfortunately, for most areas of the world this information continues to be much less systematic than what is available for income inequality. Subsequently, we describe the general trends of wealth inequality, and we provide a brief account of some of the current debates on the drivers of changes in inequality in the long run of history.

Description of the concepts used

The analyses presented in this chapter refer to the distribution of household net wealth. We measure inequality via two standard indicators: the Gini index and the wealth share of the richest 10%. Both choices require further discussion.

At the outset, we have to acknowledge that there is currently no universal consensus about what we should consider when measuring "wealth". Indeed, if we wanted to measure the distribution of all economic resources, we should take into account human capital (Davies and Shorrocks, 2000[19]), or the sum of what are sometimes referred to as "embodied" and "relational" wealth (embodied wealth includes knowledge and skills as well as physical characteristics like strength) (Borgerhoff Mulder et al., 2009[20]). Even if we restricted the analysis to non-human capital, there would still be some doubt about which components of wealth we should include. It might seem straightforward to define wealth as "net worth", i.e. the monetary value of assets (both real property and financial claims) minus the value of liabilities. But this leads to questions about the inclusion of intangible assets that could not easily be sold, for example, pension rights or entitlements to future government transfers. In practice, applied research on wealth usually focuses on net "marketable wealth" (Davies and Shorrocks, 2000, p. 60[19]). A final point to consider is that in the history of economic thought, two different ideas have confronted each other, one considering wealth as a set of tangible possessions (the notion of net wealth is related to this view), the other focusing not on tangibles per se, but on the pleasure or "utility" that they generate (Heilbroner, 2008[21]). A different way of raising not-entirely-dissimilar concerns comes from anthropology, which has suggested that wealth and affluence are not universal concepts, and that the possession of assets has different meanings and consequences in different societies (Sahlins, 1974[22]).

This being said, the most frequently used concept is that of net (marketable) wealth, which also corresponds to by far the largest part of the currently available information. This is the kind of information provided by the *World Inequality Database* (*WID*), for example. Consequently, the data presented here have to be understood as referring to net wealth (although for earlier epochs and for some areas we often have more limited information, hence some adjustments were required). This definition of wealth is also in line with OECD recommendations in this field (OECD, 2013[23]). Note that our measures refer, ideally, to wealth at the household level, without adjustments for household size, again reflecting data availability and common practice. However, for some countries and periods we rely on individual-level distributions (adults) due to the absence of information about household distributions. Based on the available literature – see the discussion in (Roine and Waldenström, 2015, p. 513[7]) – we expect that measures of wealth inequality at the individual level differ only slightly from those calculated at the household level.

Regarding the inequality measures, the Gini index of concentration is an obvious choice, not only because of its vast application in the social sciences, but also because a significant amount of the information we have available is expressed as a Gini value. The index value varies between 0 (perfect equality, i.e. all households or individuals have the same wealth) and 1 (perfect inequality, i.e. one household or individual owns all the wealth, the others have none). However, the Gini index is often multiplied by 100 for more effective presentation, as we do here. Note that this index can in principle be applied in exactly the same way to wealth and to income – hence we refer to Chapter 11 from (van Zanden et al., 2014[24]), dedicated to income inequality, for further discussion of the index's properties. When applied to wealth, the Gini index tends to be higher than when applied to income, due to the empirical tendency for wealth to be more concentrated than income, see (Davies and Shorrocks, 2000, p. 60[19]).

A well-known limitation of the Gini index is that the same value can describe very different distributions. One practical way of solving this problem is to couple the Gini coefficient with percentiles, which can capture changes in important parts of the overall distribution. In this chapter, we do this by including information on the wealth share of the richest part of the population. But there are other important reasons for considering this additional indicator. One is practical: for a significant number of years and countries, we have information only about the share of the richest 10%. Hence, using information about the top of the distribution allows to improve coverage across time and across countries. It has also been argued that the wealth share of the richest should be used because it is easier to understand even without specific training – which is clearly a desirable feature, given the socially sensitive nature of many current debates on inequality and the interest in the topic found across ample strata of civil society. This choice is further justified by the fact that there is a clear empirical tendency for changes in the wealth share of the rich to go hand-in-hand with changes in general inequality as measured by the Gini index (in other words, the Gini index increases/declines as the wealth share of the rich increases/declines). This empirical tendency has been confirmed by studies of both contemporary and historical societies (Alfani, 2015[8]; Alfani, 2017[9]; Atkinson, Piketty and Saez, 2011[25]; Alvaredo et al., 2013[26]; Roine and Waldenström, 2015[7]). Finally, it has been argued that in studies of wealth inequality the top shares are to be preferred to the Gini because wealth (unlike income) can assume negative values for indebted households at the bottom of the distribution – indeed, in the presence of negative wealth, the Gini value could theoretically exceed 1 (or 100 in our preferred format) (Balestra and Tonkin, 2018[27]). This, however, does not happen across our database, either because the original sources used do not allow for negative wealth (this would be the case of most property tax records and estate data sources), or because the research teams producing the data we rely upon set the minimum value of net household wealth at zero. This is the case for the *World Inequality Database* (Alvaredo, F.; A. Atkinson and L. Chancel, 2016[28]), in particular Annex Table A1, which is the main source of the estimates included in this chapter. This consideration enhances the case for coupling the Gini index with top shares or other indicators with similar properties.

The existence of a strong empirical correlation between the share of the richest and the Gini index presents the additional advantage of offering a simple and relatively reliable way of estimating one from the other when complete information is unavailable – which unfortunately is very often the case, as discussed in the next section. We made use of this procedure, with some adaptation, as the preferred way to complete gaps in the database of Gini indices or of the share of the top 10%.[2] This method is preferable to other options, like interpolation, because it makes use of at least some information about the actual distribution in a given country and year, although we had to resort to interpolation when neither the Gini nor the share of the top 10% were available. The estimated relationship between the Gini and the share of the richest 10% is the following.[3]

$$Gini = 0.30 + 0.72 * Top10$$

The main objective of this chapter is to provide a long-term perspective on wealth inequality in individual countries. However, we also provide some hints about the global distribution of household wealth (i.e. that obtained by considering all the households in the world as living in the same country). This task is performed with a method analogous to that employed in Chapter 11 of (van Zanden et al., 2014[24]), which provides additional details. Applying this method requires some assumptions. In particular, we rely upon available estimates of the average net wealth at the country level, and assume log-normality to estimate wealth distributions starting from the Gini coefficient and the net private wealth for each country and year. Global wealth distributions are then obtained by aggregating country level distributions for each year. Unfortunately, we could produce global estimates only for the more recent years, given the sparse nature of the available information about wealth inequality (even including new estimates and “guess-estimates”), as detailed in the next section. We do however provide estimates of overall wealth inequality since 1850 across a small group of Western countries (France, Sweden, the United Kingdom and the United States).

Historical sources

The information currently available about the distribution of wealth is much sparser than that relating to income. In order to build as comprehensive a database as possible, we needed to rely upon disparate sources as well as to produce some new estimates.

Most of the available historical evidence about wealth concentration comes from wealth and estate taxation statistics. Tax sources also allow building the longest time series (Roine and Waldenström, 2015[7]). However, they also have a number of limitations. A particularly important one is that often they include only the richest part of the population, which is why many applied studies have relied upon the wealth share of the richest to summarise inequality levels and trends. But this coverage problem does not always occur. For example, the Italian *estimi* (property tax records used until the turn of the 19th century) included at least 90-95% of the overall population (Alfani, 2015[8]; Alfani and Di Tullio, 2019[11]). Similar sources exist for other parts of Europe (for example southern France, north-eastern Spain and parts of Germany). Unfortunately, this kind of information is not widely available for later periods due to changes in the fiscal systems. In this chapter, these sources have been used only to estimate wealth inequality in Italy in the first part of the 19th century.

A general concern with all fiscal sources is the level of tax evasion and avoidance. Although it is usually very difficult to quantify this, it should be noted that at least some components of wealth – like real estate – are relatively difficult to hide, or to misevaluate dramatically. This being said, the opportunity for wealth tax evasion and avoidance probably tended to increase during the two centuries covered by this chapter, for example because in the current global economy it has become easier to hide movable wealth by means of tax havens or otherwise (Roine and Waldenström, 2015[7]).

Alternative sources feeding the data used in this chapter include estate data (which usually do not cover the entire population, but result from researchers' attempts at building representative samples) as well as probate inventories (data from the judicial process through which a will is accepted as the legal testament of the deceased). The latter are relatively popular in historical research, due to the vast archival availability of probates since early epochs. A useful feature of probates is that they exist for non-European areas, too, like the Ottoman lands in Anatolia and the Levant (Coşgel and Ergene, 2012[29]; Canbakal, Filiztekin and Pamuk, 2018[13]). However, due to the vast amount of archival research needed to produce national samples of probates, so far it has been possible to produce comprehensive and reliable wealth distributions based on them only for a few countries and years. Moreover, and most importantly, they do present specific challenges related to sampling – as, again, they often reflect the richest part of the population far better than the poorest, and require assumptions about mortality rates by age, sex and socio-economic status to move from the distribution of wealth of those dying in a given period to the hypothetical distributions characterising living cohorts in earlier periods (Bengtsson et al., 2018[12]; Lindert, 1986[30]; Roine and Waldenström, 2015[7]).

The main databases we drew information from provide us with standardised measures that usually incorporate attempts at correcting the limitations in the original sources described above. In particular, we relied heavily upon the *World Inequality Database* (*WID*) (WID.world[31]), from which comes most of the information we have for the period 1900-2000 (recent releases of the database have greatly increased the coverage of wealth inequality, although much more information is available about income). For 2010, we have also collected information from the *Global Wealth Databooks* (*GWD*) published by Credit Suisse's Research Institute (Credit Suisse[32]). The *Databooks* include estimates for a number of countries ranging from 20 to 30 according to the year, based on a combination of national accounts and household survey data supplemented by a variety of estimation methods. For this year and for a large subset of OECD countries, high-quality data from the *OECD Wealth Distribution Database* (*WDD*) (OECD[33]), based on a combination of household surveys and (for some countries) administrative data, are also available. We preferred relying upon the *GWD* for later years due to the much larger number of countries covered as well the higher (albeit far from perfect) consistency with the estimation methods used by the *WID*. However, we also used the *WDD* as a means of external validation of our findings. For earlier periods, we relied upon a greater variety of sources, reflecting the specificities of the archival and statistical information produced locally as well as the degree of advancement of historical research on inequality and related topics.

Only for very few countries covered by this chapter is information about wealth inequality available as far back as 1820. Especially for Africa and Asia, the estimates tend to begin only in the second half of the 20th century (the earliest estimate we have for these areas refers to Japan in 1870). Estimates extending back to 1820 are available only for five countries in Western Europe and the Western Offshoots: France, the United Kingdom, Italy, Sweden and the United States. The sources and information used in this chapter for these five countries are detailed below.

France is the country whose wealth distribution during the 19th century has been researched most thoroughly. This is due especially to the research conducted by (Piketty, Postel-Vinay and Rosenthal, 2006[34]; Piketty, Postel-Vinay and Rosenthal, 2014[5]; Piketty, 2014[6]). Indeed, of all the countries covered by this chapter, France is the only one for which complete series of wealth Gini indices and top shares are readily available, without the need for additional estimates, for the entire period 1820-2010. We used the standardised series recently made available by the *WID*.

For the United Kingdom, we made use of estimates of the share of the richest 10% produced by (Lindert, 1986[30]) for selected years during the late 18th and the 19th century, with some additional estimates used to fill in the gaps. Lindert's estimates are based on data from probates, coupled with various other sources to account properly for some specific components of wealth, like real estate. From 1900, the *WID* provides fairly systematic data about the wealth share of the richest 10%. Information about UK Gini index is much scarcer, and it is only from the 1970s that a series of Gini indices is available, from (Alvaredo, Atkinson and Morelli, 2018[35]). This led us to produce a large number of new estimates by means of the correlation procedure described in the earlier section.

For Italy, wealth inequality measures for 1820 have been produced using information from studies of long-term inequality trends in preindustrial times (based on fiscal sources and particularly on property tax records) in a variety of Italian pre-unification states – for example (Alfani, 2015[8]; Alfani, 2017[9]; Alfani and Di Tullio, 2019[11]; Alfani and Ryckbosch, 2016[16]). Information for later periods is scattered (Schneider, Pottenger and King, 2016[36]; Brandolini et al., 2004[37]; Davies et al., 2010[38]; Credit Suisse, 2017[39]; Gabbuti and Morelli, 2019[40]); as the *WID,* does not provide this information for Italy, hence many new and rather tentative estimates have been produced.

For Sweden during the 19th century, Bengtsson et al. (2018[12]) produced estimates of both the Gini index and the wealth share of the richest 10% for selected years by using data from probate inventories. For the 20th century, the same was done by Roine and Waldenström (2015[7]) regarding the share of the richest 10%, while Schneider, Pottenger and King (2016[36]) provided Gini estimates for 2000. Also, in this case, many additional estimates were required to fill in the gaps.

Finally, for the United States, very tentative estimates of top wealth shares during the 19th century have been produced by interpolation between an observation for 1774 based on data provided by Lindert (2000[15]), one for 1870 provided by Sutch (2016[41]) and one for 1890 by Roine and Waldenström (2015[7]); Schneider, Pottenger and King (2016[36]) provide a Gini index for 1870. For the 20th and early 21st centuries, the data we use for the share of the richest 10% and for the Gini indices mostly come from the *WID*. New tentative estimates have been produced to fill in the gaps.

Table 5.1 and Table 5.2 provide an overview of the kind of information used in this chapter, specifying the original data source and the estimation method used. As far as the share of the top 10% is concerned, the *WID* provided 21% of the data used in our analysis, while the publicly available database by Roine and Waldenström (2015[7]) furnished another 13% of the data. Regarding the Gini indexes, 33% of the observations are new estimates obtained by correlation with the share of the top 10% (as discussed earlier). The *WID* and Schneider, Pottenger and King (2016[36]) together supplied 18% of the other data.

For both the Gini indices and the share of the top 10%, 34% of observations had to be produced by simple interpolation, although sometimes between very close dates (for example, the estimate for Italy in 1990 is obtained by interpolating the observations for 1989 and 1991).

Unfortunately, as is quite clear, good-quality estimates of wealth inequality are only sparsely available, hence to provide at least an impression of long-term trends we had to rely heavily on new and often very tentative estimates. Also note that for some areas and periods more than one estimate is available, hence we had to make informed decisions about which one to rely upon. Beyond the general guidelines described in this section, we aimed to ensure consistency in the reconstructed series, to avoid reporting clearly artificial trends, and to offer a picture that is overall consistent with the most recent literature on single countries or areas.

Table 5.1. Estimates of wealth inequality used in this chapter by source and benchmarked year, 1820-2010, Gini coefficients

Number of countries covered

	All	World Inequality Database	Schneider, Pottenger and King (2016[36])	New estimate (obtained by correlation with share of top 10%)	New estimate (obtained by interpolation)	Other sources
1820	5	1		1	3	
1830	5	1			4	
1840	5	1			4	
1850	5	1			3	1
1860	5	1			4	
1870	6	1	1	1	2	
1880	6	1			5	1
1890	7	1	1	2	3	
1900	7	1	1	2	2	1
1910	11	1	1	3	5	1
1920	12	1		5	6	
1930	12	1		5	6	
1940	12	1		4	7	
1950	12	1		7	4	
1960	13	2		6	5	
1970	17	2	1	6	3	5
1980	18	3		4	6	5
1990	21	4	1	4	5	7
2000	25	4	3	3	2	13
2010	36	4	1	27	2	2
Total	240	33	10	80	81	36

Table 5.2. Estimates of the wealth share of the richest 10% used in this chapter by source and benchmark year, 1820-2010

Number of countries covered

	All	World Inequality Database	Credit Suisse Global Wealth Databooks	New estimate (obtained by correlation with Gini coefficient)	New estimate (obtained by interpolation)	Roine and Waldenström (2015[7])	Other literature
1820	5	1			3		1
1830	5	1			4		
1840	5	1			4		
1850	5	1			3		1
1860	5	1			4		
1870	6	1			2		2
1880	6	1		1	5		
1890	7	1			3	1	2
1900	7	2			2		3
1910	11	3		1	5		2
1920	12	3			6	3	
1930	12	3			6	3	
1940	12	3			7	1	1
1950	12	3			4	3	2
1960	13	3			5	3	2
1970	17	3		4	3	4	3
1980	18	4		2	6	2	4
1990	21	5		5	5	4	2
2000	25	5		5	2	4	9
2010	36	5	23		2	2	4
Total	240	50	23	18	81	30	38

Data quality

In Table 5.3 and Table 5.4 we provide an assessment of the quality of the data used in this chapter. So far, the study of wealth inequality has not produced a unique database containing comparable data in terms of both unit of analysis and methodologies used to gather them. Studies of wealth inequality in the past have relied upon national sources that provide different information about wealth distributions. Most of them are based on wealth tax statistics, estate records, probates, and for the last decades, household surveys. Moreover, given the presence of numerous gaps between the readily available measures, we had to produce many new estimates. This is why, according to the parameters defined in Chapter 1 of (van Zanden et al., 2014[24]), in most instances data quality has to be classified as "low" or "moderate". Note that the situation is better regarding measures of the share of the richest 10% compared to Gini indexes. To a large degree this is due to the *World Inequality Database* (*WID*), which is the most comprehensive database in this field of enquiry. For many countries and years, the *WID* provides us with good estimates of the wealth share of the richest percentiles. Other sources have been used for specific countries, as detailed previously, as well as for the 19th century, which is not usually covered by the *WID* (France is the notable exception).

Collecting estimates of the Gini coefficients for the period under study was more difficult. Indeed, earlier research has focused mostly on the top percentiles, also because it is much easier to gather data on the right-hand side of the wealth distribution. However, there is strong empirical evidence that the wealth share of the richest 10% and the Gini index are strongly correlated. Indeed, for the sub-set of countries and years for which we had information about both, the Pearson correlation coefficient is 0.94. This allowed us to present estimates of the Gini coefficient starting from values of the top 10% whenever we lacked a direct measure of the former. Less often we predicted the top share from the Gini coefficient. This is a simple estimation method which, in the absence of any useful additional information, seems preferable to more complex procedures. A more detailed description of our methods is provided in the section "Description of the Concepts Used". Furthermore, we have systematically assigned a lower data quality to predicted measures than to original measures.

Table 5.3. Quality of estimates of the Gini coefficient of wealth inequality by region and benchmark year, 1820-2010

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Middle East and North Africa	Sub-Saharan Africa	East Asia	South and Southeast Asia
1820	3		4					
1870	3		4		3			
1910	3		4				4	
1960	3		4				4	
1990	3	4	2	3			3	2
2010	3	3	2	3	2	4	3	3

Note: 1: High quality; 2: Moderate quality; 3: Low quality; 4: Tentative estimates. The table reports rounded mean values per world region/benchmarked year.

Table 5.4. Quality of estimates of the wealth share of the reachest 10% of the population by region and benchmark year, 1820-2010

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Middle East and North Africa	Sub-Saharan Africa	East Asia	South and Southeast Asia
1820	3		4					
1870	3		4		3			
1910	3		4				4	
1960	2		4				4	
1990	2	4	2	3			3	2
2010	2	3	2	3	2	4	3	3

Note: 1: High quality; 2: Moderate quality; 3: Low quality; 4: Tentative estimates. The table reports rounded mean values per world region/benchmarked year.

Main highlights of historical trends in wealth inequality

This section provides an overview of long-term tendencies in wealth inequality, focusing on developments within countries and underlying commonalities and differences across world regions. It also provides a tentative picture of global wealth inequality, generally restricted to the most recent periods given the limited availability of information about wealth distribution in many areas of the world until the final decades of the 20th century.

Developments in individual countries

At 1820, we have information about wealth inequality for only five of the countries included in Table 5.5 and Table 5.6, and all of them are in the West: France, the United Kingdom, Italy and Sweden in Western

Europe, plus the United States in the Western Offshoots. In this year, Gini indices range from 88 to 90 in Europe, while the United States was significantly more "egalitarian" with a Gini of 76. The wealth share of the richest 10% marks an even greater distance, being in the 81-85% range in Europe, and 62% in the United States. This difference in wealth inequality between Western European countries and the United States in the 19th century is well documented in historical research (Lindert and Williamson, 2016[4]). The gap between Europe and the United States tended to narrow during the 19th century, without disappearing entirely. By 1900, with a Gini of 84, the United States was still slightly more egalitarian than any European country with available data, although only marginally so compared to Italy where the Gini was 83. For that year, we also have information about Canada which, with a Gini of 89, was basically in line with European countries. The only non-Western country for which we have data in 1900, Japan, is significantly more egalitarian, with a Gini coefficient of 53 and a wealth share of the richest 10% of 36%, well below the 75-93% range characterising Western Europe and the Western Offshoots as a whole.

Importantly, during the 19th century we find an overall tendency for wealth inequality to increase. This is clearer in the United Kingdom, Sweden and the United States compared to France and Italy, whose trends are quite flat. This finding could be taken to fit the narrative of the "Kuznets curve" – which, although originally related to income, can reasonably be applied to wealth as well, see pp. 215-219 in Lindert (1991[42]) – with growing inequality seen as the consequence of the Industrial Revolution. However, recent research focused on longer time periods casts doubt on this, as this phase of increasing inequality seems to be the continuation of a multi-secular phase of wealth inequality growth that started around the 15th century, that is, after a phase of declining inequality (of both wealth and income) triggered by the Black Death plague of 1347-1352 (Alfani, 2017[9]; Alfani, 2019[43]; Alfani and Di Tullio, 2019[11]; Alfani and Ryckbosch, 2016[16]).

Wealth inequality continued to grow during the first decade of the 20th century. In Western Europe and its Offshoots, this tendency was halted by World Wars I and II and by the troubled times in between. The multi-faceted impact of war and large-scale societal and economic disruption has been variously presented as the direct cause of the inequality decline characterising the period from 1914 to 1945 (Piketty, 2014[6]; Scheidel, 2017[44]). Indeed, in each country for which we have information about wealth inequality in both 1910 and 1950, we find a decline over the period. For example in Australia, whose first available observation dates to 1910, the Gini coefficient declines sharply, from 86 to 63, and the share of the richest 10% falls from 75% to 64%. The tendency for a decline in wealth inequality continued for about two decades after 1950, probably (at least in advanced countries) also as a consequence of fiscal redistribution and the associated development of the welfare state: think about, for example, the spread of public programmes supporting home ownership (Atkinson, Piketty and Saez, 2011[25]; Alvaredo et al., 2013[26]; Piketty, 2014[6]). From the late 1970s – the beginning of a phase of fiscal reform leading to lower progressiveness in income taxation and lower inheritance taxes, as well as of deregulation in the financial sector (Alfani, 2019[43]; Piketty, 2015[17]) – we find in most countries signs of a resumption in the tendency for wealth inequality to grow, which continues to this day. At least to some degree, this could also be related to demographic tendencies in the West. Particularly relevant are increases in life expectancy: as wealth is positively correlated with age, due to higher cumulated savings, an aging population tends to be characterised by a more dispersed, hence more unequal, wealth distribution – for a discussion of this general problem, which is also relevant for income, see Paglin (1975[45]). According to some studies, population aging is a contributing factor of wealth inequality growth, for example in Germany (Ihle and Siebert-Meyerhoff, 2017[46]), although this effect can be expected to wane and finally disappear as Western populations grow even older.

From 1970 onwards, we can compare inequality levels between more world areas. At that date, the Gini index in Western Europe ranged from 65 to 77; in the Western Offshoots from 52 to 84 (72-84 if we restrict the observation to North America); in Latin America and the Caribbean from 61 to 86; and in Asia, 52, where the only estimate available refers to Japan. However for China in 1980 our estimate of the Gini index is 54, suggesting that overall East Asia was relatively egalitarian (differences in political regimes between China and Japan do not seem to be associated with significantly different inequality levels).

The coverage of our database improves further in subsequent decades. In 2000, the range of Gini levels was 57-89 in Western Europe; 75 in the Russian Federation, which is the only Eastern European country for which we have information; 62-83 in North America and Australia (the latter being, again, the most egalitarian of the group); 74-78 in Latin America and the Caribbean; 72 in Turkey; 55-59 in East Asia (Japan and China respectively); and 67-76 in South and Southeast Asia.

By 2010, the United States stands out as the least egalitarian country among those included in Table 5.5 and Table 5.6 , with a Gini coefficient of 88 and a wealth share of the richest 10% of 73%. An interesting feature of American history is that, starting from being the most egalitarian of Western societies in the 19th century, it has become exceptionally unequal. This change mainly reflected developments in the interwar period and in the 20-30 years following World War II, when the United States were characterised by a lower decline in inequality than that found in many European countries or Canada. It is also interesting to compare the case of the United States to that of France, which followed much the opposite path, turning from being one of the most unequal Western European countries in the first half of the 19th century to being one of the most egalitarian Western countries today – or at least, one where the inequality increase during the last 40 or 50 years has been exceptionally modest (by 2010, with a Gini of 70, France was as unequal as in 1970. The share of the richest 10% even declined marginally, from 58% to 56%).

Table 5.5. Gini coefficient of wealth inequality in selected countries, 1820-2010

	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1820	[90]		88		89		[90]					[76]													
1830	[90]		87		[88]		[89]					[78]													
1840	[90]		86		[88]		[88]					[79]													
1850	[90]		90		[87]		87					[80]													
1860	[90]		88		[87]		[90]					[82]													
1870	90		87		[86]		[93]					83										53			
1880	[92]		87		[86]		[92]					[83]										[53]			
1890	[94]		88		86		[92]				91	[82]										[53]			
1900	97		89		83		91				89	[84]										[53]			
1910	96		89		81		[94]			86	[87]	[87]										[53]			
1920	94		88		[79]		[96]			[80]	[84]	[86]										[52]			
1930	92		87		[77]		[95]			[74]	[82]	[91]										[52]			
1940	91		82		[74]		[90]			[69]	[79]	[86]										[52]			
1950	88		82		[71]		[86]			[63]	77	[79]										[52]			
1960	81		80		[68]		[79]			58	70	84										[52]			
1970	77		70	75	[65]		[72]			52	72	84	61	84	86							[52]			
1980	65		66	70	[62]		69			66	69	81	[66]	85	84						54	52			
1990	64	78	65	62	58		72		67	64	69	79	[70]	85	83						54	56	66		
2000	71	79	70	67	61	57	89		75	62	69	83	75	78	74		72				59	55	67	76	71
2010	69	81	70	73	63	67	73	58	78	61	71	88		82						82	73	65	80	78	78

Note: For an assessment of data quality, see Table 5.3 and Table 5.4. The complete database, including information about the sources used for each country/year, can be downloaded from http://didattica.unibocconi.eu/Alfani_database.
Values in brackets [ ] indicate very tentative data.

StatLink https://stat.link/swhjlm

Table 5.6. Wealth share of the richest 10% of the population in selected countries, 1820-2010

	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1820	[83]		85		81		83					[62]													
1830	[84]		82		[80]		[81]					[63]													
1840	[84]		80		[79]		[80]					[63]													
1850	[84]		87		[78]		79					[64]													
1860	[84]		83		[76]		83					[65]													
1870	84		81		[75]		87					65										36			
1880	[86]		83		[74]		87					69										[36]			
1890	[89]		83		73		87				84	72										[36]			
1900	93		84		76		86				82	75										[36]			
1910	92		85		82		88			75	[79]	78										[35]			
1920	88		82		[76]		92			[72]	[75]	78										[35]			
1930	86		80		[71]		89			[70]	[72]	85										[35]			
1940	84		72		[66]		83			[67]	[68]	77										[35]			
1950	80		72		[61]		77			[64]	65	68										[35]			
1960	71		71		[55]		68			[61]	55	70										[35]			
1970	64		58	62	[50]		58			58	56	71	46	73	76							[35]			
1980	52		52	49	[45]		55			50	56	65	[51]	75	74						41	35			
1990	46	64	50	41	39		59		53	49	56	64	[57]	75	72						41	40	52		
2000	51	58	57	44	49	42	60		65	45	56	69	62	67	61		59				48	39	53	65	58
2010	54	62	56	59	46	51	59	37	66	45	57	73		72						72	63	48	69	67	67

Note: For an assessment of data quality, see Table 5.3 and Table 5.4. The complete database, including information about the sources used for each country/year, can be downloaded from: http://didattica.unibocconi.eu/Alfani_database.
Values in brackets [ ] indicate very tentative data.

StatLink https://stat.link/8okbt9

Regional comparisons

The country-level data reported above allow for some regional comparisons, as shown in Table 5.7 (for all world regions except sub-Saharan Africa and the Middle East and North Africa). The measures shown in the table do not reflect differences in net wealth between countries nor country population sizes, but are simple averages of Gini coefficients. Table 5.7, which we restricted to 1950-2010 to avoid presenting measures based on few country observations, shows the final phase of decline in within-country wealth inequality in Western Europe and in the Western Offshoots, from the 1950s to the 1970s. This decline, which is usually associated with the development of the welfare state (see discussion above), has been followed in the last 40-50 years by stagnation and then an upward creep (which was more clearly visible in the country-level data presented in Table 5.5 and Table 5.6).

A period of declining inequality seems to have been followed by one of rising inequality also in East Asia, where average inequality today is very close to that of Western Europe. East Asia and Western Europe are the only two world regions that, in 2010, were below the world average of 72. The Gini index for the Latin America and Caribbean region matches the average perfectly, while that for the Western Offshoots (at 74) is just above it. South and Southeast Asia leads the ranking, with a Gini index of 82 (although excluding Singapore drops this to 79). This picture does not change much if we look at 2000 (when database coverage is highest) instead of 2010, with the exception of the Latin America and Caribbean region, which jumps to the top of the ranking. The very high wealth inequality in the Latin America and

Caribbean area in 2000 mirrors its exceptionally high income inequality (Milanovic, 2005[47]; Prados de la Escosura, 2007[48]).

Table 5.7. Wealth inequality across world regions, 1950-2010

Gini coefficients, unweighted averages

	Western Europe	Western Offshoots	Latin America and Caribbean	East Asia	South and Southeast Asia	World
1950	84	73		52		79
1960	80	70		52		75
1970	76	69	77	52		74
1980	72	72	78	53		71
1990	70	71	79	55	66	70
2000	72	71	76	57	71	71
2010	71	74	72	69	82	72

Note: For an assessment of data quality, see Table 5.3 and Table 5.4. This table includes all the countries for which we have information about wealth inequality, so this covers a larger group compared to Table 5.5 (hence some slight differences in the value of "World" averages between this table and Table 5.8).

StatLink https://stat.link/p6f9m7

A first glance at the global wealth distribution

We can provide a global perspective on wealth distribution, by observing the inequality levels corresponding to all the world's households taken together, i.e. as if the whole world were considered as a single country, similarly to what is done for income inequality in Chapter 11 of van Zanden et al. (2014[24]).[4] Unfortunately, this is not easy to achieve for wealth, given that we do not have complete information for any year, either about inequality levels or about per capita net wealth per country. Consequently, we focus on 2000, which is the year with the largest country coverage. Some additional assumptions were made to obtain estimates of the global wealth distribution in 2000; in particular, we use the information about inequality for South Africa in 2010 to represent the whole of sub-Saharan Africa in 2000, which clearly does not reflect perfectly the overall situation of the area.

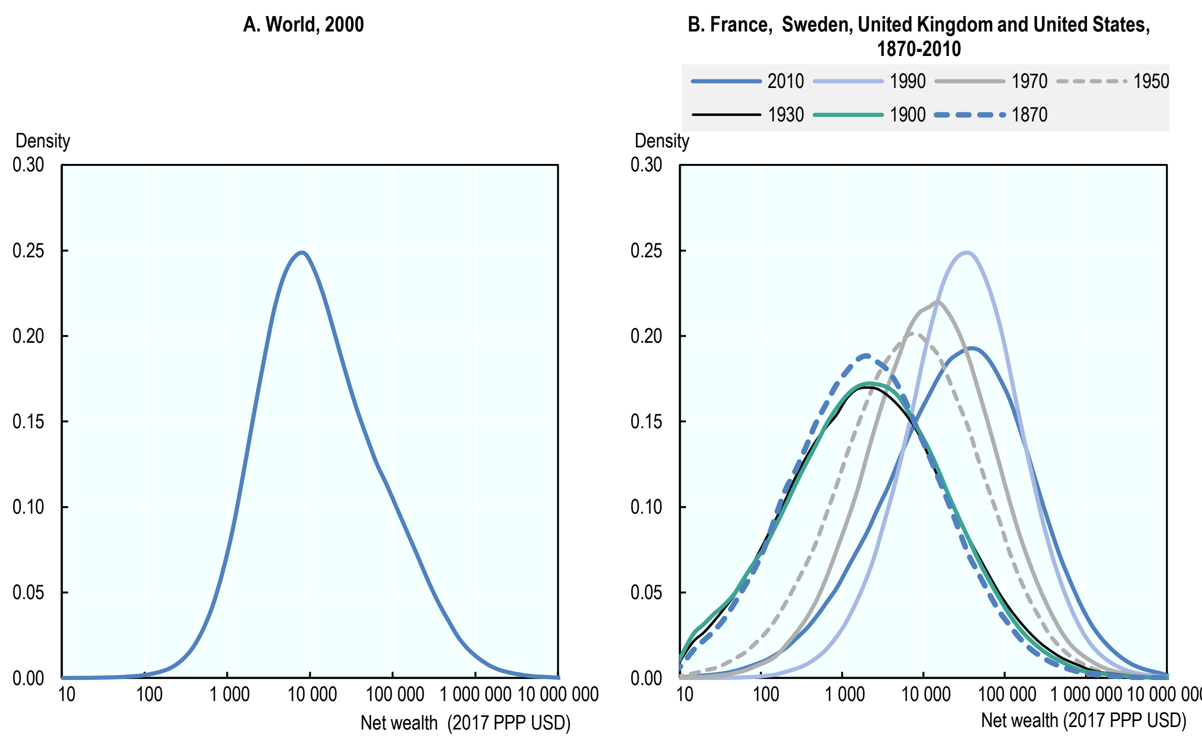
Our estimate of the global wealth distribution in 2000 is shown in Figure 5.1, panel A. As can be seen, this global distribution is log-normal (i.e. it assumes a normal shape when the x axis is expressed in a logarithmic scale), with the mode placed at a per capita net wealth of about USD 8 700 (measured as US dollars PPP 2017). The corresponding Gini level is 82, and the share of the richest 10% is 73%. Our estimate is quite a bit lower than the Gini of 93 (with a share of the top 10% of 89%) estimated for the same year by Davies, Lluberas and Shorrocks (2017[49]) (see in particular pp. 754-5), due to differences in the countries covered and in the methods used. In particular, while we apply to wealth the method used for income in Chapter 11 of van Zanden et al (2014[24]) in order to maximise the internal comparability of measures, Davies and Shorrocks (2000[19]) follow a different approach that adjusts the right tail of the distribution based on the "rich lists" published by Forbes and others.

Unlike what was shown for the income distribution in van Zanden et al (2014[24]), our wealth distribution for 2000 is unimodal. As the coverage of our database is quite imperfect, we cannot identify any time trend. Although in Table 5.8 we include measures of "Global Inequality" for the period 1960-2010, changes in the Gini index are also affected by changes in the country coverage of our database.

For the period 1870-1950, however, we have information about both the Gini index and average net wealth for four Western countries (France, Sweden, the United Kingdom and the United States). If we restrict the analysis to this group, we can indeed detect a time trend for 1870-2010. Figure 5.1, panel B, shows that, in all years, the shape of the distribution is log-normal. The mode of the distribution shifts to the right through time (with the sole exception of the troubled period between 1910 and 1930), reflecting increases in average net wealth, as well as a general spreading out of wealth across the overall population of these countries. This is reflected by an estimated Gini index that declines from 88 in 1870, to 82 in 1950, and 76 in 1990.[5] Lower wealth inequality in the second half of the 20th century goes hand in hand with convergence in net wealth across the four countries (see below). However, the tendency towards lower wealth inequality within this group of rich Western countries comes to a halt and reverses in the last thirty years or so: after the minimum reached around 1990, the Gini index increases to 80 by 2000 and to 84 by 2010.[6]

Convergence in average net wealth between countries can be analysed referring to the concept of "international inequality", which is particularly apt at detecting processes of convergence/divergence between nations (Milanovic, 2005[47]). Note that the concepts of "global" and "international" inequality differ because of the methods, not because of the territorial coverage, which in Table 5.8 is kept constant across measures. In the unweighted version, international wealth inequality is calculated on the average net wealth of each country, without regard for relative population size and without taking into account within-country inequality. As shown in Table 5.8, international wealth inequality among the four Western countries declines steadily from 1870 to 2010, until it reaches almost perfect "equality". The same trend is found when looking at the population-weighted version of the index. If we apply the same measure to all countries for which we have information we get more erratic results which, again, simply reflects sharp changes in the coverage of the database across time – hence we report "world" measures from 1960 only, which are more reliable due to better territorial coverage.

Figure 5.1. Global wealth distributions



Note: For an assessment of data quality, see Table 5.2, Table 5.3 and Table 5.4.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/7h190r

The same problem with coverage applies to one final measure: the simple average of Gini levels. While in Table 5.8 the countries considered for each year are kept constant across measures, the simple average of Gini levels could cover a larger number of countries compared to our concept of global inequality, as it does not require information about net wealth. The "World Gini" calculated on all the countries for which we have the relevant information is shown in Table 5.7.

Table 5.8. Average Gini, international inequality and global inequality, 1820-2010

	Average Gini (population weighted)		Average Gini (population unweighted)		International inequality (weighted)		International inequality (unweighted)		Global inequality	
	World	FRA, SWE, GBR, USA	World	FRA, SWE, GBR, USA	World	FRA, SWE, GBR, USA	World	FRA, SWE, GBR, USA	World	FRA, SWE, GBR, USA
1820		87		86						
1850		88		87						
1870		87		88		22		29		88
1890		87		89		15		24		89
1910		90		91		12		20		91
1930		91		91		10		18		91
1950		82		84		17		25		82
1960	74	83	73	81	13	13	17	18	83	84
1970	73	80	71	76	15	10	13	14	75	81
1980	60	76	68	70	61	10	20	17	86	77
1990	60	74	68	70	63	8	29	15	84	76
2000	64	79	70	78	55	8	22	16	82	80
2010	74	83	72	75	37	1	18	5	80	84

Note: For an assessment of data quality, see Table 5.3 and Table 5.4. Table 5.8 includes all the countries for which we also had information about net wealth. Consequently, the table covers a smaller group compared to Table 5.7, leading to some slight differences in the value of the "World" average between this table and Table 5.8.

StatLink https://stat.link/t3xbv9

Correlation with GDP per capita

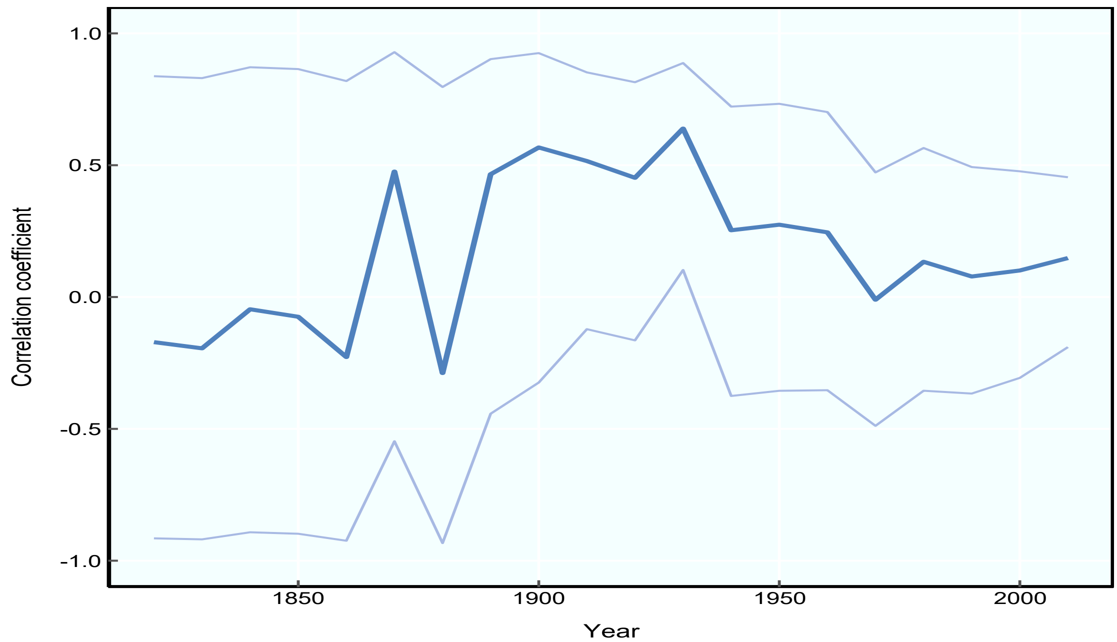
Figure 5.2 gives a general picture of the relationship between GDP per capita and wealth inequality over time and across countries. From 1820 to 1840, wealth inequality is only very weakly correlated with GDP per capita. From 1850 onwards, the correlation becomes stronger and it is positive: richer countries have a higher wealth inequality than poorer ones. Nevertheless, after World War I the relationship stops increasing and, although remaining positive until 2010. By then it is even weaker than in the first half of the 19th century.

Figure 5.3 shows the correlation between GDP per capita and the Gini index of wealth inequality for all countries. The three graphs correspond to three different periods. The first panel includes observations for the period before 1930, the second panel describes the correlation from 1950 to 1970, and the third refers to the period from 1980 to 2010.

For the first period, the correlation between GDP per capita and wealth inequality is positive for those countries with a GDP per capita less than USD 8 000; the correlation turns negative for countries with GDP per capita above that level. During the 1950s and 1970s, the relationship is slightly negative and turns positive only for the highest-income countries. However, there are very few observations in the bottom and top GDP region. For the last period, the relationship is positive until USD 16 000, it turns negative until USD 30 000, and then it becomes positive again. Due to the already-discussed limitations in the available information, the number of countries differs across period, which might affect our results.

Figure 5.2. Correlation between Gini coefficient of wealth inequality and GDP per capita, 1820-2010

Pearson correlation coefficient



Note: The thick and thin lines correspond to the Person correlation coefficient and the 95% confidence interval, respectively. For an assessment of data quality, see Table 5.3 and Table 5.4.

StatLink https://stat.link/8oprm6

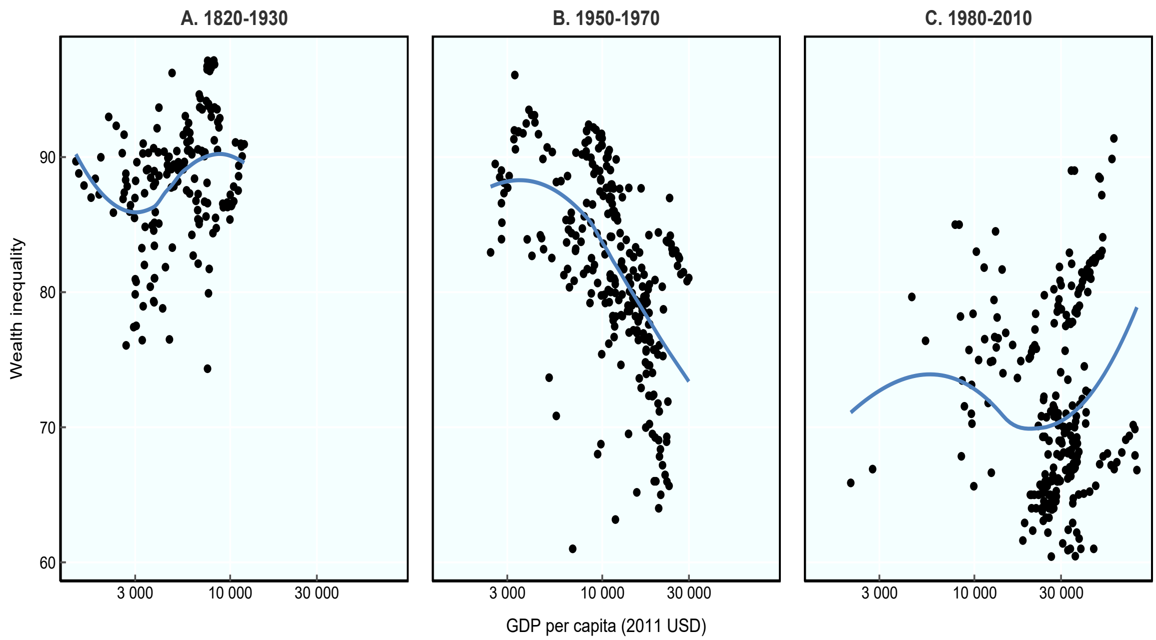
The overall conclusion is that there is no obvious and automatic correlation between GDP per capita and the change in wealth inequality in the period 1820-2010 – which confirms what has been argued for the pre-industrial period (1500-1800), when trends in wealth (and income) inequality have been reported to be largely unrelated to GDP (Alfani, 2015[8]; Alfani, 2019[43]; Alfani and Ryckbosch, 2016[16]). Overall, our findings confirm the view that simple "Kuznetsian" explanations are not sufficient to explain long-term inequality trends (Alfani, 2019[43]; Lindert, 2000[15]; Vecchi and Amendola, 2017[50]). Instead, we need to explore a broader range of causal factors beyond GDP change, embrace complexity in explanations, and pay great attention to context. An example of this approach is the recent masterful study of inequality in the United States from 1700 until today by Lindert and Williamson (2016[4]).

Priorities for Future Research

The current availability of data about wealth inequality – in both contemporary and past societies – is far from satisfying. Even in comparison with data we have regarding income inequality, which is also surely not adequate for current research needs, wealth fares significantly worse. For many areas of the world, there is no information from before the last few decades, and for many countries we have no information even for recent years. To some degree, this reflects the nature of many national fiscal systems, which today weigh much more on income than on wealth – hence generating an inferior amount of useful information about the latter.[7] But to this we have to add that research long neglected wealth in favour of income due to ideas about the long-run drivers of inequality growth that were probably misguided (chiefly, the view that wealth inequality is of only minor importance in determining total income inequality in the long run), as clearly exposed by many recent studies (Alfani, 2019[43]; Lindert and Williamson, 2016[4]; Piketty, 2014[6]; Piketty, 2015[17]; Roine and Waldenström, 2015[7]).

Figure 5.3. Correlation between the Gini coefficient of wealth inequality and GDP per capita in three time periods, 1820-2010

Semi-logarithmic scale



Note: Gini coefficient and GDP per capita at 2011 PPPs. The grey area around the correlation coefficient (shown as the blu line) corresponds to the confidence interval. Black dots corresponds to country estimates. For an assessment of data quality, see Table 5.3 and Table 5.4. This figure is based on all available annual information, not just the data included in Table 5.5. General patterns do not change when limiting the analysis to the countries in Table 5.5.

StatLink https://stat.link/c16mhj

In the current world, characterised by growing inequality of both income and wealth, questions about the causes, and about the very nature and the potential consequences, of distributional change have acquired high societal relevance. But to answer such questions, we need more and better data. This need is further strengthened by the fact that recent research highlights how levels and trends of economic inequality cannot be "deduced" from trends in other variables, like GDP per capita, but must be measured as directly as possible. In particular, recent historical evidence provides strong support for the view that inequality growth cannot be simply considered a side effect of increasing prosperity (Alfani, 2019[43]). In the foreseeable future, it seems probable that the area of the world for which the availability of data on the distribution of wealth will grow at the quickest pace is Europe, partly due to the relative abundance of useful historical sources and, even more, thanks to the many research groups that have risen to the challenge. This is undoubtedly a positive development, but the other side of the coin is that more will need to be done to better integrate other world areas – particularly Africa and Asia – into the broader picture.

References

Alfani, G. (2019), "Wealth and income inequality in the long run of history", in Diebolt, C. and M. Haupert (eds.), *Handbook of Cliometrics*, Springer International Publishing. [43]

Alfani, G. (2017), "The rich in historical perspective. Evidence for preindustrial Europe (ca. 1300-1800)", *Cliometrica*, Vol. 1/3, pp. 321-348. [9]

Alfani, G. (2015), "Economic inequality in northwestern Italy: A long-term view (fourteenth to eighteenth centuries)", *The Journal of Economic History*, Vol. 75/4, pp. 1058-1096. [8]

Alfani, G. and F. Ammannati (2017), "Long-term trends in economic inequality: The case of the Florentine State, c. 1300-1800", *The Economic History Review*, Vol. 70/4, pp. 1072-1102. [10]

Alfani, G. and M. Di Tullio (2019), *The Lion's Share. Inequality and the Rise of the Fiscal State in Preindustrial Europe*, Cambridge University Press. [11]

Alfani, G. and W. Ryckbosch (2016), "Growing apart in early modern Europe? A comparison of inequality trends in Italy and the Low Countries, 1500-1800", *Explorations in Economic History*, Vol. 62, pp. 143-153. [16]

Alvaredo, F.; A. Atkinson and L. Chancel (2016), "Distributional national accounts (DINA) guidelines: Concepts and methods used in WID.world", *WID.World Working Paper*, No. 2. [28]

Alvaredo, F., A. Atkinson and S. Morelli (2018), "Top wealth shares in the UK over more than a century", *The Journal of Public Economics*, Vol. 162, pp. 26-47. [35]

Alvaredo, F. et al. (2013), "The top 1 percent in international and historical perspective", *The Journal of Economic perspectives*, Vol. 27/3, pp. 3-20. [26]

Atkinson, A. (1980), *Wealth, Income and Inequality*, Oxford University Press. [2]

Atkinson, A., T. Piketty and E. Saez (2011), "Top incomes in the long run of history", *The Journal of Economic Literature*, Vol. 49/1, pp. 3-71. [25]

Balestra, C. and R. Tonkin (2018), "Inequalities in household wealth across OECD countries: Evidence from the OECD Wealth Distribution Database", *OECD Statistics Working Papers*, No. 2018/01, OECD Publishing, Paris, https://dx.doi.org/10.1787/7e1bf673-en. [27]

Bengtsson, E. et al. (2018), "Wealth inequality in Sweden, 1750-1900", *The Economic History Review*, Vol. 3, pp. 772-794. [12]

Borgerhoff Mulder, M. et al. (2009), "Intergenerational wealth transmission and the dynamics of inequality in small-scale societies", *Science*, pp. 682-688. [20]

Brandolini, A. et al. (2004), *Household Wealth Distribution in Italy in the 1990s*, Banca d'Italia. [37]

Canbakal, J., A. Filiztekin and S. Pamuk (2018), *Inequality of Income and Wealth in the Ottoman Empire, 1500-1800*, paper given at the World Economic History Congress (Boston, 30 July-3 August 2018). [13]

Coşgel, M. and B. Ergene (2012), "Inequality of wealth in the Ottoman Empire: War, weather, and long-term trends in eighteenth-century Kastamonu", *The Journal of Economic History*, Vol. 72/2, pp. 308-331. [29]

Credit Suisse (2017), *Global Wealth Databook 2017*, https://www.credit-suisse.com/about-us/en/reports-research/global-wealth-report.html (accessed on 14 January 2020). [39]

Credit Suisse (n.d.), *The Global Wealth Reports*, https://www.credit-suisse.com/about-us/en/reports-research/global-wealth-report.html (accessed on 12 January 2020). [32]

Davies, J., R. Lluberas and A. Shorrocks (2017), "Estimating the level and distribution of global wealth, 2000-2014", *The Review of Income and Wealth*, Vol. 63/4, pp. 731-759. [49]

Davies, J. et al. (2010), "The level and distribution of global household wealth", *The Economic Journal*, Vol. 121/551, pp. 223-254. [38]

Davies, J. and A. Shorrocks (2000), *The Distribution of Wealth*, Elsevier. [19]

Di Matteo, L. (2018), *The Evolution and Determinants of Wealth Inequality in the North Atlantic Anglo-Sphere, 1668-2013. Push and Pull*, Springer. [3]

Gabbuti, G. and S. Morelli (2019), *Inheritances and the Concentration of Personal Wealth in Italy from Unification to the Great War (1862-1915)*, Mimeo. [40]

Heilbroner, R. (2008), *Wealth*, Palgrave Macmillan. [21]

Ihle, D. and A. Siebert-Meyerhoff (2017), "The older, the richer? A decomposition of wealth inequality by age subgroups", *CAWM Discussion Paper*, Vol. 97. [46]

Kuznets, S. (1955), "Economic growth and income inequality", *The American Economic Review*, Vol. 45, pp. 1-28. [1]

Lindert, P. (2000), *Three centuries of inequality in Britain and America*, Elsevier. [15]

Lindert, P. (1991), "Toward a comparative history of income and wealth inequality", in *Income Distribution in Historical Perspective*, Cambridge University Press. [42]

Lindert, P. (1986), "Unequal English wealth since 1670", *The Journal of Political Economy*, Vol. 94/6, pp. 1127-62. [30]

Lindert, P. and J. Williamson (2016), *Unequal Gains. American Growth and Inequality Since 1700*, Princeton University Press. [4]

Milanovic, B. (2005), *Worlds Apart. Measuring International and Global Inequality*, Princeton University Press. [47]

Nicolini, E. and F. Ramos Palencia (2016), "Comparing income and wealth inequality in pre-industrial economies: Lessons from 18th-century Spain", *EHES, Working Papers*, Vol. 95. [14]

OECD (2015), *In It Together: Why Less Inequality Benefits All*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264235120-en. [18]

OECD (2013), *OECD Guidelines for Micro Statistics on Household Wealth*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264194878-en. [23]

OECD (n.d.), *OECD Wealth Distribution Database*, https://stats.oecd.org/Index.aspx?DataSetCode=WEALTH. [33]

Paglin, M. (1975), "The measurement and trend of inequality: A basic revision", *The American Economic Review*, Vol. 65, pp. 598-609. [45]

Piketty, T. (2015), “Putting distribution back at the center of economics: Reflections on capital in the twenty-first century”, *The Journal of Economic Perspectives*, Vol. 29/1, pp. 67-88. [17]

Piketty, T. (2014), *Capital in the Twenty-First Century*, Belknap Press of Harvard University Press. [6]

Piketty, T., G. Postel-Vinay and J. Rosenthal (2014), “Inherited vs self-made wealth: Theory & evidence from a rentier society (Paris 1872-1927)”, *Explorations in Economic History*, Vol. 51, pp. 21-40. [5]

Piketty, T., G. Postel-Vinay and J. Rosenthal (2006), “Wealth concentration in a developing economy: Paris and France, 1807-1994”, *The American Economic Review*, Vol. 96/1, pp. 236-256. [34]

Prados de la Escosura, L. (2007), *Inequality and poverty in Latin America: A long-run exploration*, M.I.T. Press. [48]

Roine, J. and D. Waldenström (2015), *Long run trends in the distribution of income and wealth*, Elsevier, https://www.sciencedirect.com/handbook/handbook-of-income-distribution/vol/2/suppl/C (accessed on 14 January 2020). [7]

Sahlins, M. (1974), *Stone Age Economics*, Routledge. [22]

Scheidel, W. (2017), *The Great Leveller: Violence and the Global History of Inequality from the Stone Age to the Present*, Princeton University Press. [44]

Schneider, M., M. Pottenger and J. King (2016), *The Distribution of Wealth-Growing Inequality?*, Edward Elgar Publishing. [36]

Sutch, R. (2016), *The Accumulation, Inheritance, and Concentration of Wealth during the Gilded Age: An Exception to Thomas Piketty’s Analysis*. [41]

van Zanden, J. et al. (2014), *How Was Life?: Global Well-being since 1820*, OECD Publishing, https://doi.org/10.1787/9789264214262-en. [24]

Vecchi, G. and N. Amendola (2017), *Inequality*, Oxford University Press. [50]

WID.world (n.d.), *The World Inequality Database*, https://wid.world/ (accessed on 12 January 2020). [31]

Notes

[1] The authors wish to thank Peter Lindert, Livio Di Matteo, Salvatore Morelli and Giacomo Gabbuti for supplying us with some data on wealth inequality. This chapter received support from the European Research Council under the European Union's Horizon 2020 Framework Program/ERC Grant agreement No. 725687, *SMITE-Social Mobility and Inequality across Italy and Europe, 1300-1800*, as well as from the Luxembourg National Research Fund (FNR) (12553347).

[2] Note that across our database, the Pearson coefficient between the Gini index and the share of the richest 10% is 0.94, suggesting a very strong correlation. The Pearson coefficient has been calculated on the 258 country-years for which both Gini and the share of the top 10% were available.

[3]With an R^2 of 0.8756, an adjusted R^2- of 0.8751 and a t-statistics of 42.61.

[4] Also see the section "Description of the Concepts Used" about methods.

[5] From 1870 to 1990, the share of the richest 10% declines from 83% to 65%.

[6] From 2000 to 2010, the share of the richest 10% grows from 70% to 75%.

[7] This is also why, paradoxically, for a growing number of countries, especially but not limited to Europe, we now know more about wealth inequality before 1820 than after such date: as income taxes replaced older taxes based on real estate and other wealth components.

6 Life expectancy and length of life inequality in the long run

Lamar Crombach, ETH Zürich, KOF Swiss Economic Institute
Jeroen Smits, Radboud University, Global Data Lab
Christiaan Monden, University of Oxford

This chapter describes trends in adult length of life and its distribution based on 15 144 life tables derived from various sources that for some countries cover a period of over 200 years. Since 1800, life expectancy in the most developed countries has increased from around 55 to 81 years for men and from 57 to 87 for women, an increase of about 50%. Concurrently, inequality in length of life in the best-performing countries has been cut by over 2/3 for both men and for women. This decrease, however, is not independent of the change in average life expectancy, as the two are strongly connected. Our data show that, in spite of great improvements in average life expectancy and reductions in length of life inequality since 1800, at each level of life expectancy there is substantial variation in inequality of length of life between different countries.

Introduction

The increase in life expectancy (LE) over the last 150 years is one of humankind's greatest achievements. Throughout the 19th century, LE at birth in the best-performing countries was less than 60 years (Maddison, 2001[1]). Nowadays, LE at birth is over 85 for women and over 80 for men (Oeppen and Vaupel, 2002[2]), and there are good reasons to expect this increase to continue in the coming decades[1] (Oeppen and Vaupel, 2002[2]; Vaupel et al., 1998[3]). The intuitive appeal of the concept of LE and its wide availability make it one of the most important indicators of the overall performance of societies (UNDP, 2017[4]).

While LE gives an excellent indication of the longevity of the members of a population, it is not informative of another fundamental aspect of length of life: the degree to which the available years of life are distributed equally among the population members. Length of life inequality (LI) is one of the most fundamental forms of inequality. Whereas other forms of inequality, such as in income, wealth, education or occupation, might be compensable using redistributive policies, a high level of LI indicates that a substantial share of the population has died prematurely, a situation from which no recovery is possible (Pradhan, Sahn and Younger, 2003[5]; Smits and Monden, 2009[6]). Redistributive policies might improve the situation of the current population but not for those who died prematurely in the past.

This chapter studies historical trends in both life expectancy and life inequality of those aged 15 and over. These indicators should be studied simultaneously, as they are highly correlated. Increases in LE due to, for instance, improvements in health care will in most cases be associated with a decrease of LI. If this correlation is not accounted for, changes in LI may, largely, reflect changes in LE (Smits and Monden, 2009[6]). Studying LI while taking into account the effect of LE thus provides additional valuable information.

The next section will discuss the central concepts in more detail. Thereafter, the historical data sources are described, followed by an assessment of their quality. We then present the main trends in the indicators for the different countries and regions of the globe over the last 150 years and then assess their correlations with GDP per capita (GDPpc). The final section will discuss priorities for future research.

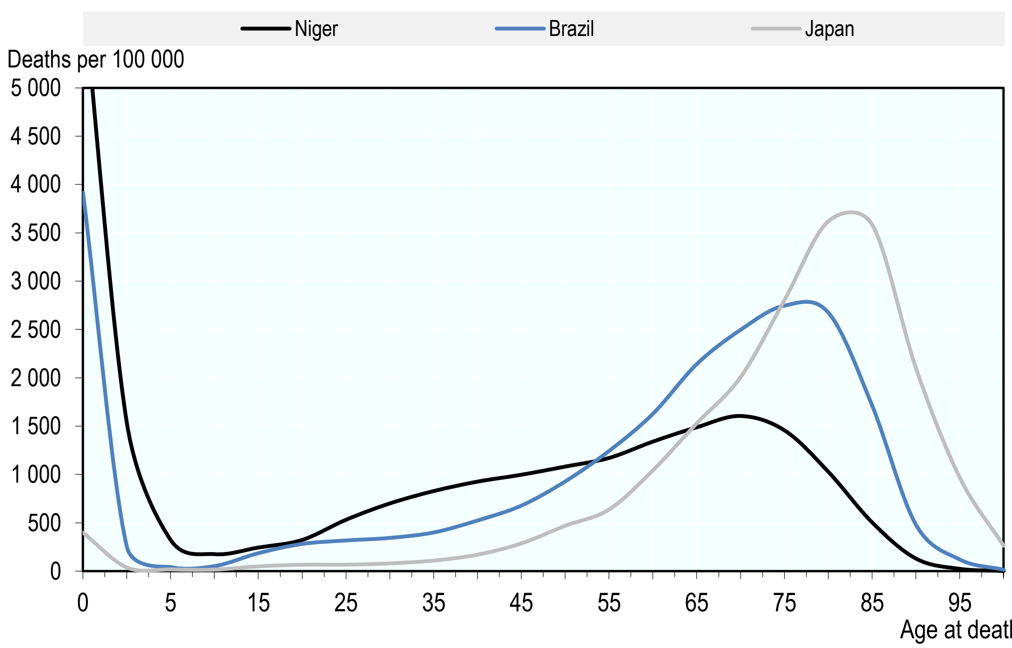
Description of the concepts used

Life expectancy (LE) – also called period life expectancy – is "the age that a person of a particular age is expected to reach based on the age-specific mortality rates prevailing at a specific point in time" (van Zanden et al., 2014[7]). LE does not accurately reflect how long individuals may expect to live, as it reflects the mortality rates prevailing in the year when they are born, rather than the rates experienced over the course of their life. It does, however, provide an overall indication of the health performance of a society at that moment in time (Smits and Monden, 2009[6]).

We illustrate the meaning of length of life inequality (LI) by depicting the distribution of length of life for countries at different levels of development. Figure 6.1 does this for men in Niger, Brazil and Japan in the year 2000. This figure, derived from Smits and Monden (2009[6]), shows that the distribution generally has two peaks. For all countries, there is a first peak at the start of life, reflecting the relatively high mortality at birth or soon thereafter. This peak is high in Niger and low in Japan, which reflects the strong reduction of infant and child mortality that countries experience in the course of their modernisation process. From the age category of 5-10 onwards, the number of individuals reaching a given age gradually increases until a second peak occurs for the group aged 65 and over. After this peak, the number of individuals reaching an older age quickly decreases. The difference in LI among the three countries is clearly reflected in the variation of the distributions surrounding the old age peak, which is much larger in Niger than in Japan.

This chapter focuses solely on the old age peak, because the two peaks are to a certain extent a reflection of different underlying processes, e.g. Edwards and Tuljapurkar (2005[8]). A reduction of infectious diseases, mainly due to improved sanitary practices and the diffusion of new and more effective medicine, is the main cause of the reductions of infant and child mortality that countries experience. Such developments, however, have less influence on adult mortality patterns, which change at a slower pace and are partly determined by other factors (Bloom, Canning and Sevilla, 2003[9]; Cutler and Meara, 2004[10]; Marmot, 2005[11]). A comparison of inequality measures of a two-peaked distribution is also problematic, as one would be unable to determine whether changes in the inequality measures are a result of changes around the first peak or the second one. As we are mostly interested in distributions among adults, we have chosen to focus on individuals aged 15 and older.

Figure 6.1. Distribution of length of life for men in Niger, Brazil and Japan, 2000



Source: (Smits and Monden, 2009[6]), "Length of life inequality around the globe", *Social Science & Medicine*, Vol. 68/6, pp. 1114-1123, http://dx.doi.org/10.1016/j.socscimed.2008.12.034.

StatLink https://stat.link/vg6l9j

As Figure 6.1 illustrates, most differences in the distribution of length of life between countries at different levels of development are on the left side of the old-age peak. This is largely due to the fact that the distribution of length of life is bounded at the right (Smits and Monden, 2009[6]). It is not bounded completely, as over time the proportion of persons who become very old grows steadily, but it is bounded in a practical sense, as the actual number of very old persons remains limited. There is no fundamental reason why human beings could not become 150 years old, but as far as we know, nobody has ever reached that age. Moreover, the technology to allow more individuals to reach old age is more readily available than the technology necessary for old-age individuals to reach an even older age.

Given this right-hand side boundedness of the distribution, most of the dynamics in the distribution of length of life take place left of the second peak. This implies that, with increasing LE, length of life becomes more and more concentrated in a small age band around the second peak. As such, the variation in length of life diminishes as LE increases. Indeed, previous research has shown a strong negative correlation between LE and LI (a Pearson's *r* exceeding 0.9) (Shkolnikov, Andreev and Begun, 2003[12]; Wilmoth and Horiuchi, 1999[13]; Edwards and Tuljapurkar, 2005[8]; Seaman, Leyland and Popham, 2016[14]; Németh, 2017[15]; Smits and Monden, 2009[6]). This raises the question whether changes in LI as such offer much new insight into social inequality, beyond what is already known from changes in LE.

The concept of relative length of life inequality (RLI), introduced by Smits and Monden (2009[6]), addresses this issue by studying variation in LI between countries at similar levels of LE. In this manner, an indicator of LI is obtained that is independent of LE. In this chapter, this approach is illustrated by looking at variation in LI and in premature mortality within country-year combinations with similar levels of LE, and by showing trajectories of countries through the LE-LI space. In this way we aim to increase our understanding of the degree to which changes in LI are the result of changes in LE or of other factors, such as social distribution mechanisms.

Measurement

LE is calculated as the mean age of death from age 15 onwards, while LI is measured by computing the Gini coefficient over the distribution of age at death from the same population. The distributions of age at death were obtained by applying the age- and sex-specific mortality rates from the life table to a population of 100 000 individuals aged 15, thus standardising for differences in adult population structure among countries and time periods. In the remainder of this chapter, LE and LI refer to the population aged 15 and over. Premature mortality is defined as the sum of all life-table deaths per 1 000 in the 15-50 age group (Smits and Monden, 2009[6]).

Historical sources

The data needed for the computation of LE and LI consists of period life tables (LTs). These are tables with information on the total population and the number of deaths for a specific region or country in a specific time period. Given the simplicity of the required information, this kind of data has been routinely collected by statistical offices of many countries for centuries. For this reason, LT series in many (particularly European) countries date back to the 19th, 18th or even 17th century. As there are substantial differences between the mortality patterns of men and women, separate LTs for the two groups are generally constructed.

For developed countries, LTs are regularly produced by statistical offices and national health monitoring organisations on the basis of vital and population statistics information. The most important source of life tables is the *Human Mortality Database (HMD)* through which a large number of high-quality life tables is made available. These life tables have been constructed according to a standardised procedure from birth and death counts derived from vital statistics and population data from censuses and official population estimates. The HMD provides life tables for about 40 countries, including long time series that sometimes date back to the 19th or even 18th century (HMD[16]). The *HMD* is the main data source used in this chapter. For countries not represented in the *HMD*, a second preferred source is life tables derived directly from statistical offices. These life tables are generally also of good quality, but because of differences in methods and data quality they might be less comparable than those derived from the *HMD*.

Because low- and middle-income countries are not well represented in the *HMD*, and because the statistical offices of these countries generally do not publish life tables, we had to rely on other sources for them. In order of preference, we used life tables from the *Global Health Observatory* of the World Health Organization (WHO, 2018[17]), from the *Human Lifetable Database* (Shkolnikov, 2017[18]), or from the *GDL Length of Life Database* (Smits and Monden, 2009[6]). In total, our database includes 15 144 sex-specific LTs, of which 7 582 for men and 7 562 for women.

The time period for which the deaths are counted in an LT is often a year, but it may also be a shorter or longer period. For the population size of the age groups, the mid-year population is generally (but not always) used. The size of the age groups is usually one or five years. When five-year groups are used, the first two groups deviate from the general pattern, as the first group includes only children under the age of one and the second group only children aged one to four. With the third group aged five to nine, the regular five-year pattern starts. The highest group is always an open-ended group that includes all persons of and above the highest age (e.g. 85+, 95+, 100+, 110+).

To achieve comparability between the different sources, all LTs used in this chapter are abridged (85+) period LTs, using 5-year age intervals. We chose to work with abridged 85+ LTs, because for part of the countries and time periods it is the only type available. Additionally, we wanted to include as many LTs as possible to achieve the widest possible coverage. LTs with more detail in our database were recalculated into the abridged 85+ form to make them comparable to the others. Hence LI among the eldest (85+) age group is not reflected in our results.

In total, our database includes 15 144 life tables for 203 countries from all regions of the world. The database is available at the GDL Length of Life website (Global Data Lab[19]).

Data quality

An overview of the quality of the data sources is provided in Table 6.1. As with data from other historical sources, life tables are not without problems. Although the required data (number of deaths and total population) are not very complex, the quality of the available records depends on the state of development and procedures used by administrative offices, which might have differed considerably between countries and regions. Besides such obvious differences in the original data sources, the quality of the life tables that are used in this chapter also varies depending on the method that was used to construct them.

As discussed above, the 6 886 life tables derived from the *Human Mortality Database* are of highest quality (Rank 1 in Table 6.1, as they are constructed according to a standardised method from the source material. Life tables from the HMD are therefore fairly comparable between countries and years.

The life tables of national statistical offices and most other sources are more varied and therefore of somewhat lower quality (Rank 2). This is also true for the life tables from the Global Health Observatory of the World Health Organization, which includes the best possible values and estimates for all WHO member states except for a few island states (WHO, 2018[17]). The quality may vary however between countries and regions, and sometimes life tables are estimated. This also applies to the *Human Lifetable Database* (Max Planck/Berkeley), which contains published life tables from a broad range of sources for which no further information on data quality is available (Shkolnikov, 2017[18]). Both sources are therefore indicated with Quality Rank 2 in Table 6.1. Of the life tables used for this chapter, 7 984 are of Rank 2. The remaining 324 life tables from other – miscellaneous – sources derived from the Smits and Monden (2009[6]) life table database are indicated with Quality Rank 3.

Table 6.1. Data quality

	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1800s	..	..	2	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1810s	..	..	2	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1820s	..	..	1	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1830s	..	..	1	..	..	..	1	..	..	..	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1840s	2	..	1	..	..	..	1	..	..	..	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1850s	2	1	1	..	..	..	1	..	..	..	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..
1860s	2	1	1	..	..	..	1	..	..	..	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..
1870s	2	1	1	2	1	..	1	..	..	..	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..
1880s	2	1	1	2	1	..	1	..	..	2	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..
1890s	2	1	1	2	1	..	1	..	2	2	2	2	..	..	..	..	..	..	..	..	..	2	..	..	..
1900s	2	1	1	2	1	2	1	..	..	2	2	2	..	..	..	..	..	..	..	..	..	2	2	..	..
1910s	..	1	1	..	1	1	1	..	..	..	..	2	..	..	..	..	..	..	..	..	..	2	2	..	..
1920s	2	1	1	2	1	1	1	3	2	1	1	2	..	..	..	..	..	..	..	..	..	2	2	..	..
1930s	1	1	1	2	1	1	1	2	2	1	1	2	..	..	..	..	..	..	..	..	2	2	2	..	..
1940s	1	1	1	2	1	1	1	..	..	1	1	1	..	..	..	2	..	..	..	2	..	..	..	..	..
1950s	1	1	1	2	1	1	1	2	2	1	1	1	3	..	..	2	..	..	..	2	..	1	2	..	2
1960s	1	1	1	2	1	1	1	1	1	1	1	1	2	..	3	2	3	..	..	2	..	1	3	3	..
1970s	1	1	1	2	1	1	1	1	1	1	1	1	2	2	3	3	3	3	..	2	3	1	3	3	3
1980s	1	1	1	2	1	1	1	1	1	1	1	1	2	2	2	..	..	..	..	..	3	1	3	..	2
1990s	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	..	..	..	..	..	3	1	3	..	2
2000s	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	..	2	2	2	2	1	3	2	2
2010s	1	1	2	2	2	2	1	2	2	2	2	1	2	2	2	2	2	2	2	2	2	1	2	2	2

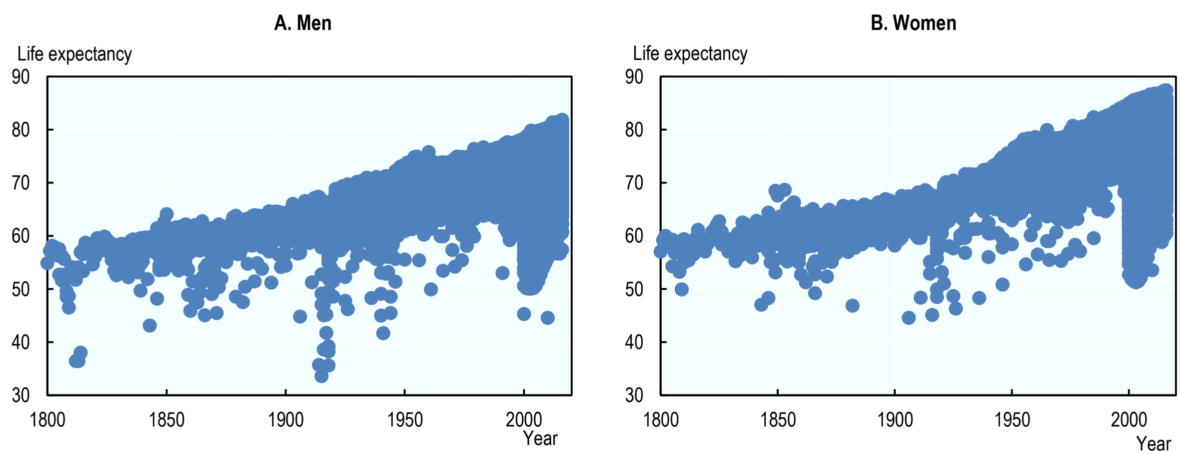
Main highlights of trends in life expectancy and length of life inequality

Life expectancy

Figure 6.2 provides an overview of LE (15+) for men (left) and women (right) for the period 1800-2016. We see that, over time, the number of LTs has increased gradually until 2000. A large increase in the number of LTs is observable since 2000, when our dataset starts to include life tables from the WHO for all the world's countries. The strong increase of observations (country-years) with a low level of life expectancy is due to the fact that the WHO database contains many life tables for low- and middle-income countries, for which none or only a few LTs were available earlier.

Until 1900, LE in the best-performing countries was less than 65 years for men and less than 67 years for women. For men, life expectancy then increased until about 72 around 1950, plateaued at this level until about 1970, and then increased gradually until its current value of a little bit above 82. For women, a similar pattern can be observed.

Figure 6.2. Life expectancy at birth for people aged 15 and over by gender, 1800-2016



Note: Blue dots represent country-year observations.
Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/l9jvb2

Periods of war are observable in the left panel, showing severe drops in LE for men in selected years. Clear examples are the Napoleonic wars in the 1803-1815 period, including his 1812 campaign into Russia, and the First and Second World Wars. For women, wars have less influence on LE than for men, but the disastrous Spanish flu of 1918 is clearly visible.

Table 6.2 and Table 6.3 provide an overview per decade of LE for all men and women in the 25 countries on which this book focuses for the period 1800-2016. One can observe a large degree of variation between countries and over time. The tables show that LE has increased in all 25 countries. A man living in Sweden in the first decade of the 19th century was expected to live for about 57 years, but in the 2010s he is expected to live 80.4 years. Other countries have made similar gains, though at a different pace. It took Sweden about two centuries to improve LE by 23 years. This same improvement was achieved in India in about half the time – a clear indication that less developed countries are leapfrogging the stages of health innovation as late adopters of new technologies. However, many countries have not yet reached the final stage, as large differences in LE across countries persist in the 2010s. Although India has improved at a fast pace, its LE for men is currently still 70.3 years, compared to the LE of 80.9 of the current frontrunner Australia. This is a difference of 10.6 years. African countries lag even further behind, with an LE of 61.7 for men and 62.4 for women in Nigeria, and an LE of 62.0 for men and of 68.1 for women in South Africa. As such, there is still great potential for improvement.

Some variation also exists amongst high-income countries. While Australia and Canada have the highest male LE in the 2010s (of 80.9 and 80.6 years respectively), male LE in the United States and Germany is clearly lower (77.2 and 78.5 respectively). For women, all Western European and Western Offshoot countries have an LE of over 83 in the 2010s, while the United States' LE remains below 82.

Regarding gender differences, the figures reveal higher LE for women than for men in almost all cases. The most extreme case is Russia, where women on average lived 11.1 years longer than men in the 2010s, and 12.8 years longer in the 2000s: a staggering difference. In Nigeria, women also lived longer than men in the 2010s, but this difference was limited to only 0.7 years. In some exceptional cases, men lived longer than women, as in India in the period between the 1900s and the 1970s, reflecting the "missing women" phenomenon noted by Sen (1990[20]).

Periods of war are clearly observable as well. A large drop in LE of 10.8 years is observed in the 1910s for French and Italian men during World War I. We also see a drop in LE in the 1940s during World War II in the Netherlands, France and Italy. After World War II, LE recovered rather quickly and since then shows a steady increase over time, with a faster pace in less developed countries.

Table 6.2. Life expectancy of men aged 15 and over, 1800s-2010s

Years, decadal averages

	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1800s	..	..	35.1	..	..	..	40	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1810s	..	..	35.4	..	..	..	40.4	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1820s	..	..	44	..	..	..	41.4	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1830s	..	..	42.8	..	..	..	40.7	..	..	..	42.1	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1840s	43.2	..	42.9	..	..	..	42.6	..	..	..	43.1	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1850s	43.3	42	42.8	..	..	..	42.5	..	..	..	43.6	43.1	..	..	..	..	..	..	..	..	..	..	..	..	..
1860s	43.1	43.1	44.3	..	..	..	45.1	..	..	..	44.0	44.6	..	..	..	..	..	..	..	..	..	..	..	..	..
1870s	43.1	44.1	42.3	42.5	42.6	..	46.1	..	..	..	44.2	44.4	..	..	..	..	..	..	..	..	..	..	..	..	..
1880s	44.1	46.2	44.3	43.5	44.9	..	48.1	..	..	44.8	45	44.2	..	..	..	..	..	..	..	..	..	..	..	..	..
1890s	44.6	47.4	44.6	45.2	46.4	..	48.7	..	44.6	46.6	45.7	45.1	..	..	..	..	..	..	..	..	..	43.6	..	..	..
1900s	47.1	49.8	44.8	46.7	47.3	43.3	49.6	..	..	48.6	47.1	46.1	..	..	..	..	..	..	..	..	..	44.2	29.9	..	..
1910s	..	50.8	34.0	..	40.4	43.8	49.2	..	..	..	..	46.6	..	..	..	..	..	..	..	..	..	43.9	30.1	..	..
1920s	50.3	53.8	47.0	50.8	49.4	45.8	52.5	49.6	46.5	51.9	52.8	49.3	..	..	..	..	..	..	..	..	..	43.0	31.1	..	..
1930s	51.5	55.6	47.9	52.3	51.3	45.3	53.8	48.7	46.4	53.4	53.9	50.3	..	..	..	..	..	..	..	..	44.2	44.0	33.1	..	..
1940s	52.2	53.8	45.8	..	48.3	47.9	56.0	..	..	54.2	54.7	52.3	..	..	..	45.1	..	..	..	41.4	..	..	..	..	..
1950s	54.8	58.5	53.9	55.3	55.7	54.6	57.9	53.1	51.5	54.6	55.8	54.1	49.5	..	..	50.0	..	..	..	45.0	..	52.3	39.9	..	52.9
1960s	55.4	58.0	54.8	54.9	56.0	56.6	58.3	54.9	52.1	54.8	56.2	54.2	52.2	..	54.0	51.0	51.0	..	..	45.2	..	55.0	44.8	39.7	..
1970s	56.0	57.9	55.6	55.0	56.8	57.5	58.3	54.4	50.0	55.8	56.8	55.2	53.9	56.5	54.1	50.4	53.0	48.6	..	42.9	55.0	57.9	48.7	45.3	53.7
1980s	57.7	59.0	57.3	56.8	58.4	59.3	59.5	53.7	50.2	58.4	58.9	57.2	55.2	54.7	56.1	..	..	..	..	..	55.9	60.6	51.2	..	53.9
1990s	59.7	60.3	59.5	58.8	60.5	60.2	61.6	53.8	47.3	61.0	60.8	58.7	57.3	52.6	56.6	..	..	..	..	..	57.0	62.2	52.9	..	54.5
2000s	62.3	62.6	62.1	61.7	63.4	62.6	63.7	56.4	46.2	63.9	63.3	60.8	59.4	55.3	57.8	55.0	..	46.3	45.1	44.0	59.6	64.0	54.0	53.8	55.6
2010s	64.4	64.9	64.3	63.5	65.5	65.1	65.4	58.5	50.8	65.9	65.6	62.2	59.7	57.1	59.1	55.1	60.2	52.9	46.7	47.0	60.6	65.5	55.3	54.3	57.5

Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/tw1j4g

Table 6.3. Life expectancy of women aged 15 and over, 1800s-2010s

Years, decadal averages

	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1800s	..	..	42.7	..	..	..	42.3	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1810s	..	..	43.4	..	..	..	43.0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1820s	..	..	44.2	..	..	..	45.0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1830s	..	..	43.3	..	..	..	44.4	..	..	..	43.3	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1840s	44.0	..	43.8	..	..	..	46.2	..	..	..	44.5	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1850s	44.2	42.9	43.5	..	..	..	45.7	..	..	..	45.3	44.1	..	..	..	..	..	..	..	..	..	..	..	..	..
1860s	44.5	43.8	44.8	..	..	..	48.1	..	..	..	45.7	45.5	..	..	..	..	..	..	..	..	..	..	..	..	..
1870s	45.0	44.9	44.3	44.4	42.4	..	48.8	..	..	..	46.0	45.4	..	..	..	..	..	..	..	..	..	..	..	..	..
1880s	46.3	47.4	45.9	45.6	44.2	..	50.3	..	..	47.9	46.9	44.9	..	..	..	..	..	..	..	..	..	..	..	..	..
1890s	47.1	48.9	46.7	47.3	46.0	..	50.6	..	45.5	49.6	47.8	46.1	..	..	..	..	..	..	..	..	..	44.3	..	..	..
1900s	49.8	51.0	47.7	48.9	47.4	44.7	51.3	..	..	51.5	49.3	48.0	..	..	..	..	..	..	..	..	..	44.5	29.7	..	..
1910s	..	51.6	48.2	..	47.3	45.2	51.2	..	..	..	..	48.9	..	..	..	..	..	..	..	..	..	44.2	30.1	..	..
1920s	53.1	53.9	50.4	52.3	50.4	48.4	53.8	50.9	51.3	54.9	53.2	50.7	..	..	..	..	..	..	..	53.0	..	44.0	31.2	..	..
1930s	54.7	56.0	52.6	54.5	53.3	50.9	55.1	51.0	52.7	56.5	55.1	53.2	..	..	..	..	..	..	..	..	43.4	46.2	33.1	..	..
1940s	56.9	56.9	53.6	..	54.7	53.9	57.9	..	..	57.9	57.4	56.7	..	..	..	51.1	..	..	..	44.5	..	..	..	..	..
1950s	59.8	60.9	59.5	59.0	59.3	58.8	60.5	57.9	59.5	60.0	60.4	59.8	52.4	..	..	54.4	..	..	..	49.0	..	55.9	39.1	..	57.1
1960s	61.1	62.5	61.6	60.1	61.1	61.3	62.2	60.4	60.8	61.0	62.2	60.9	55.3	..	59.9	55.4	53.3	..	..	51.4	..	59.7	44.2	41.7	59.2
1970s	62.0	63.7	63.2	61.7	62.8	63.0	63.9	61.4	60.5	62.4	63.7	62.6	58.2	61.4	60.6	53.6	55.3	50.9	..	51.2	57.5	62.8	49.8	47.0	59.8
1980s	63.4	65.4	65.3	63.6	64.8	65.6	65.4	61.7	60.5	64.8	65.5	64.2	60.9	60.7	62.2	..	..	..	..	..	59.3	66.1	53.3	..	60.9
1990s	64.8	65.9	67.4	65.2	66.9	67.3	66.7	62.4	59.3	66.7	66.6	65.0	62.6	59.7	63.6	..	..	..	..	..	61.2	68.5	55.5	..	62.1
2000s	66.6	67.0	69.1	67.3	68.9	69.1	68.0	64.8	59.0	68.7	68.0	65.8	64.5	62.8	64.8	59.8	..	47.4	45.6	48.0	62.5	70.8	56.7	57.2	62.2
2010s	68.1	68.4	70.5	68.3	70.1	70.7	69.1	66.5	62.0	69.9	69.6	66.9	65.2	64.5	65.8	59.8	65.7	56.5	47.4	53.1	63.5	71.9	58.6	58.0	64.6

Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

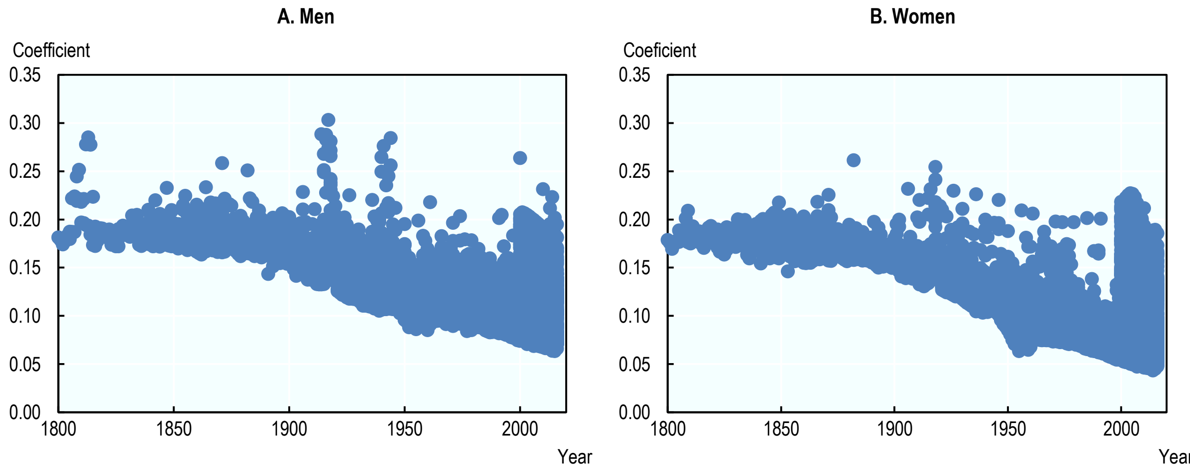
StatLink https://stat.link/4v5faz

Length of life inequality

Figure 6.3 shows the Gini coefficients for length of life inequality (LI) in the population aged 15 and over for the period 1800 to 2016. The high correlation between LE and LI (discussed in Section 6.2) is clearly visible, as Figure 6.3 (on length of life inequality) almost exactly mirrors Figure 6.2 (on average life expectancy).

Figure 6.3. Length of life inequality among people aged 15 and over by gender, 1800-2016

Gini Coefficient



Note: Blue dots represent country-year observations.
Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/qfhkxl

The Gini coefficient for length of life inequality for men runs from about 0.07 to 0.30 and for women from 0.04 to 0.25. These are low levels compared to Gini values for income inequality at the country level, which run from about 0.25-0.30 in the best-performing countries to about 0.60-0.65 in the worst-performing countries (Solt, 2019[21]). This means that the distribution of years of life in the 15+ population is clearly narrower than the income distribution. This does not mean, however, that length of life inequality is an unimportant issue. While income inequality can be compensated using redistributive policies, a high value of LI means that people have died at too young an age, a situation that cannot be compensated.

Figure 6.3 indicates that LI has decreased over time. This is confirmed by Table 6.4 and Table 6.5, which show the trends in LI for the 25 countries in the period 1800-2016. In the large majority of cases, we see a clear decrease of LI. There are some notable exceptions in the 2000s and 2010s, when Mexico and Egypt faced increased inequality. Over the past few decades, changes in LI have been rather small in India, Indonesia, Thailand, Argentina, Brazil and Mexico. In the developed world, the United States has a relatively large LI compared to other developed countries, an issue that is discussed in more detail later in this chapter.

Table 6.4. Length of life Inequality among men aged 15 and over, 1800s-2010s

Years, decadal averages

	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1800s	..	..	0.235	..	..	..	0.189	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1810s	..	..	0.227	..	..	..	0.187	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1820s	..	..	0.187	..	..	..	0.183	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1830s	..	..	0.193	..	..	..	0.184	..	..	..	0.191	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1840s	0.182	..	0.195	..	..	..	0.176	..	..	..	0.185	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1850s	0.185	0.188	0.194	..	..	..	0.181	..	..	..	0.182	0.192	..	..	..	..	..	..	..	..	..	..	..	..	..
1860s	0.187	0.186	0.181	..	..	..	0.170	..	..	..	0.180	0.184	..	..	..	..	..	..	..	..	..	..	..	..	..
1870s	0.180	0.182	0.194	0.173	0.188	..	0.170	..	..	..	0.179	0.186	..	..	..	..	..	..	..	..	..	..	..	..	..
1880s	0.167	0.170	0.184	0.167	0.177	..	0.164	..	..	0.177	0.174	0.185	..	..	..	..	..	..	..	..	..	..	..	..	..
1890s	0.161	0.161	0.177	0.158	0.166	..	0.164	..	0.172	0.163	0.170	0.177	..	..	..	..	..	..	..	..	..	0.174	..	..	..
1900s	0.150	0.147	0.172	0.149	0.161	0.176	0.162	..	..	0.149	0.162	0.165	..	..	..	..	..	..	..	..	..	0.176	0.228	..	..
1910s	..	0.142	0.227	..	0.211	0.175	0.167	..	..	..	..	0.162	..	..	..	..	..	..	..	..	..	0.18	0.228	..	..
1920s	0.136	0.122	0.160	0.131	0.150	0.163	0.143	0.147	0.161	0.133	0.135	0.144	..	..	..	..	..	..	..	..	..	0.181	0.225	..	..
1930s	0.128	0.110	0.153	0.124	0.138	0.170	0.129	0.143	0.161	0.123	0.126	0.140	..	..	..	..	..	..	..	..	0.181	0.178	0.221	..	..
1940s	0.128	0.127	0.172	..	0.168	0.158	0.114	..	..	0.112	0.118	0.130	..	..	..	0.166	..	..	..	0.184	..	..	..	..	..
1950s	0.102	0.095	0.115	0.103	0.108	0.114	0.098	0.117	0.133	0.110	0.111	0.119	0.165	..	..	0.130	..	..	..	0.162	..	0.123	0.199	..	0.144
1960s	0.100	0.097	0.112	0.101	0.106	0.102	0.095	0.111	0.133	0.108	0.109	0.118	0.148	..	0.118	0.131	0.132	..	..	0.157	..	0.106	0.165	0.178	..
1970s	0.098	0.097	0.112	0.104	0.101	0.099	0.097	0.113	0.141	0.107	0.109	0.117	0.141	0.130	0.120	0.127	0.126	0.156	..	0.164	0.110	0.097	0.132	0.154	0.138
1980s	0.094	0.091	0.108	0.098	0.097	0.097	0.093	0.117	0.137	0.099	0.100	0.11	0.129	0.131	0.111	..	..	..	..	..	0.104	0.009	0.123	..	0.139
1990s	0.090	0.086	0.104	0.097	0.093	0.098	0.085	0.118	0.150	0.091	0.092	0.107	0.117	0.131	0.109	..	..	..	..	..	0.102	0.085	0.127	..	0.141
2000s	0.085	0.079	0.093	0.088	0.081	0.087	0.077	0.111	0.156	0.081	0.084	0.100	0.111	0.125	0.106	0.106	..	0.179	0.167	0.171	0.087	0.081	0.124	0.114	0.138
2010s	0.079	0.072	0.085	0.081	0.072	0.076	0.072	0.105	0.142	0.075	0.078	0.098	0.113	0.123	0.103	0.107	0.084	0.144	0.159	0.154	0.084	0.075	0.120	0.111	0.127

Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/6id1j3

Table 6.5. Length of life Inequality among women aged 15 and over, 1800s-2000s

Years, decadal averages

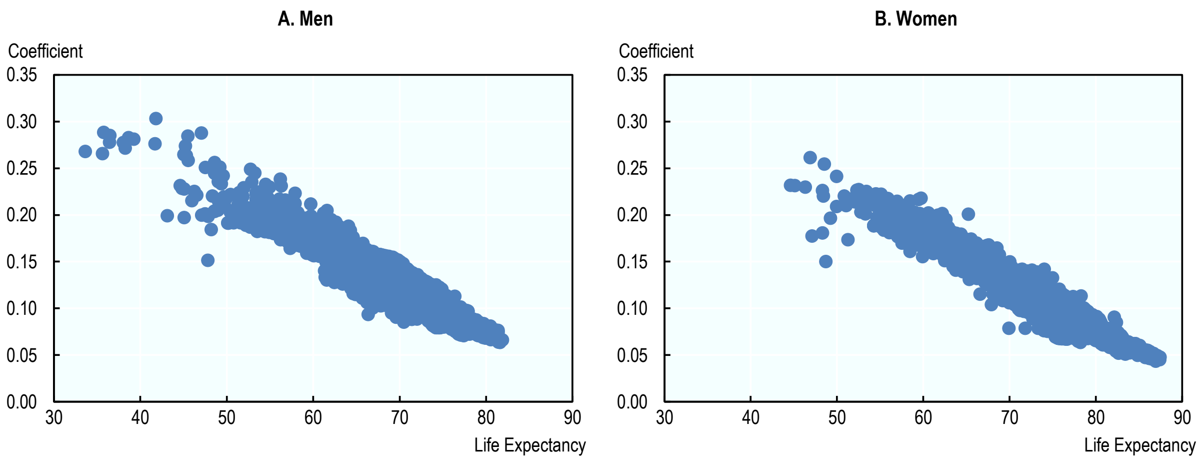
	Western Europe							Eastern Europe		Western Offshoots			Latin America and Caribbean			Middle East and North Africa		Sub-Saharan Africa			East Asia		South and Southeast Asia		
	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1800s	..	..	0.181	..	..	..	0.183	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1810s	..	..	0.183	..	..	..	0.181	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1820s	..	..	0.188	..	..	..	0.173	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1830s	..	..	0.191	..	..	..	0.173	..	..	..	0.196	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1840s	0.185	..	0.191	..	..	..	0.166	..	..	..	0.189	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1850s	0.188	0.190	0.192	..	..	..	0.170	..	..	..	0.184	0.193	..	..	..	..	..	..	..	..	..	..	..	..	..
1860s	0.187	0.187	0.185	..	..	..	0.160	..	..	..	0.181	0.185	..	..	..	..	..	..	..	..	..	..	..	..	..
1870s	0.182	0.183	0.189	0.169	0.187	..	0.162	..	..	..	0.179	0.187	..	..	..	..	..	..	..	..	..	..	..	..	..
1880s	0.165	0.170	0.181	0.162	0.182	..	0.159	..	..	0.169	0.174	0.189	..	..	..	..	..	..	..	..	..	..	..	..	..
1890s	0.156	0.158	0.172	0.152	0.171	..	0.159	..	0.173	0.158	0.168	0.179	..	..	..	..	..	..	..	..	..	0.192	..	..	..
1900s	0.143	0.145	0.166	0.144	0.165	0.174	0.158	..	..	0.145	0.158	0.160	..	..	..	..	..	..	..	..	..	0.194	0.232	..	..
1910s	..	0.143	0.165	..	0.167	0.176	0.160	..	..	..	..	0.154	..	..	..	..	..	..	..	..	..	0.199	0.231	..	..
1920s	0.130	0.126	0.153	0.126	0.149	0.161	0.138	0.143	0.150	0.127	0.140	0.141	..	..	..	..	..	..	..	0.130	..	0.198	0.230	..	..
1930s	0.122	0.110	0.139	0.114	0.132	0.147	0.124	0.139	0.145	0.116	0.127	0.132	..	..	..	..	..	..	..	..	0.201	0.186	0.226	..	..
1940s	0.111	0.106	0.134	..	0.121	0.133	0.103	..	..	0.104	0.110	0.113	..	..	..	0.142	..	..	..	0.186	..	..	..	..	..
1950s	0.090	0.083	0.094	0.086	0.091	0.100	0.084	0.099	0.101	0.092	0.093	0.098	0.146	..	..	0.115	..	..	..	0.155	..	0.115	0.209	..	0.140
1960s	0.085	0.078	0.086	0.081	0.084	0.085	0.078	0.088	0.094	0.089	0.086	0.094	0.130	..	0.097	0.109	0.128	..	..	0.138	..	0.090	0.173	0.181	0.129
1970s	0.084	0.076	0.081	0.079	0.078	0.077	0.076	0.083	0.093	0.085	0.083	0.090	0.115	0.101	0.096	0.105	0.12	0.145	..	0.142	0.108	0.078	0.140	0.153	0.118
1980s	0.080	0.072	0.074	0.073	0.071	0.069	0.071	0.083	0.091	0.076	0.077	0.085	0.095	0.097	0.086	..	..	..	..	..	0.098	0.067	0.129	..	0.107
1990s	0.075	0.070	0.068	0.072	0.066	0.063	0.067	0.081	0.098	0.069	0.071	0.082	0.086	0.096	0.082	..	..	..	..	..	0.091	0.061	0.122	..	0.098
2000s	0.069	0.066	0.061	0.064	0.058	0.056	0.061	0.074	0.102	0.061	0.066	0.078	0.087	0.088	0.080	0.091	..	0.183	0.172	0.184	0.079	0.055	0.115	0.103	0.097
2010s	0.064	0.061	0.056	0.060	0.053	0.050	0.057	0.069	0.092	0.056	0.062	0.076	0.085	0.085	0.079	0.090	0.062	0.128	0.162	0.151	0.076	0.050	0.105	0.100	0.085

Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/y1cubv

It remains to be seen what the decrease of length of life inequality observed for most countries exactly means. As discussed earlier in this chapter, there exists a strong negative correlation between the trends in LE and LI, with LI almost automatically decreasing when LE increases. This correlation, which is -0.94 for men and -0.97 for women in our data, is clearly visible in Figure 6.4, which depicts the relationship between LI and LE without the time dimension. Each dot in this figure represents the values of LE and LI for a specific country in a given year; the left figure is based on 7 582 life tables for men and the right figure on 7 562 life tables for women.

Figure 6.4. The relationship between average life expectancy and length of life inequality for men and women



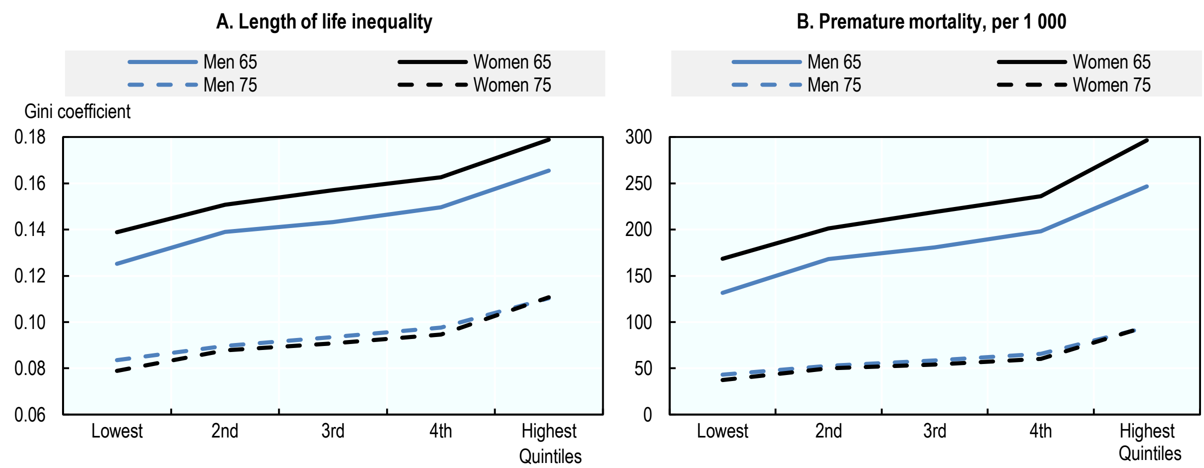
Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/zbwksn

The extremely strong association between LE and LI, as depicted in Figure 6.4, raises the question whether changes in LI as such may offer many new insights beyond what is known from changes in LE. Smits and Monden (2009[6]) answer this question confirmatively, by pointing to the fact that at each level of LE there is substantial variation between observations (country-years) with more and with less inequality. In their data, the number of premature deaths in the 15-50 age group is significantly and considerably higher in the 20% most unequal country-years compared to the 20% least unequal country-years for a given level of LE. They therefore introduce the concept of Relative Length of Life Inequality (RLI), which refers to the variation in LI between observations with a similar LE.

To gain insight into the importance of this pattern in our data, Figure 6.5 zooms in on it by providing information on length of life inequality and on premature deaths within LI quintiles, for two levels of average life expectancy: an LE of 65 (i.e. 64.5-65.5) and an LE of 75 (i.e. 74.5-75.5). The left panel shows that there is substantial variation in LI across observations (country-years). For a life expectancy of 65 (the top lines in the panel), length of life inequality varies between 0.139 and 0.179 for women (shown in red) and between 0.125 and 0.165 for men (shown in blue). For a life expectancy of 75, it varies between 0.079 and 0.111 for women, and between 0.084 and 0.110 for men. All these differences are significant at $p<0.001$, which makes clear that they are not negligible.

Figure 6.5. Length of life inequality and premature mortality for a given life expectancy



Note: Premature mortality by quintiles of length of life inequality. The lines in black refer to women, those in blue to men; solid and dotted lines refer to life expectancy of 65 and 75, respectively.
Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/enlkwb

To gain insight into the importance of this pattern in our data, Figure 6.5 zooms in on it by providing information on the variation in length of life inequality and premature deaths between countries with higher inequality and countries with lower inequality at the same level of LE. The left panel (A) shows how LI varies between the lowest, 2nd, 3rd, 4th and highest LI quintiles at an LE of 65 (64.5-65.5) and at an LE of 75 (74.5-75.5). For a life expectancy of 65 (the top lines in the panel), length of life inequality varies between 0.139 and 0.179 for women (shown in red) and between 0.125 and 0.165 for men (shown in blue). For a life expectancy of 75, it varies between 0.079 and 0.111 for women, and between 0.084 and 0.110 for men. All these differences are significant at p<0.001, which makes clear that they are not negligible.

A similar pattern is visible in the right panel of Figure 6.5, which shows premature mortality – as measured by the number of deaths per 1 000 people between age 15 and 50 – for the two life expectancy levels. At both levels of LE, premature mortality in this age group varies considerably between the lowest, 2nd, 3rd, 4th and highest LI quintiles. At an LE of 65, it varies between 168 and 297 deaths per 1 000 for women, and between 132 and 247 deaths per 1 000 for men. At a life expectancy of 75, it varies between 37 and 95 deaths per thousand for women, and between 43 and 94 deaths per thousand for men. Hence, in both cases, the most unequal quintile has substantially more premature deaths than the most equal quintile, both for women and for men. All differences are again significant at p<0.001.

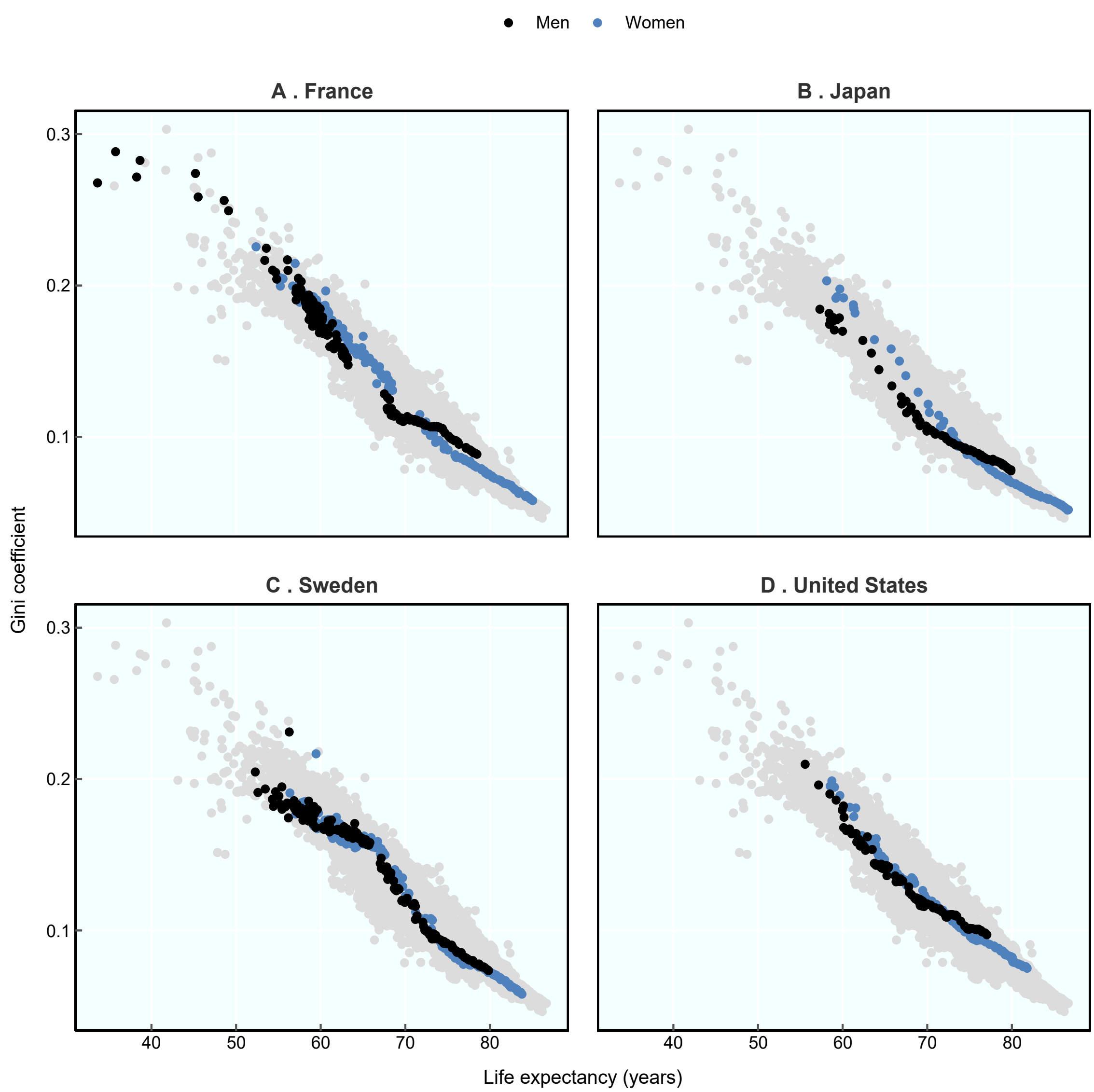
Trajectories of countries

An effective way to increase our understanding of the dynamics of inequalities in the age of death is to highlight the trajectories described by individual countries through the dotted areas in Figure 6.4, hence against the background of the LE-LI values for all other country-year combinations in our database. In Figure 6.6, this is done for four countries – the United States, France, Sweden and Japan. Each dot represents a gender-specific LE-LI value of a specific country in a specific year. Values for men are in blue, values for women in red.

The country-specific trajectories through the LE-LI space are not steadily decreasing chronological lines. During wars or epidemics, when LE may drop and LI increase, a country tends to jump in the upper-left direction. In the following years, when things go back to normal, the country may jump back to around its earlier position, to continue its path towards (on average) a higher LE and lower LI.

Figure 6.6 shows that during much of its history, the United States has been at an average level of LI given its LE, but that at the higher levels of LE (above an LE of 74) the country tended to move upward in the LE-LI space. France had a high level of length of life inequality for the lower values of average life expectancy, but it had lower inequality between LE 55 and LE 70. After LE attained levels of 75 for males and 80 for females, the decrease of LI decelerated, and the country has moved upwards in the LE-LI distribution.

Figure 6.6. The relationship between average life expectancy and length of life inequality in selected countries



Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/2o3018

Japan's trajectory is less clear-cut, as the trajectories of men and of women differ. Until an LE of about 74, inequality among Japanese men was relatively low, while among Japanese women it was much higher. After LE 74, the pattern reversed, with inequality among men being higher than among women. However, while for Japanese men the highest LE is associated with an average level of LI, Japanese women were at the upper end of the distribution when reaching their highest level of LE.

In the case of Sweden, for an LE between 50 and 60, both men and women enjoyed one of the lowest length of life inequalities. However, when LE grew beyond 60, the decrease in LI slowed down until LE 68, when Sweden was among the most unequal countries within this LE group. From LE 68 on, further increases in LE were associated with substantial decreases in LI, particularly for men. At the highest level of LE, Swedish men are therefore among the most equal populations in their LE group, whereas Swedish women are in the middle.

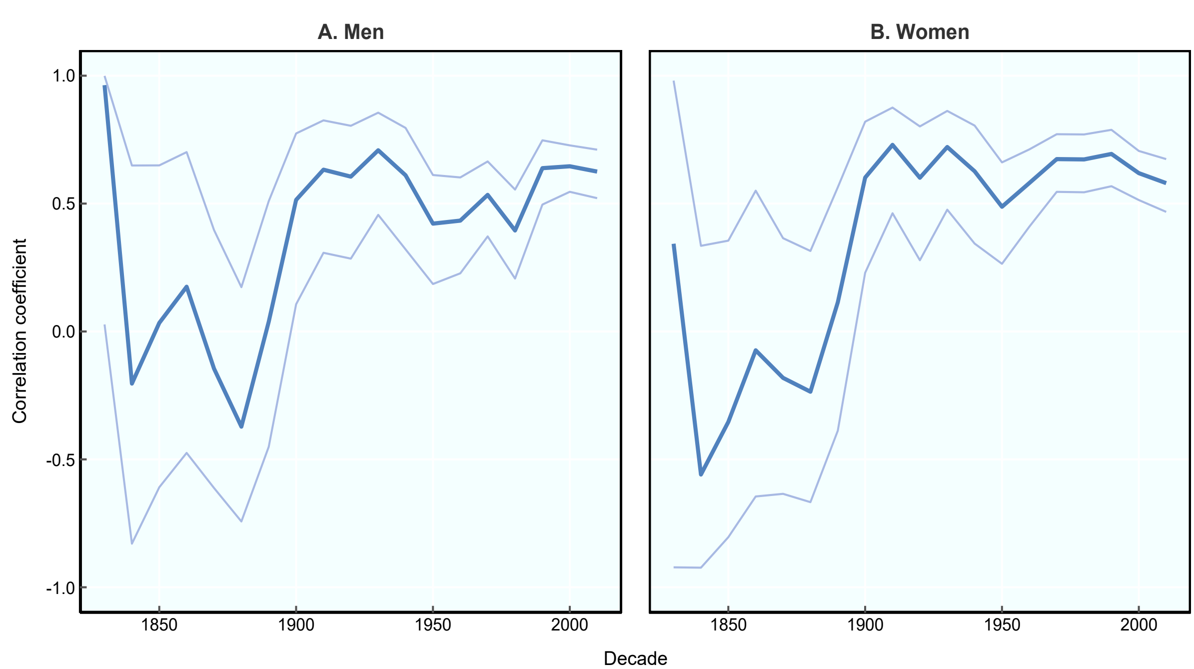
The comparatively high level of LI among women in times of low LE is most likely due to high maternal mortality rates associated with pregnancy and childbirth, which until the 1960s contributed substantially to premature mortality among women, even in the most developed countries. The strong reduction of this form of mortality most likely is responsible for the crossing of the trajectories of women and men somewhere around an LE of 70.

Correlations with GDP per capita

In this section, trends in life expectancy and length of life inequality are analysed alongside trends in GDP per capita. Correlation coefficients for each decade are shown in Figure 6.7 for life expectancy, and Figure 6.8 for length of life inequality, respectively.

Figure 6.7. Correlation between life expectancy among the population aged 15 and over and GDP per capita, 1800s-2010s

Pearson correlation coefficient



Note: Thin lines depict 95% confidence intervals. Coefficients calculated from decennial averages.
Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/0u73q9

Figure 6.7 shows that the correlation between average life expectancy and GDP per capita was weak during most of the 19th century, if it existed at all. A clear positive correlation emerges only around 1820-1830. The wide confidence intervals reflect the differences in data availability over time. After 1900, the positive correlation becomes stronger, on the order of 0.3-0.5, which means that countries with higher GDP per capita had higher levels of LE. This could indicate that higher life expectancy leads to more productivity, or that an increase in income allows individuals to live healthier – and consequently longer – lives.

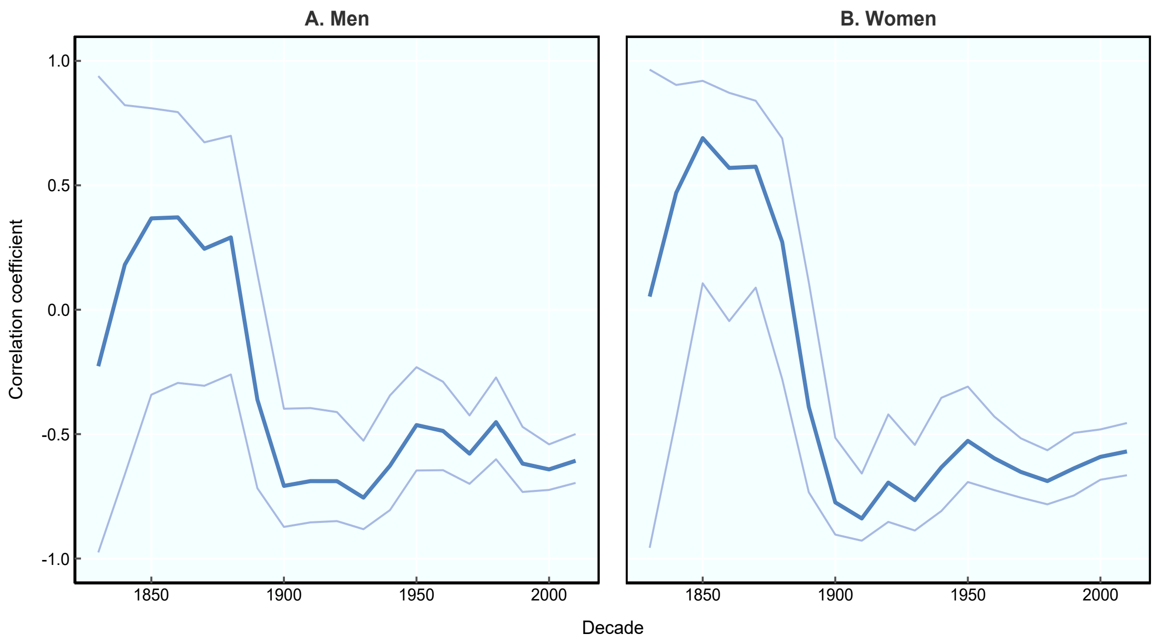
An additional explanation for the fact that the correlation between LE and GDP per capita was relatively strong in the 1910s-1950s period is World War I and World War II, as large drops in LE were accompanied by large drops in GDP per capita. The correlation then weakened up until the 1970s, reflecting a return to normal. The correlation then starts to strengthen again, up to the same high levels as during World War II. A possible explanation for this is the rise of the welfare state, allowing for increases in GDP per capita to lead to stronger increases in LE, as populations started to gain access to health care.

Figure 6.8 presents the correlation between length of life inequality and GDP per capita. As in the case of LE, we observe a shift in the nature of the correlation over time. During the 19th century, the correlation was mostly significant and positive, indicating that higher GDP was associated with higher levels of LI. This might be explained by the fact that, in the earlier phase of industrialisation, the GDP per capita of industrialising countries was growing substantially, but this came at the cost of the poorer health of the mass of workers, reflecting low labour standards and living in large cities with poor quality housing, bad sanitation and weak public health facilities (Fogel, 2004[22]; Steckel, 2001[23]).

From the 1890s onwards, the correlation turns negative, with more developed countries showing lower values of LI. This might reflect an improvement in labour conditions and public (health) services in these countries. The negative association peaked in the 1930s at a relatively strong level of -0.6, after which it decreased again to around -0.3 in the 1970s. Since then it has increased again to a value of about -0.6 in the 2010s. In other words, from the 1930s onwards, the association between GDP per capita and LI weakened, up until the 1970s. From the 1970s onwards, the association started to regain strength. This is in line with the observed pattern of the correlation between LE and GDP per capita. During the First and Second World Wars, LI increased strongly as a result of the loss of soldiers, while GDP per capita decreased strongly. The introduction of the welfare state in the 20th century may have translated part of the average increases in GDP per capita into better health for the more vulnerable part of the population, leading to a lower LI.

Figure 6.8. Correlation between length of life inequality among the population aged 15 and over and GDP per capita, 1800s-2010s

Pearson correlation coefficient



Note*:* Thin lines depict 95% confidence intervals. Coefficients calculated from decennial averages.
Source: (Global Data Lab[19]), *Length of Life Database*, https://globaldatalab.org/lengthoflife/.

StatLink https://stat.link/6olgsh

Priorities for future research

Life expectancy and length of life inequality are among the most fundamental indicators of the status of human societies. A low LE compared to what is potentially possible means that more people die than is necessary. Moreover, a high level of LI within a country means that too many people die at too young an age. Given that death is irreversible, and that the deceased cannot be compensated for their disadvantage, LE and LI are directly linked to the success or failure of societies in satisfying their citizens' basic needs. As such, these indicators should be among the first to consider when assessing the performance of societies.

Given the extremely high correlation between LE and LI, it is essential that, besides the absolute value of LI, its relative value, for a given value of life expectancy, is also considered. As LI more or less automatically decreases as countries achieve higher LE, the major question is not what the level of LI in a country is, but to what extent the country's level of LI is higher or lower than that of other countries at the same level of LE. As shown in Figure 6.5, the values of LI and the number of premature deaths can vary greatly between countries at similar levels of LE. A high level of LI, or premature deaths, at a similar level of LE suggests that not enough resources are being spent on improving the health situation of the younger and middle-aged population groups. As such, these inequality measures can provide a broad indication of distribution problems related to the population's health within a society.

The value of LE and LI as indicators of societies' performance is strengthened by their ease of construction. All one requires is information on the male and female population and deaths within age categories in a specific time-period. Hence, the required information is very factual and more easily observable than that for most other indicators, such as, for instance, (national) income or morbidity.

Given the fundamental character of indicators related to life and death, and the simplicity of their measurement, the registration of population data already started centuries ago in more industrialised societies. Official population registrations date back to the 19th century in many European countries, and even to the 1750s in Sweden and Finland. Given that over the whole period since then the same data has been collected, these registrations offer us the possibility to look further back in time than is possible with data for most other indicators.

This does not mean, however, that the population-based measures presented in this chapter are without problems. Besides the obvious measurement issues associated with historical data, there is a more fundamental problem related to the use of these measures to indicate societies' performance: They suffer from a time lag, as their current values are determined by the situation and behaviour of people in the past. Food scarcity during childhood or smoking behaviour during young adulthood influences old-age mortality that is used for computing LE for later-born cohorts.

This problem to a certain extent also exists for LI, as determining inequality in length of life involves a comparison between mortality in older age groups with mortality at a younger age. However, given that a relatively high level of LI at a certain level of LE indicates that more people die at too young an age, this indicates that not enough resources are being spent on improving the health situation of the younger and middle-aged population groups. Hence, by combining LI with LE, a broad indication of distribution problems related to population health within a society can be obtained.

We believe that even more is possible, and that substantial progress can be achieved without great additional data collection efforts. The life tables used for constructing LE and LI contain more information that can be used for this purpose. They provide, for each five-year age category, and for men and women separately, the group-specific mortality rates. These mortality rates can be compared between different age groups within the same country, as well as between the same age groups in different countries and years. In this way, we are able to observe – very specifically and currently – which age groups in which countries and years do worse – or better – than would be expected based on what is observed in other situations.

As a priority for future research, we therefore recommend developing a dashboard based on these age-specific mortality rates, with which the situation and life chances of men and women in the different age groups can be monitored. The life table database underpinning this chapter offers ample possibilities for conducting cross-national and cross-temporal comparative research to validate such a dashboard by determining which societal situations are associated with deviant patterns in the different phases of life.

References

Bloom, D., D. Canning and J. Sevilla (2003), *The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change*, RAND, http://dx.doi.org/10.7249/MR1274. [9]

Edwards, R. and S. Tuljapurkar (2005), "Inequality in life spans and a new perspective on mortality convergence across industrialized countries", *Population and Development Review*, Vol. 31/4, pp. 645-674, http://dx.doi.org/10.1111/j.1728-4457.2005.00092.x. [8]

Fogel, R. (2004), *The Escape from Hunger and Premature Death, 1700-2100: Europe, America and the Third World*, Cambridge University Press, http://dx.doi.org/10.1017/CBO9780511817649. [22]

Global Data Lab (n.d.), *Length of Life Database*, https://globaldatalab.org/lengthoflife/ (accessed on 14 January 2020). [19]

HMD (n.d.), *The Human Mortality Database*, https://www.mortality.org/ (accessed on 14 January 2020). [16]

Kontis, V. et al. (2017), "Future life expectancy in 35 industrialised countries: Projections with a Bayesian model ensemble", *The Lancet*, Vol. 389/10076, pp. 1323-1335, http://dx.doi.org/10.1016/S0140-6736(16)32381-9. [24]

Maddison, A. (2001), *The World Economy: A Millennial Perspective*, Development Centre Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264189980-en. [1]

Marmot, M. (2005), "Social determinants of health inequalities", *The Lancet*, Vol. 365/9464, pp. 1099-1104, http://dx.doi.org/10.1016/S0140-6736(05)71146-6. [11]

Mathers, C. et al. (2015), "Causes of international increases in older age life expectancy", *The Lancet*, Vol. 385/9967, pp. 540-548, http://dx.doi.org/10.1016/S0140-6736(14)60569-9. [25]

Németh, L. (2017), "Life expectancy versus lifespan inequality: A smudge or a clear relationship?", *PLoS One*, Vol. 12/9, http://dx.doi.org/10.1371/journal.pone.0185702. [15]

Oeppen, J. and J. Vaupel (2002), "Broken limits to life expectancy", *Science*, Vol. 296/5570, pp. 1029-1031, http://dx.doi.org/10.1126/science.1069675. [2]

Pradhan, M., D. Sahn and S. Younger (2003), "Decomposing world health inequality", *The Journal of Health Economics*, Vol. 22/2, pp. 271-293, http://dx.doi.org/10.1016/S0167-6296(02)00123-6. [5]

Seaman, R., A. Leyland and F. Popham (2016), "Increasing inequality in age of death at shared levels of life expectancy: A comparative study of Scotland and England and Wales", *SSM - Population Health*, Vol. 2, pp. 724-731, http://dx.doi.org/10.1016/j.ssmph.2016.10.001. [14]

Sen, A. (1990), "More than 100 million women are missing", *New York Review of Books*, pp. 61-66, https://www.nybooks.com/articles/1990/12/20/more-than-100-million-women-are-missing/. [20]

Shkolnikov, V. (2017), "Methods protocol for the human life-table database", *Working Paper*, https://www.mortality.org/Public/Docs/MethodsProtocol.pdf. [18]

Shkolnikov, V., E. Andreev and A. Begun (2003), "Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples", *Demographic Research*, Vol. 8/11, pp. 305-358, http://dx.doi.org/10.4054/DemRes.2003.8.11. [12]

Smits, J. and C. Monden (2009), "Length of life inequality around the globe", *Social Science & Medicine*, Vol. 68/6, pp. 1114-1123, http://dx.doi.org/10.1016/j.socscimed.2008.12.034. [6]

Solt, F. (2019), "Measuring income inequality across countries and over time: The standardized world income inequality database", *Social Science Quarterly*, Vol. 101/3, pp. 1-23, https://doi.org/10.1111/ssqu.12795. [21]

Steckel, R. (2001), *Industrialization and health in historical perspective*, Oxford University Press, http://dx.doi.org/10.1093/acprof:oso/9780192631961.001.0001. [23]

UN (2019), *World Population Prospects 2019: Data Booklet*, Statistical Papers - United Nations (Ser. A), Population and Vital Statistics Report, United Nations, New York, https://dx.doi.org/10.18356/3e9d869f-en. [26]

UNDP (2017), *"Human Development Report 2016"*, United Nations Development Programme, http://dx.doi.org/10.18356/6d252f18-en. [4]

van Zanden, J. et al. (eds.) (2014), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-en. [7]

Vaupel, J. et al. (1998), "Biodemographic trajectories of longevity", *Science*, Vol. 280/5365, pp. 855-860, http://dx.doi.org/10.1126/science.280.5365.855. [3]

WHO (2018), *"WHO methods and data sources for life tables 1990-2016"*, World Health Organization, https://www.who.int/healthinfo/statistics/LT_method.pdf. [17]

Wilmoth, J. and S. Horiuchi (1999), "Rectangularization revisited: Variability of age at death within human populations", *Demography*, Vol. 36/4, pp. 475-495, http://dx.doi.org/10.2307/2648085. [13]

Wise, D. (ed.) (2004), *"Changes in the age distribution of mortality over the twentieth century"*, University of Chicago Press. [10]

Annex: 6.A. Inequality measurement

Length of life inequality (LI) is measured by computing the Gini coefficient over the distribution of age at death based on the population aged 15 and over. The distributions of age at death were obtained by applying the age- and sex-specific mortality rates from the life table to a population of 100 000 individuals aged 15, thus standardising for differences in adult population structure among countries and time periods. The Gini coefficient varies between zero and one, with zero indicating a situation of maximum equality (everybody has the same age at death) and one indicating maximum inequality. The choice of the inequality measure for computing length of life inequality is not a very critical one. There is a high correlation (generally over 0.90) among inequality measures computed over the distribution of age at death, and the use of different inequality measures leads to similar results (Wilmoth and Horiuchi, 1999[13]).

Note

[1] Recent sources confirm a projected increase in the limits of LE, albeit at a potentially slower pace (Kontis et al., 2017[24]; UN, 2019[26]). The main causes for these increases are reduced tobacco usage and a reduction in cardiovascular disease mortality (Mathers et al., 2015[25]). There is, however, large variation among developed countries in terms of LE trends. For instance, the United States, which has a relatively low level of LE when compared to other developed countries, is projected to face declines in LE (Kontis et al., 2017[24]).

7 Inequality in educational achievement

Bas van Leeuwen and Jieli Li
International Institute of Social History, Amsterdam

Education is far from universally attainable, with the resulting educational inequality having wide implications for both individuals and societies. Because of this importance, this chapter reports trends in educational inequality from the 19th century to the present. We do this by using relative and absolute measures of inequality in (formal) educational attainment. Overall, we observe a strong decline in the Gini coefficient of years of schooling over the period, a decline that is caused mostly by a reduction in the share of persons with no formal education. Looking at absolute inequality, as measured by the standard deviation, we find a rising trend in the first half of the 20th century and a decline afterwards.

Introduction

As noted recently by the OECD Secretary-General Angel Gurría, “the dream of ‘quality education for all’ is still far off”.[1] Indeed, even today, education is far from being universally available, with educational inequality having wide implications for both individuals and societies. First, educational differences translate into inequalities in a range of other well-being outcomes, such as earnings, longevity, morbidity and subjective well-being.[2] Second, higher educational inequality may also reduce benefits for society as a whole, for example, a lower demand for public goods such as health care, threats to the rule of law, a lower investment in R&D and the adoption of new technologies, or higher crime rates (Card and DiNardo, 2002[1]). A third rationale for reducing educational inequality is equity and equality of opportunity. Access to education must be fair, i.e. not dependent on personal characteristics, and inclusive, i.e. with the same rules applying to all (Field, Kuczera and Pont, 2007[2]).[3]

Educational inequality thus has significant effects on both individuals and societies. This explains the wide scholarly interest in the topic. Many studies such as Meyer (2011[3]) have shown educational inequality in a country to be strongly related to the historical development of its education system.[4] This is one reason why this chapter documents trends in educational inequality from the 19th century to the present. We do this, following the literature, by using relative and absolute measures of inequalities in the completion of formal education. After describing concepts, sources and data quality, we present the main features identified from the historical evidence. Overall, our evidence shows a strong decline in the Gini coefficient of educational attainment, which is caused mostly by a reduction in the share of people without any formal education. Measures of absolute inequality show an increase in the first half of the 20th century and a decline afterwards. We conclude by briefly discussing the link between educational inequality and GDP per capita, and by highlighting some priorities for future historical research.

Description of the concepts used

Education, defined as the facilitation of learning, has varied strongly across countries and over time. In order to enhance international comparability, an International Standard Classification of Education (ISCED) was created in 1976. Even though this prompted some standardisation of educational statistics, it has not been fully implemented by all countries. Even though most countries are currently using the 2011 ISCED classification when reporting their educational statistics, the basic definitions remain the same as those used in previous versions, with formal primary education starting roughly from age 5-7, with a duration of around 6 years, and offering basic reading, writing and arithmetic; secondary education, covering lower and higher secondary schools, also commonly lasting 6 years in total; and tertiary education usually lasting around 3 or 4 years (UNESCO Institute for Statistics, 2012[4]). As a continuous measure is needed for the purpose of computing educational inequalities, some scholars have mapped these layers into mean years of schooling, e.g. Morrisson and Murtin (2009[5]) and Lee and Lee (2016[6]).

A commonly used measure of inequality is the Gini coefficient, which has the property of being independent of mean and population size and it is easy to interpret, ranging between 0, in the case of perfect equality, and 1 (or 100 if reported as an index) in the case of perfect inequality. For those reasons, this measure is widely used in the literature on educational inequality, e.g. Lopez, Thomas and Wang (1999[7]), Thomas, Wang and Fan (2003[8]); Földvári and Leeuwen (2011[9]).[5] Other authors prefer alternative measures for two main reasons.

First, being a relative measure, the Gini coefficient does not capture the scale effect of higher educational attainment: if, for every individual, the number of years of education doubles, the Gini coefficient does not change even as the absolute difference in the number of years in education between any two individuals doubles. Whether this is important or not depends on how one views inequality. If one considers inequality as intrinsically relative (e.g. person A has 50% more years of education than person B), then the Gini coefficient is the preferred measure; conversely, if one considers that what matters are absolute differences between people (e.g. person A has 5 years more education than person B), then an absolute measure such as the standard deviation is preferable. As pointed out by Amiel and Cowell (1999[10]) even though a sizeable group of respondents to surveys considers absolute differences as important, most respondents consider inequality as a relative concept.

Second, the Gini coefficient declines when the average number of years of education increases. That is, changes in the value of the Gini coefficient for education are, to a large extent, driven by the declining share of persons without formal education (Castelló and Doménech, 2002[11]; Morrisson and Murtin, 2009[5]). On the contrary, absolute measures of inequality[6] – which include the absolute Gini (the Gini index times mean years of education) (Jenkins and Jäntti, 2005[12]; Niño-Zarazúa, Roope and Tarp, 2017[13]; Bandyopadhyay, 2018[14]) and the standard deviation (Ram, 1990[15]; Meschi and Scervini, 2014[16]) – record the absolute dispersion (in years of education) of education among individuals in a society. With absolute measures, when more people complete basic education, inequality will initially rise but, when more than 50% of the population has achieved basic education, additional persons with education will cause inequality to decline. For those reasons, absolute measures of inequality, such as the standard deviation, have been used alongside the Gini index to provide a more complete picture of educational inequality, e.g. Ram (1990[15]).

Historical sources

Because of a lack of systematic historical censuses, educational inequality is often calculated based on datasets of educational attainment (the highest level of formal education completed by each person), converted into a continuous measure of "years of schooling". There are a great number of such databases, e.g. Cohen and Soto (2007[17]), Barro and Lee (2013[18]), de la Fuente and Doménech (2015[19]), and Samir and Lutz (2017[20]), but these data refer mostly to the period after 1950, with two datasets providing projections for the future (Barro and Wha Lee, 2015[21]; Samir and Lutz, 2017[20]). Yet, there are fewer datasets covering historical periods, e.g. Morrisson and Murtin (2013[22]); Van Leeuwen and Li (2014[23]); Tamura et al (2019[24]); and Lee and Lee (2016[6]). Only two datasets have the advantage of covering a broad range of countries spanning back to 1870. The first is the one presented by van Leeuwen and Li (2014[23]), which is based on a combination of the dataset by Morrisson and Murtin (2009[5]) and by Földvári and van Leeuwen (2014[25]); the second is the one from Lee and Lee (2016[6]).

Choosing which dataset to use is complicated since all have their merits. An advantage of the Van Leeuwen and Li dataset is that it is annual, while the Lee-Lee data are quinquennial. Conversely, Lee and Lee (2016[6]) provides data for men and women separately. A further difference is that the Van Leeuwen and Li dataset covers the population aged 15 and older, while the Lee-Lee data covers the population aged 15-64. This obviously causes some differences in trends, but levels of both databases have very high correlations of close to 96%. For these reasons, both databases are often used in historical research.

Yet, there are also more critical observations about both datasets. For example, de la Fuente and Doménech (2015[19]) note that, for OECD countries after 1950, the reliability of the Barro-Lee database, on which the Lee-Lee data is partially based, ranks considerably below other available series. Conversely, Lee and Lee (2016[6]) argue against the use of literacy data in the Van Leeuwen and Li dataset to gauge educational attainment in all types of educational institutions (i.e. both public and private), preferring to limit their remit to public schools. Yet, since private schools make up a significant portion of schools both in Europe and in the non-Western world, using data limited to public schools implies a significant underestimate of attainment when going back in time, leading to higher estimates of inequality. For example, the share of the population over age 15 with some attainment (irrespective of the level) in 1870 in China is 24.8% in the Van Leeuwen and Li dataset and 0% in Lee and Lee. To assess the reliability of both estimates, one may look at Chinese enrolment ratios for 1870. These vary between close to 0% (for

Yugan county in Jiangxi province) to over 40% (for Chongning county in Sichuan province), with an average of close to 20% (Yugan, 1947[26]; Cao, 2013[27]), thus providing evidence in favour of the higher estimate of educational attainment reported by Van Leeuwen and Li. Similar observations apply to other countries. These differences in mean years of schooling have significant effects on inequality estimates: whereas, using the Van Leeuwen and Li data, the global Gini coefficient is 79.0 in 1870 (Table 7.2), Lee and Lee provide a measure of 90.5. Because of these considerations, the dataset from Van Leeuwen and Li is used in this chapter to calculate educational inequality over the period 1870-2010.

Data quality

Our assessment of the quality of the educational inequality estimates is based on the method used to collect the underlying data, sources and indicators; we rate this quality as high quality (1), medium quality (2), moderate quality (3) or low quality (4). An assessment of the quality of data on mean years of education was already reported in van Leeuwen and Li (2014[23]). The data used for this chapter, being a slightly improved version of those in van Leeuwen and Li (2014[23]), are quite comparable.

The quality measures are reported in Table 7.1. Starting with the data from 1950 to 2010, these are sourced from Van Leeuwen and Li, which for the modern period are based on direct education data and for earlier years on back-projection. Starting with the most recent period, comparison of these attainment-based data with the highest quality level individual-level data available from the International Social Survey Programme (ISSP) Research Group (2017[28]) results in a strong correlation. Since these ISSP data are high quality, but not formally produced by an official statistical agency, we assess data quality at level 2 for all world regions since 1980, as well as for Europe, the Western Offshoots, and Latin America and the Caribbean for the period between 1950 and 1980. Varying by country, most data between 1870 and 1950 are based on a backward extrapolation of censuses using various data such as population and enrolments, which deteriorate in quality the further one goes back in time; hence, these are assessed at level 3.

Table 7.1. Quality of data on educational inequality

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Sub-Sahara Africa	Middle East and North Africa	East Asia	South and Southeast Asia
1870	3	3	3	3	3	3	4	3
1913	3	3	3	3	3	3	3	3
1950	2	2	2	2	3	3	3	3
1980	2	2	2	2	2	2	2	2
2010	2	2	2	2	2	2	2	2

Note: High quality: the product of official statistical agency (national or international); 2. Medium quality: the product of economic-historical research using the same sources and methods as applied by official statistical agencies; 3. Moderate quality: economic historical research, but making use of indirect data and estimates; and 4. Low quality: estimates based on a range of proxy information. In case of multiple sources, the lowest quality source is given.

Main highlights of trends in educational achievement

We start by reporting global trends in educational inequality. Table 7.2 shows that the Gini coefficient of years of schooling at the world level decreases over time. This decline mainly reflects a fall in the share of people without any formal education, from around 73% in 1870 to 43% in 1950 and 15% in 2010. The standard deviation points to increasing world inequality.

Table 7.2. World educational inequality, 1870-2010

	Gini index	Standard Deviation
1870	79.0	1.06
1890	73.9	1.22
1910	67.9	1.41
1930	61.5	1.65
1950	54.1	1.93
1970	43.8	2.49
1990	36.1	2.85
2010	29.6	3.03

Sources: (van Leeuwen and van Leeuwen-Li, 2014[23]), "Education since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en; (Morrisson and Murtin, 2009[5]), "The century of education", *Journal of Human Capital*, Vol. 3/1, pp. 1-42, https://doi.org/10.1086/600102.

StatLink https://stat.link/utgcrm

Yet, this global pattern obscures the trends for the various regions of the world. As shown in Table 7.3, the Gini coefficient of educational inequality declines in all regions, but more so for less developed regions. For the standard deviation, regions with lower education levels (all regions except Western Europe and the Western Offshoots) experienced rising inequality up to around 1990. Western Europe and the Offshoots, however, show broadly constant standard deviations.

It is important to note that, even though for both measures of inequality trends among all world regions are roughly similar, this does not apply to their levels. Whereas in terms of the Gini coefficient, the Middle East and North Africa, sub-Saharan Africa, and Latin America and the Caribbean display the highest levels of inequality, for the standard deviation prior to 1950 Western Europe and the Western Offshoots had the highest levels of inequality, arguably because of the scale effect of higher average levels of education.

One question is whether cross-country differences in educational inequality can be attributed either to structural differences in countries' education systems or simply to lagging development in mean years of schooling. To test this, Figure 7.1 looks at whether some countries have a different Gini index when they reach a certain level of mean years of education later than other countries. We find no evidence of rising or declining inequality for the (mostly developed) countries that reached an average of 5 years of schooling before the 1950s, suggesting that, within this group of countries, differences in inequality are mainly due to a lagging educational development. Even though there is no strong effect, we do see some evidence of higher educational inequality among the (mostly less developing) countries that reached an average of 5 years of schooling after around 1950, thus suggesting that countries experiencing educational expansion only recently featured increasing structural inequalities in their education systems.

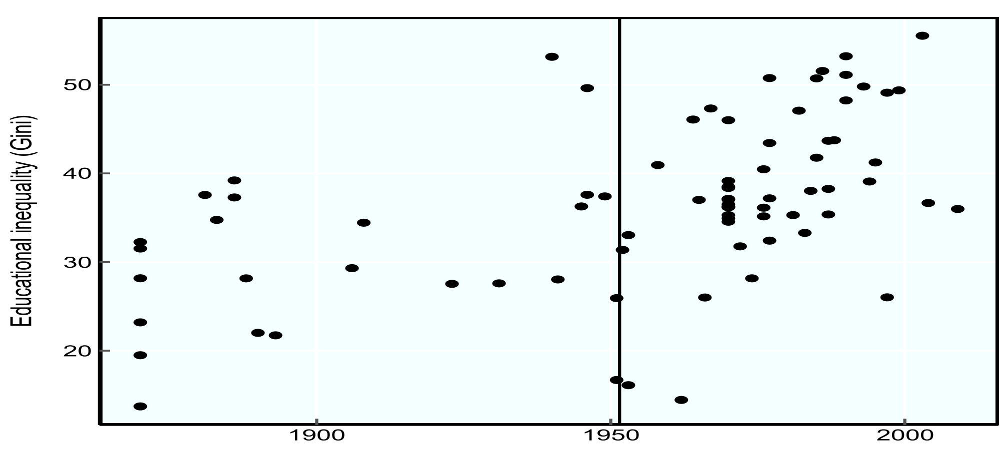
Table 7.3. Educational inequality by major world regions, 1870-2010

		Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Sub-Saharan Africa	Middle East and North Africa	East Asia	South and Southeast Asia
1870	Gini Index	50.1	82.9	32.0	88.5	97.1	96.1	77.4	98.2
	Standard Deviation	2.2	1.2	2.7	1.0	0.2	0.5	1.2	0.1
1890	Gini Index	36.3	77.9	24.5	82.9	95.8	95.3	74.5	97.7
	Standard Deviation	2.1	1.4	2.4	1.3	0.2	0.7	1.3	0.5
1910	Gini Index	29.8	69.5	22.1	76.9	94.7	93.8	69.6	94.6
	Standard Deviation	2.2	1.5	2.5	1.6	0.3	0.8	1.4	0.7
1930	Gini Index	26.0	48.2	19.6	69.0	91.9	92.4	66.2	90.2
	Standard Deviation	2.2	2.0	2.5	2.0	0.4	1.0	1.5	1.2
1950	Gini Index	23.3	26.0	17.3	53.4	85.5	87.7	59.9	83.6
	Standard Deviation	2.3	2.0	2.5	2.3	0.9	1.5	1.9	1.7
1970	Gini Index	17.6	24.8	14.0	45.0	75.0	75.9	37.9	71.5
	Standard Deviation	2.5	2.5	2.5	2.9	1.9	2.4	2.4	2.5
1990	Gini Index	11.4	19.4	9.7	32.3	58.0	54.0	31.9	56.0
	Standard Deviation	1.7	2.7	2.0	3.4	2.7	3.5	2.9	3.2
2010	Gini Index	9.9	18.2	8.1	24.1	41.9	38.6	24.4	44.2
	Standard Deviation	1.7	2.6	1.8	3.1	2.8	3.8	3.0	3.6

Source: (van Leeuwen and van Leeuwen-Li, 2014[23]), "Education since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en.

StatLink https://stat.link/7wmxy0

Figure 7.1. Gini index of educational inequality and year when the educational attainment of the total population in each country reaches a level of 5 years of schooling



Source: (van Leeuwen and van Leeuwen-Li, 2014[23]), "Education since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en.

StatLink https://stat.link/9mscjr

These patterns of less equal education systems in less developed countries have to be viewed in a historical perspective. Indeed, the highest literacy rates (implying a lower Gini index) in the 19th century occurred in Western Europe and the Western Offshoots. These high literacy rates reflected the existence of an education system that was linked to economic, religious and political developments of individual countries (Fuller, Edwards and Gorman, 1986[29]; Boli, 1989[30]). A very different pattern prevailed in former colonies, whose education system was often partly modelled on the coloniser's conditions rather than locally prevailing conditions (Bolt and Bezemer, 2009[31]). The importance of this factor can be seen when examining why the Gini index in former French colonies is higher than that of former British colonies (Garnier and Schafer, 2006[32]). For example, Cogneau and Moradi (2014[33]) try to isolate this effect of the coloniser country and conclude that the lower educational Gini of former British colonies can be attributed to the more flexible British system based on missionary schools (White, 1996[34]), which connected better to local society than the highly regulated French system and hence resulted in higher literacy rates (Cogneau and Moradi, 2014[33]).

This historical persistency does not mean that the educational systems of countries are always inexorably tied to their past. Educational development depends, to a large extent, on government policy. As pointed out in the introduction, reducing educational inequality can deliver both private benefits (benefitting the individual who invested in education) and social returns (accruing to society as a whole), such as enhancing equity and opportunity, and the priority attached to each goal can change over time. After World War II, capitalist countries generally pursued an educational policy aimed at enhancing private returns to education by stimulating post-primary education. Other countries, such as socialist countries, often started mass literacy campaigns based on the expansion of basic education, which enhanced social returns and furthered equity (Stites and Semali, 1991[35]). In China, for example, Mao's education drive between 1950 and 1970 resulted in a reduction of the share of persons with no education from 61% to 32%, causing a drop in the Gini coefficient of East Asia over this period. Likewise, Dupraz (2019[36]) studied the effect of the colonial legacy for Cameroon, which was divided into English and French parts, but reunited after independence. Using this as a natural experiment, he found that, even though people born in the English part of the country had, on average, one year more schooling than those born in Cameroon's regions under French rule, this difference disappeared after World War II due to a policy of stimulating schools in the French part. It has to be stressed though that, rather than conforming to an ideal-type, policies are frequently mixed or changing over time, e.g. Samoff (1991[37]). For example, in the 1950s-1970s, various socialist countries moved away from their focus on social returns and equity and prioritised increasing private returns instead.

Table 7.4. Gini index of educational inequality, 25 major countries, 1870-2010

	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1870	37.3	23.2	47.9	28.2	83.2	79.1	37.8		79.5	50.8	19.5	32.3	89.4	92.5	74.5	98.3	96.4	95.0	99.8	81.3	77.2	79.8	98.1	99.3	95.8
1880	23.5	21.0	37.7	25.9	73.2	71.9	32.8		77.3	42.3	15.9	29.6	86.0	91.5	71.9	98.3	95.4	93.0	99.8	81.3	76.4	77.3	97.9	99.5	95.0
1890	23.4	19.0	32.0	21.4	62.2	62.9	27.0		75.0	38.2	15.9	24.4	82.5	90.5	69.2	97.4	94.3	91.1	99.8	81.3	75.0	68.7	97.8	99.5	93.9
1900	16.3	16.7	29.7	20.6	52.4	56.7	25.3		72.2	28.1	17.7	21.6	79.1	89.2	64.9	96.5	93.5	89.2	99.8	81.3	73.5	55.9	95.4	99.3	92.5
1910	16.1	14.4	26.4	23.1	44.2	51.6	22.7		70.6	24.9	18.4	22.3	76.0	87.3	61.6	94.5	92.5	87.3	99.8	80.9	72.2	44.0	94.4	99.0	91.7
1920	13.5	13.3	24.5	19.1	36.2	48.9	22.5	70.5	58.7	20.0	18.8	21.3	73.8	84.6	55.7	94.8	91.8	85.3	99.8	74.7	70.8	32.8	92.8	98.3	90.6
1930	14.5	13.6	23.5	21.7	31.0	50.1	22.8	76.5	41.9	16.8	18.4	19.9	71.0	82.6	52.1	91.4	91.1	84.6	99.1	67.8	70.6	28.2	90.4	97.0	88.9
1940	16.2	14.9	22.0	18.1	26.7	36.0	22.5	71.7	15.0	15.8	18.0	19.0	63.7	77.4	42.5	89.3	90.2	82.7	95.2	60.7	68.2	24.0	87.9	94.6	76.6
1950	17.4	16.0	24.7	17.8	26.5	33.9	20.5	70.7	16.0	14.3	17.1	17.4	52.0	66.2	32.0	85.3	86.2	74.9	91.8	59.5	65.7	23.9	85.6	88.8	58.6
1960	18.3	18.0	23.9	15.4	23.4	26.1	18.9	58.7	19.5	14.7	18.9	16.2	47.3	66.8	29.9	84.8	67.4	73.2	79.5	57.2	56.6	18.0	83.5	75.5	50.4
1970	15.0	19.3	16.7	10.9	24.6	20.9	18.3	41.7	21.0	14.7	20.1	13.3	31.1	60.4	28.0	84.2	54.3	63.4	77.2	54.0	41.7	14.1	78.7	60.1	43.9
1980	8.6	18.3	11.4	4.0	22.6	16.2	14.9	32.5	18.9	9.5	17.1	11.6	31.7	51.4	29.4	71.8	52.7	49.7	76.8	48.9	41.1	11.8	73.6	49.8	32.1
1990	3.8	16.8	9.8	4.4	17.7	19.8	11.5	29.8	17.7	6.3	14.0	9.3	26.4	36.0	25.3	51.1	42.7	32.4	65.6	41.5	35.3	9.9	64.5	33.2	24.8
2000	4.9	15.4	9.8	5.2	14.3	21.8	11.9	24.7	18.9	5.1	10.7	8.9	23.3	28.8	23.2	39.3	35.8	21.6	47.5	32.5	30.1	9.2	52.9	25.7	19.1
2010	5.1	14.5	6.9	5.2	10.5	21.2	9.6	26.4	17.4	5.3	9.9	8.1	21.8	24.4	21.3	30.4	32.6	16.8	46.0	29.5	27.7	8.0	51.0	22.3	19.8

Source: (van Leeuwen and van Leeuwen-Li, 2014[23]), "Education since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en.

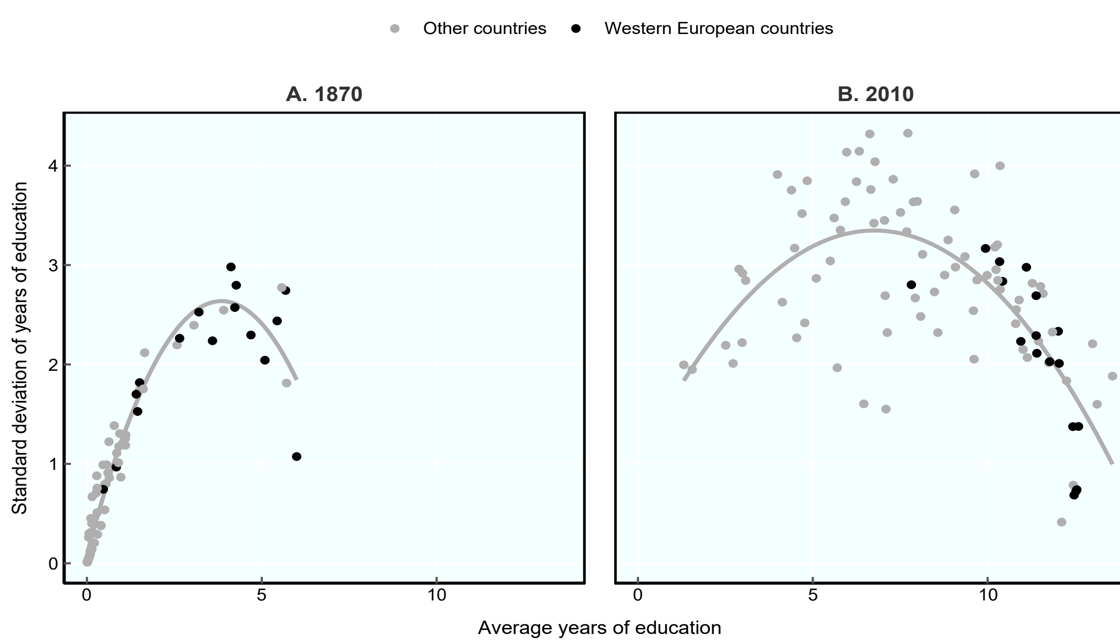
StatLink https://stat.link/xz1emv

Large declines in the Gini index of educational inequality occurred in all world regions. Yet, as pointed out before, this does not apply to the standard deviation, as higher shares of persons with some education imply higher standard deviations up to the point where more than 50% have some level of education, after which the standard deviation declines again. These trends are often captured in the literature by the notion of an educational Kuznets curve (Ram, 1990[15]; Thomas, Wang and Fan, 2003[8]; Castelló and Doménech, 2002[11]). This notion implies that, when a country moves from zero to maximum education, it will first experience an increase in inequality and later a reduction (Ram, 1990[15]), creating an inverse U-shaped curve (see Figure 7.2). In 1870, the United States, with a standard deviation of 2.77 and an average of 5.6 years of schooling, was on the declining path of this curve, although inequality there was still higher than in most developing countries, which remained on the path of rising educational inequality. This inflection point for educational inequality had however significantly increased by 2010. As found by Ram (1990[15]) and Thomas, Wang and Fan (2003[8]), the standard deviation in years of schooling peaks at around 7 years of education, i.e. a level around 3.5 years higher than in 1870. In 1870, only Western Europe (with the exception of some southern European countries) was on the declining portion of the curve. Hence, increasing educational attainment in Western Europe has, since 1870, led to lower inequality when measured by the standard deviation.

Looking at other world regions, various countries from sub-Saharan Africa were in 1870 on the increasing part of the curve, and remain there even today. In terms of policy, an absolute measure of educational inequality such as the standard deviation implies that less developed countries face a trade-off between increasing the mean years of education and reducing inequality, while developed economies, whose mean years of schooling are well above 7, can, by investing further in education, reap both benefits at the same time.

No such trade-off exists when considering a relative measure of educational inequality such as the Gini index. Yet, this does not mean that countries that have passed the 7-year threshold do not face choices in the structure of their educational investments: investments in tertiary education are more likely to increase inequality, even expressed in the Gini index, than investments in basic education.

Figure 7.2. Mean years of education and standard deviation, 1870 and 2010



Note: Data for Western European countries are shown in black. See https://clio-infra.eu/.

StatLink https://stat.link/mrw60d

Table 7.5. Standard deviation of years of schooling for selected countries, 1870-2010

	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1870	2.24	2.04	2.98	2.44	0.97	1.82	2.57		1.01	2.39	1.81	2.77	0.99	0.99	1.74	0.67	0.71	0.20	0.01	1.19	1.20	0.87	0.08	0.26	0.31
1880	1.89	1.91	2.90	2.33	1.28	2.17	2.48		1.09	2.75	1.36	2.78	1.14	1.10	1.81	0.79	0.92	0.28	0.01	1.19	1.24	1.15	0.29	0.26	0.43
1890	2.01	1.76	2.88	1.96	1.57	2.38	2.22		1.16	3.16	1.28	2.48	1.25	1.21	1.88	0.85	0.86	0.35	0.01	1.19	1.26	1.42	0.48	0.26	0.45
1900	1.56	1.55	2.93	1.87	1.86	2.61	2.23		1.20	2.84	1.63	2.29	1.45	1.31	2.08	0.95	1.08	0.43	0.01	1.19	1.29	1.70	0.65	0.26	0.47
1910	1.58	1.27	2.88	2.43	2.03	2.67	2.07		1.29	2.84	1.75	2.52	1.56	1.41	2.41	1.06	1.12	0.50	0.01	1.20	1.31	2.00	0.77	0.37	0.56
1920	1.20	1.08	2.88	1.99	2.11	2.73	2.13	2.26	1.70	2.47	1.89	2.52	1.61	1.54	2.69	1.18	1.19	0.56	0.01	1.46	1.33	2.21	1.07	0.46	0.58
1930	1.28	1.14	2.84	2.53	2.23	2.81	2.25	2.82	1.75	2.19	2.02	2.53	1.70	1.71	2.92	1.57	1.26	0.63	0.04	1.76	1.38	2.42	1.24	0.66	0.68
1940	1.44	1.27	2.76	2.03	2.25	2.67	2.30	3.11	1.43	2.22	2.13	2.40	1.95	2.00	2.98	1.80	1.56	0.69	0.20	2.05	1.49	2.32	1.48	0.91	1.17
1950	1.67	1.45	3.24	2.05	2.46	2.70	2.16	3.21	1.61	2.25	2.21	2.52	2.13	2.32	2.80	2.41	1.94	1.05	0.36	2.64	1.74	2.51	1.73	1.21	1.68
1960	2.14	1.80	3.34	2.38	2.33	2.42	2.05	3.32	1.85	2.36	2.48	2.57	2.50	3.04	2.91	2.70	2.37	1.99	1.16	3.31	2.21	2.14	2.05	1.72	2.17
1970	2.41	2.36	2.97	2.46	2.58	2.21	2.55	3.18	2.21	2.51	2.81	2.45	2.38	3.37	2.99	2.96	2.76	2.54	1.38	3.76	2.41	2.38	2.64	2.48	2.41
1980	1.95	2.66	2.37	0.54	2.70	1.71	2.61	3.10	2.57	1.97	2.94	2.32	2.91	3.60	3.50	3.71	3.32	2.93	1.60	3.67	2.87	2.31	3.14	2.77	2.71
1990	0.53	2.73	2.09	0.60	2.53	2.26	2.31	3.03	2.58	1.01	2.80	1.96	2.94	3.67	3.04	4.28	3.45	2.69	2.63	3.56	3.02	2.08	3.29	3.14	2.89
2000	0.69	2.75	2.06	0.74	2.60	2.96	2.35	2.52	2.60	0.75	2.34	1.89	2.87	3.39	3.01	4.69	3.48	2.11	2.93	3.73	3.20	1.91	3.68	2.89	2.20
2010	0.73	2.69	1.37	0.74	2.23	3.17	2.03	2.85	2.54	0.79	2.21	1.88	2.73	3.11	2.90	4.33	3.42	1.60	3.17	3.64	3.34	1.60	3.85	2.67	2.32

Source*:* (van Leeuwen and van Leeuwen-Li, 2014[23]), "Education since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en.

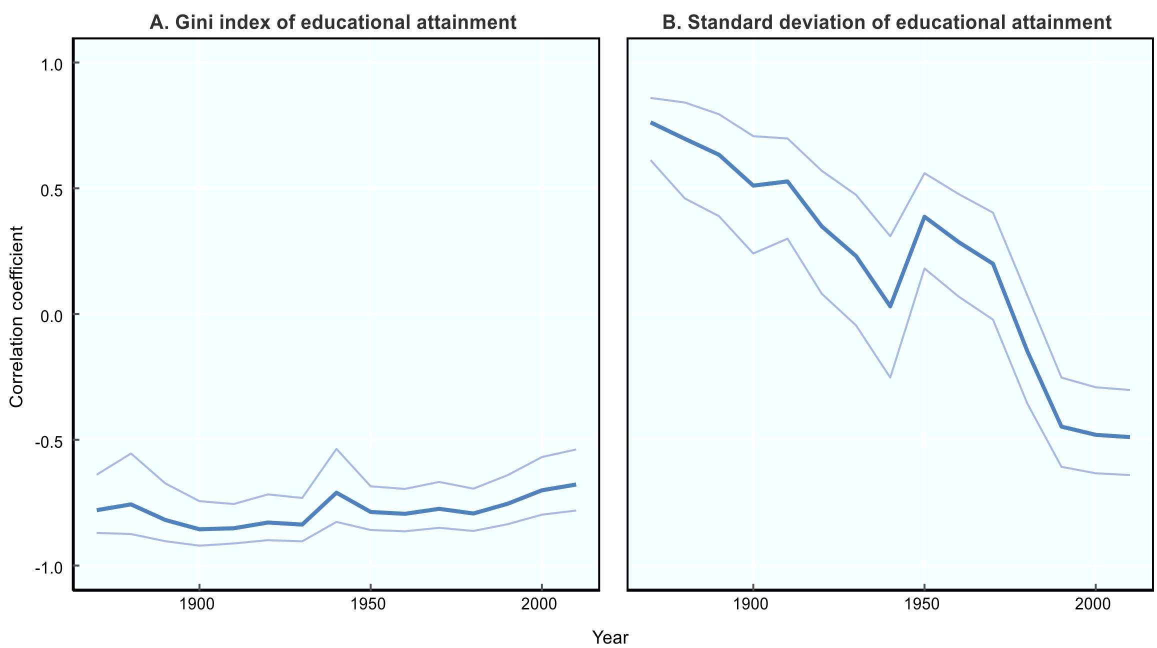
StatLink https://stat.link/4sra6q

Correlation of educational inequality with income

Since, as described in this chapter, measures of educational inequality may, at times, differ substantially from one another, it is unsurprising that the same differences exist in the correlation of these measures with per capita GDP. Looking at the Gini index, based on the models from Lopez, Thomas and Wang (1999[7]) and Földvári and van Leeuwen (2011[9]), we might expect a negative effect, i.e. the higher the per capita GDP, the lower the educational inequalities. Yet, while some studies found this effect to be significant (Lopez, Thomas and Wang, 1999[7]; Castelló and Doménech, 2002[11]), others, correcting for missing variables such as physical capital, found it to be small or insignificant (Földvári and van Leeuwen, 2011[9]; van Damme, 2014[38]). Since our correlations do not correct for missing factors, Figure 7.3 shows a negative and significant correlation between GDP per capita and the Gini index for educational inequalities (Panel A), a correlation that is broadly unchanged over time.

Figure 7.3. Correlation between educational inequality and GDP per capita

Pearson correlation coefficient



Note: Thick and thin lines depict the correlation coefficient and 95% confidence intervals respectively.
Source: (van Leeuwen and van Leeuwen-Li, 2014[23]), "Education since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en.

StatLink https://stat.link/j41mri

The result is different for the standard deviation (Panel B), which runs from positive (i.e. higher GDP per capita is associated with higher educational inequality) before 1950, to negative (i.e. higher GDP per capita is associated with lower educational inequality) in the second half of the 20th century (Figure 7.3).This may be attributed partly to a scale effect: European countries and their Offshoots had both higher levels of GDP per capita and higher educational inequality before around 1950, while the less developed countries had low levels of both per capita GDP and educational inequality, resulting in positive correlations. Yet, as shown in Figure 7.2, over time educational inequality in the less developed countries increased, whereas it declined for the developed countries, thus leading to declining, and eventually negative, correlations.

Priorities for future research

Inequality in educational attainment can be expressed in various ways, depending on one's perspective. If relative differences are considered more important, measures such as the Gini index are preferred. However, if absolute differences in years of education are seen as relevant, absolute measures of inequality, such as the standard deviation, are preferred.

Overall, the Gini index of educational inequality at the world level declined significantly over the 19th century and even more so in the 20th and early 21st centuries. This reduction has been caused largely by the reduction of the share of people without any formal education. On the contrary, the standard deviation of educational inequality first increases over time when a growing share of the population participates in formal education, and then declines when this share exceeds 50% of the population. These patterns of rising and declining educational inequality, as measured by the standard deviation, are captured by the notion of the educational Kuznets curve. With the exception of some less developed countries that are still on the rising part of this curve, most countries reached the peak of the standard deviation at some point in the 20th century. These exceptions are noteworthy since, when countries are on the upward trajectory of educational inequality, they face a trade-off between increasing the mean years of schooling and reducing educational inequality.

These patterns in educational inequality are, however, tentative, and much more research is necessary. For example, in this chapter we focused on inequality in years of education, ignoring the large literature on the quality of education. After all, an increase of one year in tertiary education will have a different impact on educational inequality than the same increase in primary education. Likewise, there may be large differences in the quality of primary education between countries. At present, various studies use the OECD's Programme for International Student Assessment (PISA) to capture the ability of 15-year-olds in e.g. reading and mathematics. Unfortunately, PISA scores are unavailable for historical periods. One way to overcome this would be to use proxies for educational quality, such as wages or pupil-teacher ratios, which may be available in historical periods. Finally, future research should also focus on personal characteristics, such as gender or place of residence, which may influence trends in educational inequality, e.g. Lee and Lee (2016[6]) and Ali, Benjamin and Mauthner (2004[39]).

References

Ali, S., S. Benjamin and M. Mauthner (2004), *The Politics of Gender and Education: Critical Perspectives*, Palgrave Macmillan, London. [39]

Amiel, Y. and F. Cowell (1999), *Thinking about Inequality: Personal Judgment and Income Distributions*, Cambridge University Press, Cambridge. [10]

Bandyopadhyay, S. (2018), "The Absolute Gini is a more reliable measure of inequality for time dependent analyses (compared with the Relative Gini)", *Economics Letters*, Vol. 162, pp. 135-139, https://doi.org/10.1016/j.econlet.2017.07.012. [14]

Barro, R. and J. Wha Lee (2015), *Education Matters: Global Schooling Gains from the 19th to the 21st Century*, Oxford University Press, Oxford. [21]

Barro, R. and J. Wha Lee (2013), "A new data set of educational attainment in the world, 1950-2010", *The Journal of Development Economics*, Vol. 104, pp. 184-98, https://doi.org/10.1016/j.jdeveco.2012.10.001. [18]

Boli, J. (1989), *New Citizens for a New Society. The Institutional Origins of Mass Schooling in Sweden*, Pergamon Press, Oxford. [30]

Bolt, J. and D. Bezemer (2009), “Understanding long-run African growth: colonial institutions or colonial education?”, *The Journal of Development Studies*, Vol. 45/1, pp. 24-54, https://doi.org/10.1080/00220380802468603. [31]

Bukowski, P. (2019), “How history matters for student performance. Lessons from the partitions of Poland”, *The Journal of Comparative Economics*, Vol. 47/1, pp. 136-75, https://doi.org/10.1016/j.jce.2018.10.007. [41]

Cao, N. (2013), *Population Sequel Collection In The Period of the Republic of China*, National library Press, Beijing. [27]

Card, D. (2001), “Estimating the return to schooling: Progress on some persistent econometric problems”, *Econometrica*, Vol. 69/5, pp. 1127-60, https://www.jstor.org/stable/2692217?seq=1#metadata_info_tab_contents. [42]

Card, D. and J. DiNardo (2002), “Skill-biased technological change and rising wage inequality: Some problems and puzzles”, *The Journal of Labor Economics*, Vol. 20/4, pp. 733-783, https://doi.org/10.1086/342055. [1]

Castelló, A. and R. Doménech (2002), “Human capital inequality and economic growth: Some new evidence”, *Economic Journal*, Vol. 112/478, pp. 187-200, https://doi.org/10.1111/1468-0297.00024. [11]

Cogneau, D. and A. Moradi (2014), “Borders that divide: Education and religion in Ghana and Togo since colonial times”, *The Journal of Economic History*, Vol. 74/3, pp. 694-724, https://doi.org/10.1017/S0022050714000576. [33]

Cohen, D. and M. Soto (2007), “Growth and human capital: Good data, good results”, *The Journal of Economic Growth*, Vol. 12/1, pp. 51-76, https://doi.org/10.1007/s10887-007-9011-5. [17]

de la Fuente, A. and R. Doménech (2015), “Educational attainment in the OECD, 1960-2010. Updated series and a comparison with other sources”, *Economics of Education Review*, Vol. 48, pp. 56-74, https://doi.org/10.1016/j.econedurev.2015.05.004. [19]

Dupraz, Y. (2019), “French and British colonial legacies in education: Evidence from the partition of Cameroon”, *The Journal of Economic History*, Vol. 79/3, pp. 628-68, https://doi.org/10.1017/S0022050719000299. [36]

Eyre, L., T. Lovell and C. Smith (2004), “Gender equity policy and education: Reporting on/from Canada”, in *The Politics of Gender and Education*, London: Palgrave Macmillan. [43]

Feldmann, H. (2016), “The long shadows of Spanish and French colonial education”, *Kyklos*, Vol. 69/1, pp. 32-64, https://doi.org/10.1111/kykl.12102. [44]

Field, S., M. Kuczera and B. Pont (2007), *“No More Failures: Ten Steps to Equity in Education”*, https://www.oecd.org/education/school/45179151.pdf. [2]

Földvári, P. and B. van Leeuwen (2014), “Educational and income inequality in Europe, ca. 1870-2000”, *Cliometrica*, Vol. 8/3, pp. 271-300, https://doi.org/10.1007/s11698-013-0105-3. [25]

Földvári, P. and B. van Leeuwen (2011), "Should less inequality in education lead to a more equal income distribution?", *Education Economics*, Vol. 19/5, pp. 537-54, https://doi.org/10.1080/09645292.2010.488472. [9]

Fourie, J. and C. Swanepoel (2015), "When selection trumps persistence. The lasting effect of missionary education in South Africa", *TSEG/ Low Countries Journal of Social and Economic History*, Vol. 12/1, pp. 1-30. [40]

Fuller, B., J. Edwards and K. Gorman (1986), "When does education boost economic growth? School expansion and school quality in Mexico", *Sociology of Education*, Vol. 59/3, pp. 167-81, https://doi.org/10.2307/2112341. [29]

Garnier, M. and M. Schafer (2006), "Educational model and expansion of enrollments in Sub-Saharan Africa", *Sociology of Education*, Vol. 79/22, pp. 153-75. [32]

ISSP Research Group (2017), *"International Social Survey Programme: Social Inequality IV - ISSP 2009."*, Cologne: GESIS Data Archive. [28]

Jenkins, S. and M. Jäntti (2005), "Methods for summarizing and comparing wealth distributions", *Social Research. ISER Working Paper Series*. [12]

Lee Wha, J. and H. Lee (2016), "Human Capital in the Long Run", *The Journal of Development Economics*, Vol. 122, pp. 147-169, https://doi.org/10.1016/j.jdeveco.2016.05.006. [6]

Lopez, R., V. Thomas and Y. Wang (1999), *Addressing the Education Puzzle: The Distribution of Education and Economic Reform*, The World Bank. [7]

Meschi, E. and F. Scervini (2014), "Expansion of schooling and educational inequality in Europe: The educational Kuznets Curve revisited", *Oxford Economic Papers*, Vol. 66, pp. 660-80, https://doi.org/10.1093/oep/gpt036. [16]

Meyer, H. (2011), ""Path dependence in German and American public education: The persistence of institutional difference in a globalizing world"", in Mitchell, D., R. Crowson and D. Shipps (eds.), *Shaping Education Policy: Power and Process*, Routledge, New York, https://doi.org/10.4324/9780203837351. [3]

Morrisson, C. and F. Murtin (2013), "The Kuznets Curve of human capital inequality: 1870-2010", *The Journal of Economic Inequality*, Vol. 11/3, pp. 283-301, https://doi.org/10.1007/s10888-012-9227-2. [22]

Morrisson, C. and F. Murtin (2009), "The century of education", *Journal of Human Capital*, Vol. 3/1, pp. 1-42, https://doi.org/10.1086/600102. [5]

Niño-Zarazúa, M., L. Roope and F. Tarp (2017), "Global inequality: Relatively lower, absolutely higher", *Review of Income and Wealth*, Vol. 63/4, pp. 661-84, https://doi.org/10.1111/roiw.12240. [13]

Psacharopoulos, G. (1979), "On the weak versus the strong version of the screening hypothesis", *Economics Letters*, Vol. 4, pp. 181-85. [48]

Ram, R. (1990), "Educational expansion and schooling inequality: International evidence and some implications", *The Review of Economics and Statistics*, Vol. 72/2, pp. 266-74, https://doi.org/10.2307/2109716. [15]

Samir, K. and W. Lutz (2017), “The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100”, *Global Environmental Change*, Vol. 42, pp. 181-92, https://doi.org/10.1016/j.gloenvcha.2014. [20]

Samoff, J. (1991), “Socialist education?”, *Comparative Education Review*, Vol. 35/1, pp. 1-22, https://www.jstor.org/stable/1188013. [37]

Stites, R. and L. Semali (1991), “Adult literacy for social equality or economic growth? Changing agendas for mass literacy in China and Tanzania”, *Comparative Education Review*, Vol. 35/1, pp. 44-75, https://www.jstor.org/stable/1188015. [35]

Subramanian, S. and D. Jayaraj (2013), “The evolution of consumption and wealth inequality in India: A quantitative assessment”, *Journal of Globalization and Development*, Vol. 4/2, pp. 253-81, https://doi.org/10.1515/jgd-2013-0018. [45]

Tamura, R. et al. (2019), “Economic Growth in the Long Run”, *The Journal of Development Economics*, Vol. 137/March, pp. 1-35, https://doi.org/10.1016/j.jdeveco.2018.10.010. [24]

Thomas, V., Y. Wang and X. Fan (2003), “A new dataset on equality in education: Gini and Theil Indices of schooling for 140 countries, 1960-2000”, *The Journal of Educational Planning and Administration*, Vol. 17/1, pp. 5-33. [8]

Trent, W. (1984), “Equity considerations in higher education: Race and sex differences in degree attainment and major field from 1976 through 1981”, *The American Journal of Education*, Vol. 92/3, pp. 280-305, https://www.journals.uchicago.edu/doi/abs/10.1086/443750. [46]

UNESCO Institute for Statistics (2012), *International standard classification of education: ISCED 2011*, Montreal: UNESCO Institute for Statistics. [4]

van Bavel, J. (2012), “The reversal of gender inequality in education, union formation and fertility in Europe”, *Vienna Yearbook of Population Research*, pp. 127-54. [47]

van Damme, D. (2014), “How Closely is the Distribution of Skills Related to Countries’ Overall Level of Social Inequality and Economic Prosperity?”, *OECD Education Working Papers*, No. 105, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jxvd5rk3tnx-en. [38]

van Leeuwen, B. and J. van Leeuwen-Li (2014), “Education since 1820”, in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en. [23]

White, B. (1996), “Talk about school: Education and the colonial project in French and British Africa (1860-1960)”, *Comparative Education*, Vol. 32/1, pp. 9-26, https://doi.org/10.1080/03050069628902. [34]

Yugan (1947), *Fifth Township, Fourth District, Household Registration Books*, Shangrao City: Yugan county, Shangrao City, Jiangxi Province. [26]

Notes

[1] http://www.oecd.org/education/launch-of-oecd-2015-education-at-a-glance.htm.

[2] Some authors have claimed that private returns cannot be calculated, as education does not signal higher productivity but rather functions as a screening tool for innate abilities. Yet, as pointed out by Psacharopoulos (1979[48]) and Card (2001[42]), even though screening might occur at the hiring phase (when employers have little knowledge about one's productivity), this effect ebbs away over time when employers become more capable of assessing the true productivity of individual workers.

[3] Equity depends on a great many factors. The role of race (Trent, 1984[46]), gender (Eyre, Lovell and Smith, 2004[43]; van Bavel, 2012[47]) and income has been widely discussed in relation to educational differences.

[4] Often these historical differences are proxied by border effects (Bukowski, 2019[41]), the experience of colonialism (Feldmann, 2016[44]) or religion (Fourie and Swanepoel, 2015[40]).

[5] There are some studies that included, besides the Gini, also the Theil index. See e.g. Thomas, Wang and Fan (2003[8]) or Morrisson and Murtin (2013[22]).

[6] There also exist intermediate measures between relative and absolute inequality, but we will ignore those here (Subramanian and Jayaraj, 2013[45]).

8 Gender equality since 1900

Selin Dilli, Sarah G. Carmichael and Auke Rijpma, Utrecht University

This chapter provides an overview of global trends in gender equality in health, socio-economic resources and politics over the entire 20th century. It does so by extending the historical gender equality index (HGEI) introduced in the previous *How Was Life?* report back to 1900 and forward to 2010, and by including additional indicators. While progress since 1900 towards gender equality is visible especially in the dimensions of health and socio-economic resources, cluster analysis reveals that the groupings of countries by level of gender equality remains similar through time. The main exceptions are Southern Europe and the Nordic countries, which witness substantial improvements in the post-1950 period.

Introduction

Gender inequalities manifest themselves at all levels of society and are significant at a global scale, especially when women's decision-making capacity in politics, business and research are considered (OECD, 2016[1]). However, our knowledge of what instruments work best to eliminate gender inequalities remains limited by the fact that gender statistics provide information mostly on cross-sectional patterns. In the absence of systematic time-series data, we know little about the historical development of gender equality, let alone about the relative importance of different theories in explaining changes over time and between world regions (Carmichael, Dili and Rijpma, 2014[2]; Dilli, Carmichael and Rijpma, 2019[3]). Without this knowledge, it is much more difficult to develop the right tools to help eliminate gender inequalities.

Gender inequality implies economic and social costs. A recent estimate by Ferrant and Kolev (2016[4]) puts the economic loss at 7.5% of world GDP. Moreover, progress towards addressing this issue can be slow. Using the Global Gender Gap (GGG) index, which captures gender differences in health, economics, politics and household decision-making since 2011, the World Economic Forum (WEF, 2018[5]) estimates that at the current pace of progress it will take another 108 years to close the gender gap. For wages, it might take as long as 200 years (WEF, 2018[5]).

Such estimates are based on the assumption that the progress towards gender equality is linear and independent of the historical experiences of countries. But historical evidence shows that progress towards gender equality has not been linear (Carmichael, Dili and Rijpma, 2014[2]). The U-shaped hypothesis formulated by Goldin (1995[6]) implies that, in the case of the United States, industrialisation at the beginning of the 20th century coincided with a decline in the share of women participating in the labour force, which hit its lowest point around the 1920s. However, in the absence of historically comparable data across countries, it is impossible to conclude whether this profile is unique to the United States or holds globally. Progress towards gender equality has also experienced sudden setbacks. For example, with the collapse of the Soviet Union in 1989, the share of women in national parliament declined substantially as the previous quota system was phased-out (Carmichael, Dili and Rijpma, 2014[2]). These reversals in countries' progress towards gender equality cast doubt on even the WEF estimate of 108 years needed to achieve full equality between women and men.

In recent years policy analysts and researchers have come to a greater appreciation of the importance of a longer-term perspective for understanding the roots of gender inequalities. This is reflected in the development of datasets that allow for long-term comparisons across countries for a broad range of well-being outcomes. For instance, the World Bank (2015) has presented a dataset on women's legal rights that is available at a global level and covers the period since 1960. However, most datasets on gender equality lack a long-term historical perspective (Carmichael, Dili and Rijpma, 2014[2]; Dilli, Carmichael and Rijpma, 2019[3]).

In a chapter of the previous *How Was Life?* report, Carmichael, Dilli and Rijpma (2014[2]) presented data on women's well-being in socio-economic status, household position, politics and health from 1950 onwards for 20 major countries, as well as regional averages based on data for 112 countries. As gender equality has many facets (Sen, 1990[7]), that chapter relied on the Historical Gender Equality Index (HGEI) to capture its multidimensional nature. While this composite index highlighted that there has been significant global progress towards gender equality over the 20th century, it also confirmed the WEF view that there is still a long way to go. Among the countries covered in Carmichael, Dilli and Rijpma (2014[2]), none had achieved perfect equality by the 2000s, and large gender gaps persisted in many dimensions. In another study, we also showed that, despite progress towards gender equality in health, household decision-making, politics and socio-economic resources, there is limited evidence of global convergence towards a gender equal world (Dilli, Carmichael and Rijpma, 2019[3]). Despite the importance of developing a historical understanding of gender inequalities, and despite recent efforts in this direction, comparable historical gender statistics remain scarce on a global scale, especially for the period before the mid-20th century.

The lack of data limits the possibilities to understand the historical position of women. In the previous *How Was Life?* report, our HGEI covered only the period after 1950 due to lack of data on educational attainment and labour force participation by gender before that year. Given that for many of the industrialised countries, the gap in gender equality started to close before this period – as in the case of labour force participation in the United States (Goldin, 1995[6]) – coverage prior to 1950 is crucial to identify the conditions under which gender gaps might narrow. Moreover, the insight that the previous *How Was Life?* report provided into gender differences in economic empowerment remained limited, as we lacked data on key indicators such as women's wages. This chapter addresses these two shortcomings in the following ways. First, we extend the HGEI index back to 1900, and forward to 2010. This is as far back as current data allow, given our method, implying that for the first time we can provide an encompassing overview of gender equality over the full 20th century. Second, we present new data on the gender gap in wages in the 20th century. Third, we present new data on the share of female-headed households and how this has changed over time. In addition, we group countries according to their similarities and differences in the inequalities that women face in different dimensions of daily life, and identify countries where women are most disadvantaged. We do so by applying a cluster analysis on the composite HGEI index in two sub-periods (i.e. pre-1950 and post-1950) to assess whether the clustering of countries in terms of gender inequalities persists or changes over time as a result of improvements in any of the dimensions included in the HGEI index.

Description of the concepts used

This section describes how we measure the multifaceted nature of gender equality. Because gender inequality manifests itself in different dimensions of daily life, measuring it is often done using several indicators, as no single measure can capture all the dimensions in which gender inequality occurs. This chapter takes the extra step of introducing an overall composite index, based on indicators of how women fare relative to men in terms of health status, socio-economic resources, presence in national parliaments and age of first marriage. Combining these data into a composite index summarises the different indicators in a single standardised measure, giving insight into overall gender equality at national level (Dilli, Carmichael and Rijpma, 2019[3]). While the availability of historical data plays a role in the choice of our indicators, we rely on previous composite indices of gender equality, in particular the World Economic Forum's Global Gender Gap (GGG) and on the dimensions they identify as crucial in measuring women's position. Dilli, Carmichael and Rijpma (2019[3]) discuss the comparability of the composite index used in this chapter (which we label HGEI) with the existing indices such as the OECD's Social Institutions and Gender Index (SIGI) and the GGG, as well as the theoretical justification of our indicators and the method followed in constructing the HGEI in more detail. Here we provide a brief summary to guide the reader.

In constructing the HGEI, our goal is to provide a global overview of gender equality. Our choice of dimensions included in our composite index was constrained by both international comparability and data availability. Conceptually, our goal is to provide an overview of the differences between men and women in their agency to achieve desired life outcomes (Sen, 1999[8]; Nussbaum, 2000[9]). We include health, as it is key to being able to enjoy other aspects of well-being (Sen, 1999[8]). To capture this dimension, we consider differences between women and men in terms of life expectancy at birth and ratios of infant girls to boys. Second, we look at gender differences in ages at first marriage as an indicator of women's agency in the household; this is an important dimension, as the SIGI 2019 report has shown that discrimination in the family sphere is pervasive,, particularly in sub-Saharan Africa, MENA and Southeast Asia (OECD, 2019[10]). Third, we look at the share of women elected in national parliaments as a reflection of whether women are involved in the political decision-making process. Fourth, as a measure of gender inequalities in socio-economic resources, we use the differences in educational attainment and labour force participation between women and men. While education contributes to women's capacity to make meaningful life decisions, employment matters, because the income it generates gives women the ability to live their lives independently from men and strengthens their bargaining position in the household. Since the HGEI is a composite indicator, we apply slightly different rules to those used in other chapters in this book when reporting regional averages. The 40% threshold of regional population covered by non-imputed data (the threshold used in Carmichael, Dilli and Rijpma (2014[2]) is now applied to the average of the six indicators, which means that if one indicator has lower coverage, it can be compensated by the other indicators.

While the HGEI summarises the multidimensional nature of gender equality, we also discuss in this chapter two new measures of gender equality, namely women's wages and the share of female-headed households. These measures are not included in the composite HGEI index for two main reasons. First, and most importantly, the time and country coverage of these indicators are more limited than for the indicators included in the HGEI (pre-1950 and global). Second, while the indicators included in the HGEI index provide insight into women's position relative to men, this is not the case for the share of female-headed households. This indicator captures whether women manage autonomously their own households, making this an institutional arrangement that in some cases could benefit women (Van Driel, 1994[11]; Chant, 1985[12]; World Bank, 1999[13]). In some countries, female headship appears to be a voluntary choice, and one that, in the case of Vietnam, is not associated with detrimental effects (World Bank, 1999[13]). However, female-headed households often face greater hardship and are substantially poorer than those headed by men. In an historical context, Horrell and Humphries (1997[14]) show that in England during the Industrial Revolution female-headed households were those where deprivation was greatest. In recent years, female-headed households are over-represented among the poor, and are often identified as the "poorest of the poor" (Bradshaw, Chant and Linneker, 2017[15]). As, depending on national contexts and periods, this indictor can have different implications for women's empowerment, it is excluded from the composite measure presented below.

Historical sources

This section describes the sources used to capture the various dimensions of gender inequality. Table 8.1 presents an overview of the variables used to construct the composite index, which is now extended to 1900 and 2010, as well as the additional indicators of gender equality presented in this chapter, alongside their source and summary statistics. Table 8.1 also presents summary statistics for the underlying indicators of the HGEI.

The HGEI measures the position of women relative to men in terms of health status, household conditions, political representation and socio-economic resources. To measure gender differences in health status, we use the ratio of female-to-male life expectancy at birth and the ratio of girls to boys in the population between the ages of 0 and 5. We use data on the age at first marriage of women and men to capture the relative position of women in the household; those on the share of women (relative to men) among members of national parliaments to measure gender bias in political representations; and the female-to-male ratio in average years of schooling and labour force participation to measure socio-economic standing. For each indicator, we express gender differences in terms of ratios between women and men, implying that a score below 1 indicates bias against women, whereas a score above 1 indicates that women are outperforming men. The ratios allow for a straightforward interpretation of countries' progress over time in each dimension and in the overall composite index (see Dilli, Carmichael and Rijpma (2019[3]) for a more detailed explanation of the use of ratios). Because of our choice to use ratios, as well as the similarity of the indicators included in the GGG, we follow the method of the GGG index to construct our composite index (WEF, 2013[16]). First, for each variable, we truncate the ratios at the equality benchmark. Second, we take the weighted average of two sub-indices, i.e. health status and socio-economic resources, as these two sub-indices include two indicators, whereas gender equality in the household and in politics are captured by a single indicator, namely the gap in ages at first marriage and the share of women among members of national parliaments. Third, we normalise the underlying variables of each dimension by using their standard deviations. Last, we take the arithmetic average of the four sub-indices multiplied by 100. We provide a more detailed discussion of our index in Dilli, Carmichael and Rijpma (2019[3]). The data in Table 8.1 refer to non-imputed and non-truncated data. Due to increased country coverage, some of the values shown in Table 8.1 differ from those presented in Carmichael, Dilli and Rijpma (2014[2]).

Table 8.1. Descriptive statistics for measures of gender inequality

All indicators are measured in ratios of female to male

Dimension	Indicator (all expressed as ratios between women and men)	Included in HGEI	Range	Mean (standard deviation)	Number of countries	Years	Source
Health	Life expectancy	YES	0.80-1.48	0.99 (0.06)	139	1870-2010	United Nations (2013[17]); lifetable.de, Human Mortality Database; Preston (1975[18])
	Sex ratio at young age (0 to 5)	YES	0.83-1.21	0.97 (0.03)	142	1870-2010	Mitchell (2007[19]); United Nations (2013[17])
Socio-economic standing	Average years schooling	YES	0.01-1.50	0.71 (0.30)	89	1870-2010	Barro and Lee (2013[20]); Barro and Lee (2015[21]); Lee and Lee (2016[22])
	Labour force participation	YES	0.05-1.11	0.58 (0.24)	140	1870-2010	Mitchell (2007); ILO (2010[23])
	Monthly wages	NO	0.57-1.03	0.88 (0.12)	34	1955-1985	De Zwart et al. (2014); ILO (2018)
Household decision-making	Age at first marriage	YES	0.64-0.96	0.87 (0.06)	70	1870-2010	Carmichael (2011[24])
	Female-headed households	NO	0.03-0.54	0.24 (0.097)	91	1970-2003	Minnesota Population Center (2018)
Political representation	Parliament seats	YES	0.00-1.29	0.09 (0.14)	137	1900-2010	Paxton, Green and Hughes (2008[25]); online electoral archives; Inter-parliamentary Union (2011[26])
HGEI			40.35-92.96	61.59 (7.50)	138	1900-2010	Carmichael, Dilli and Rijpma (2014[2]); Dilli, Carmichael and Rijpma (2019[3])

StatLink https://stat.link/f7krq4

Until recently, while measures on gender inequalities in health status, household conditions and political representation were available from 1900 onwards, data on socio-economic inequalities were available only from 1950 onwards. Recently, however, there have been improvements in the availability of educational attainment statistics disaggregated by gender: both Barro and Lee (2015[21]) and Lee and Lee (2016[22]) updated their datasets, which we use to compute our HGEI index. This allows tracking ratios of female-to-male educational attainment back to 1820. Progress has also been made in the digitalisation of statistics on labour force participation by gender. Using Mitchell (2007[19]), we have extended data on women's and men's labour force participation for a selected number of countries back to the 1900s. This allows us to re-compute the HGEI to explore trends in gender equality for the entire 20th century.

This chapter also presents new historical evidence on the economic position of women. The International Labor Organization (ILO) provides mean nominal monthly earnings of employees by sex and by economic activity since the 1970s. Additionally, for a selected number of countries, it is possible to gather information on the real wages of women and men since the post-war period. For instance, de Zwart, B. van Leeuwen and J. van Leeuwen-Li (2014[27]) present wage data from ILO statistics for a number of developing countries from the mid-1950s onwards.

The indicator we use captures the share of all households that are female-headed using historical census data calculated for four points in time (1970, 1980, 1990 and 2000) from data available in the IPUMS database (Minnesota Population Center, 2018[28]). While the data from IPUMS are mostly available from the 1960s, for some Western countries these data extend back to the early 19th century. Together with efforts made by the Mosaic Project (n.d.[29])to digitalise historical censuses, these data will offer the opportunity to expand coverage to a number of Eastern and Southern European countries in the future.

We apply a cluster analysis on the HGEI index to identify countries where women are more disadvantaged in a global comparison, and whether countries remain or move out of this cluster over time. Clustering orders and organises data in such a way that each cluster consists of countries that are similar in terms of the underlying variable and dissimilar to other clusters (Zeumo, Tsoukiàs and Some, 2012[30]). This method has been used in multidimensional poverty measurement based on the capability approach (Ferro, Fluckiger and Weber, 2008[31]). It provides some insight into gender inequality in the world, without applying a pre-determined threshold level of disadvantage that women face across different societies.

Data quality

While recent years have seen important improvements in the collection and availability of historical gender statistics, the data used in this chapter have their limitations. First, while we extend the HGEI back to the early 20th century, we can do this only for countries for which Mitchell (2007[19]) provides statistics on women's labour force participation.

Second, while we extend the measures that provide insight into the position of women, we still lack data on many other aspects that matter for gender equality, such as the unequal allocation of time use in the home and violence against women.

Regarding women's wages, only data for a handful of occupations (bookbinders, spinners, sewing machine operators) are available in a sufficiently large number of countries for the period 1955-1985. This means that we can report only within-occupation wage ratios or ratios of the unweighted average wage across this limited number of occupations. Our analysis therefore considers simply the gender wage gap in specific occupations. This is important as occupations are often gender-segregated, with women clustering into lower paid occupations or occupations dropping in relative remuneration as the share of women rises.

The historical ILO wage data have a number of other limitations. For a large number of occupational categories, wage ratios are equal to 1, which suggests that some countries reported minimum (rather than actual) wages, or simply filled in the ILO questionnaire with identical wages for men and women. The fact that the ILO wage data for many developing countries (e.g. in sub-Saharan Africa) reported the same wages for men and women suggests this might especially be an issue with the data from that region. Therefore, rather than regional averages, below we present the trends in wages only for a selected number of countries. Moreover, the fact that we have data only for a number of occupations such as textiles, in which women often make up a majority of the labour force, means we should not discount the possibility that, in those limited cases, wages were in fact equal.

The data used for the HGEI that were included in the previous report are also subject to some caveats (Carmichael, Dili and Rijpma, 2014, p. 224[2]). For one, our measure does not capture how women are faring in absolute terms, but only relative to men. We are also not able to provide a full overview of gender disparities in some of the dimensions captured by other composite indices such as access to financial services. Finally, women often bear the burden of domestic duties and unrecorded care tasks, which our measures do not reflect (see Dilli, Carmichael and Rijpma (2019[3]) for a more elaborated discussion).

A further concern is the limited comparability of these measures across countries. Our indicators (except the educational attainment indicator) rely on official statistics, implying that our measures are collected within a given legislative framework. For instance, we can make comparisons based only on legally documented marriages, whose definition differs in different national contexts. Similar issues occur when measuring women's political participation as well as women's engagement in economic activity. For instance, our measure of political representation of women in national parliaments does not provide any insight into the grassroots political activity of women.

When available, data quality is typically quite high (see Table 8.2 and Annex Table 8.A.1). The overwhelming majority of the data used in this chapter comes from official statistical agencies or is the product of historical research using similar sources and methods.

Table 8.2. Quality of gender data on education and labour force participation

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Sub-Saharan Africa	Middle East and North Africa	East Asia	South and Southeast Asia
Average years of schooling								
1820	2	2	2	2	2	2	2	2
1870	2	2	2	2	2	2	2	2
1913	2	2	2	2	2	2	2	2
1950	2	2	2	2	2	2	2	2
1973	2	2	2	2	2	2	2	2
2008	2	2	2	2	2	2	2	2
Labour Force Participation								
1820	2	2	2	2	2	2	2	2
1870	2	2	2	2	2	2	2	2
1913	2	2	2	2	2	2	2	2
1950	1	1	1	1	1	1	1	1
1973	1	1	1	1	1	1	1	1
2008	1	1	1	1	1	1	1	1

Note: 1. High quality: the product of official statistical agency (national or international); 2. Medium quality: the product of economic-historical research using the same sources and methods as applied by official statistical agencies; 3. Moderate quality: economic historical research, but making use of indirect data and estimates; and 4. Low quality: estimates based on a range of proxy information. In case of multiple sources, the lowest quality source is given.

Main highlights of historical developments in gender equality

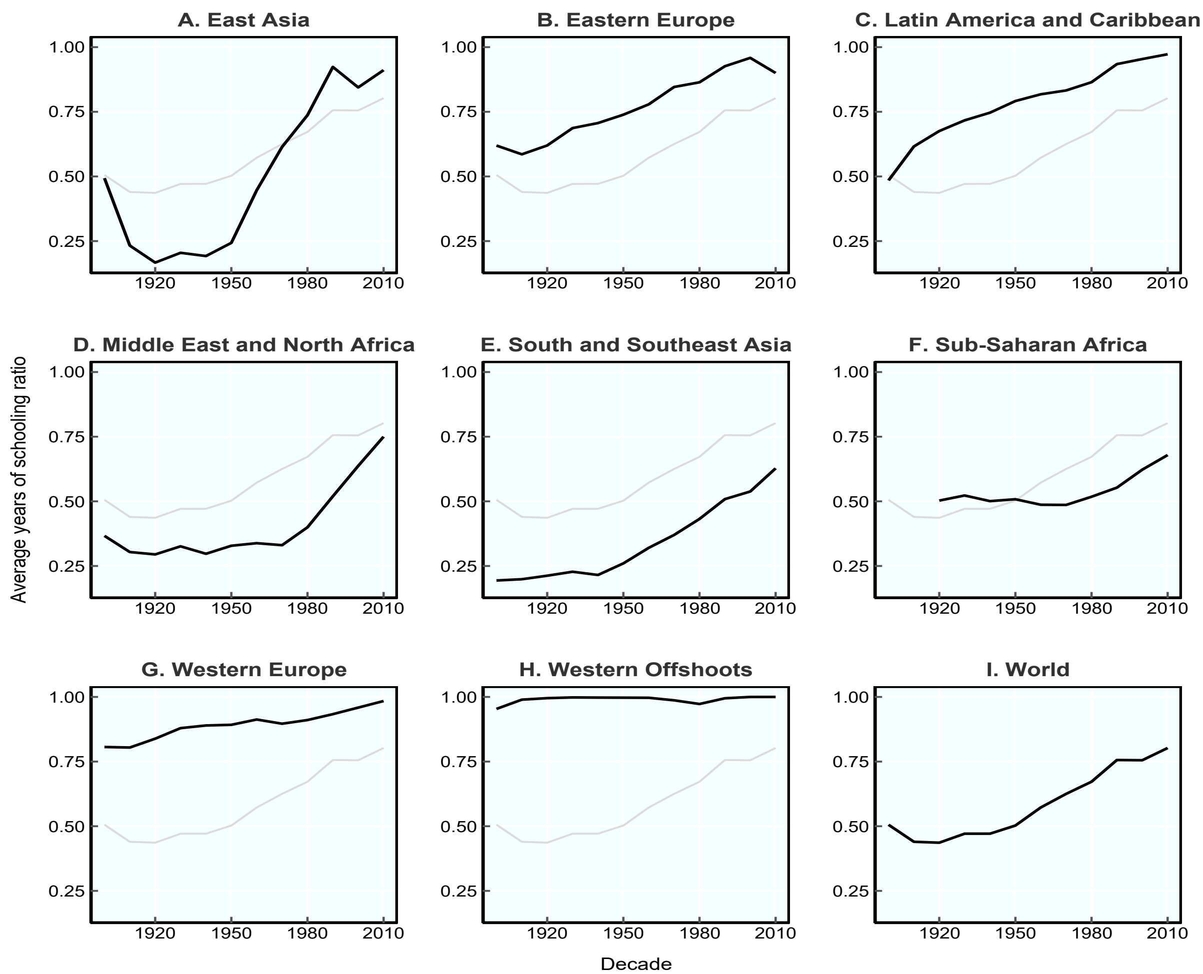
This section presents trends in individual measures of gender equality and then trends in the HGEI index by world region since 1900, as well as the results of our cluster analysis.

Historical developments in different aspects of gender inequality

While Carmichael et al. (Carmichael, Dili and Rijpma, 2014[2]) presented educational attainment data from 1950 onwards, Figure 8.1 presents gender differences in educational attainment since 1900. It is clear that a few regions experienced progress towards gender equality in education prior to the mid-20th century. This trend is especially visible in Western and Eastern Europe and in former Soviet Union countries, which show a tendency towards greater equality in this measure since the 1920s. While Latin American countries displayed greater gender equality in education already in the early 20th century, much of the narrowing of the gender gap took place from the mid-20th century onwards. As noted by Carmichael et al. (2014[2]), there was no gender gap in education in the Western Offshoots already in the 1950s, and Figure 8.2 shows that a similar pattern was already visible around the 1920s. For the regions that lag behind in the process of closing the gender gap in education, trends prior to the mid-20th century show that they experienced the most progress after the 1950s. A clear improvement in women's education is visible in the MENA region, which made the most progress starting from the 1970s, and in East Asia from the 1960s onwards. In the other countries in Asia and sub-Saharan Africa, progress in closing the gap has been more limited.

Figure 8.1. Ratios of female-to-male average years of schooling across world regions

Decennial averages



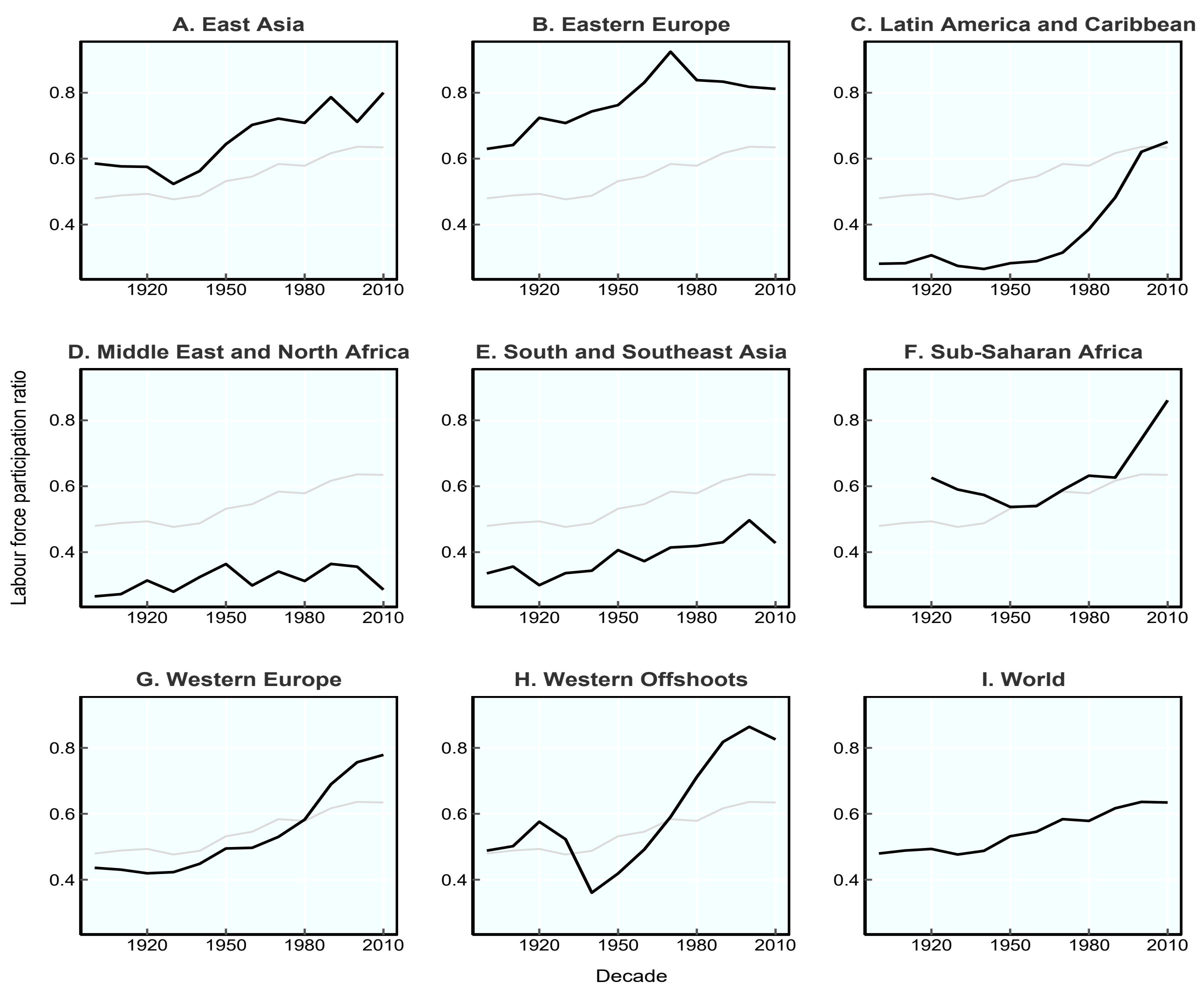
Note: Grey lines show world average.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/d9mte3

Figure 8.2 presents trends in the share of female-to-male labour force participation for the 20th century. While data are scarce, Eastern Europe (including former Soviet Union countries) was at one point by far the most equal region in this dimension of gender equality, with the highest ratio (at 0.9) in 1970. While this region's position has since slipped, it was still one of the best-performing regions in the world in 2010 (together with the Western Offshoots). While Western European countries and the Western Offshoots were close to the world average in the first half of the 20th century, progress towards gender equality in labour force participation accelerated after the 1970s, which Goldin (1995[6]) describes as part of the "Quiet Revolution" associated with the expansion of the service sector. The Middle East and North Africa as well as South and Southeast Asia perform poorly in this aspect of gender equality, being among the worst-performing regions throughout the period and showing only limited progress. Overall, we can observe a small U-shaped curve in this measure in a number of the regions.

Figure 8.2. Ratios of female-to-male labour force participation across world regions

Decennial averages

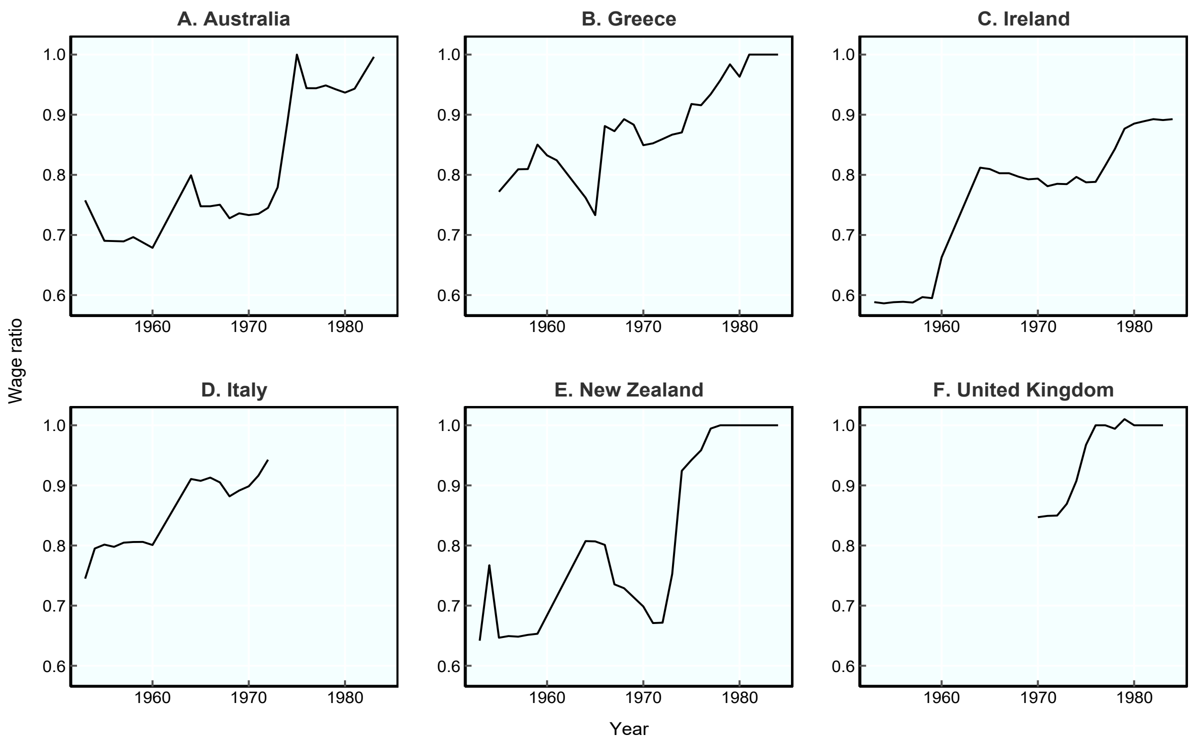


Note: Grey lines show world average.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/xtb3jc

Figure 8.3 shows trends in the gender wage gap for the handful of occupations for which data are available for both men and women from the ILO for a 30-year period. This figure shows that, in those occupations in which women's participation was deemed important enough to record their wages, the gender wage gap declined in the second half of the 20th century. By the 1980s, women were receiving wages similar to those of men in four (Austrialia, Greece, New Zealand and the United Kingdrom) of the six reported countries. These occupations are, however, those that typically earn near minimum wages. At the same time though, the historical record also shows that the status of these occupations declined over time, leading to a drop in both the number of men in these occupations and in their relative remuneration, hence exaggerating the narrowiing of the gender wage gap (Pan, 2015[32]).

Figure 8.3. Ratios of female-to-male average wages across selected countries, 1955-1985



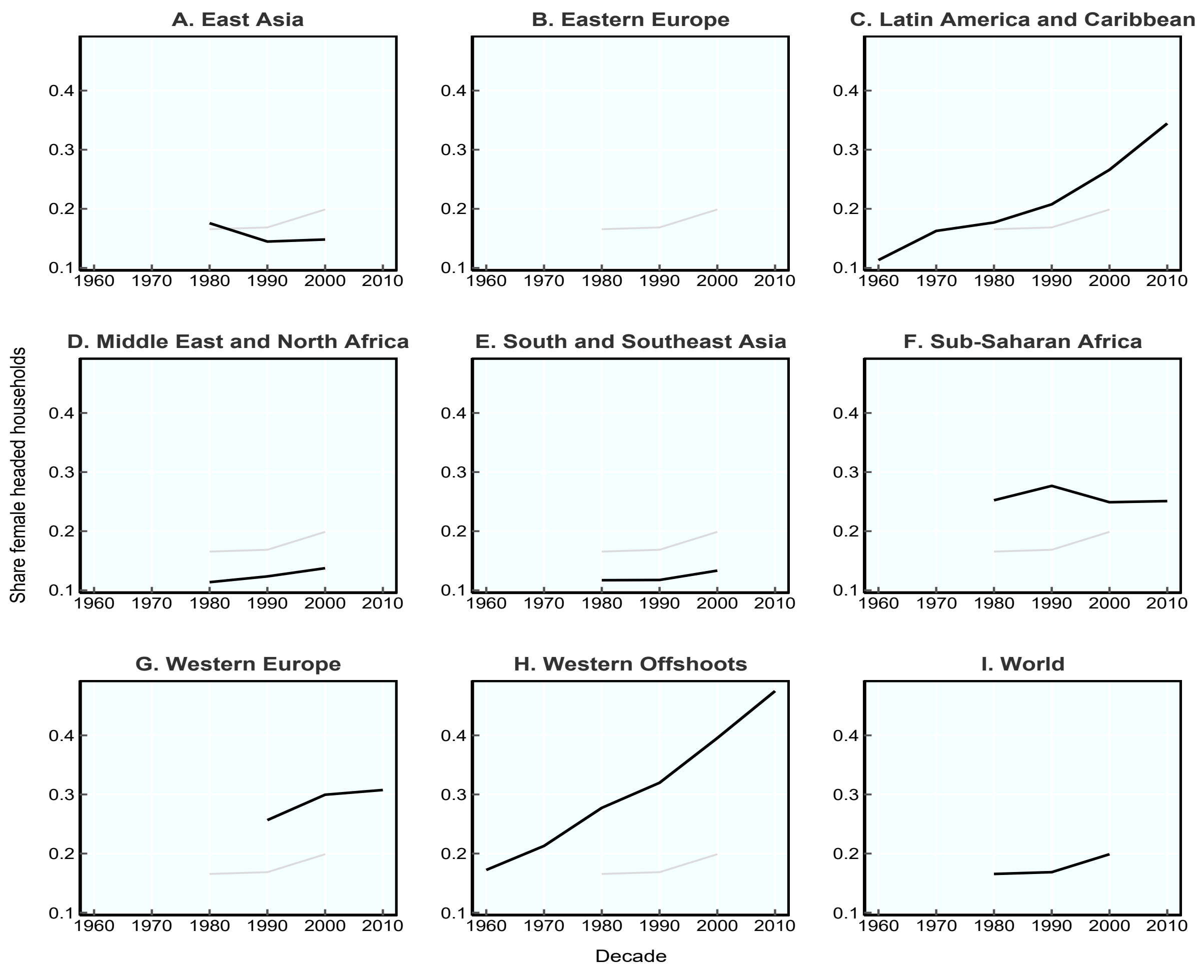
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/t2yv8o

Figure 8.4 presents developments in the share of female-headed households by major world region. This share increased quite strongly in Western Europe, Latin America and North America from the 1970s onwards, with increases also recorded in other regions. The higher life expectancy of women compared to men is an important factor driving these changes, as it results in a higher number of widows following the death of their male partner. Conversely, the share of female-headed households is quite low in the Asian and MENA regions.

Figure 8.4. Female-headed households as a share of all households, 1970-2010

Decennial averages



Note: Grey lines show world average.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/z7q9jh

Historical developments in a composite measure of gender inequality

Before discussing historical trends in our composite indicator of gender equality (HGEI), we briefly compare it to similar indices for the more recent period. Annex Figure 8.A.1 shows the relation of the HGEI with the SIGI, the GGG and the Gender Inequality Index (GII) in 2000. Overall, the HGEI is strongly correlated with the GGG and the GII over recent years, which gives us confidence in the ability of the HGEI to measure gender equality in an historical perspective. The correlation with the SIGI is somewhat weaker (-0.54) but is in line with expectations.[1] While the SIGI focuses on social institutions as determinants of gender equality, the HGEI focuses on outcomes. While we expect drivers and outcomes to be correlated with each other, many dimensions beyond institutions shape gender disparities.

Figure 8.5 shows trends in our composite gender equality index (HGEI), extended back to 1900, across major world regions. As already documented by Carmichael, Dilli and Rijpma (2014[2]), our composite measure of gender equality exhibits a steady upward trend (i.e. lower inequality), especially in the second half of the 20th century, which can be observed for all regions of the world. However, global progress is limited. While in 1900 the value of our gender equality measure at the world level was 56, it reached a value of 70 (out of a possible 100) in the 2010s, well short of the theoretical maximum.

Looking at regional averages reveals differences. The highest gender equality scores are achieved in Western Europe, the Western Offshoots and East Asia. Gender equality in other regions, particularly the Middle East and North Africa and South and Southeast Asia, remain well below the world average. Latin America shows similar trends as the world average. Sub-Saharan Africa reaches the world average around the 1960s, but in the period since then it falls slightly below the global mean.

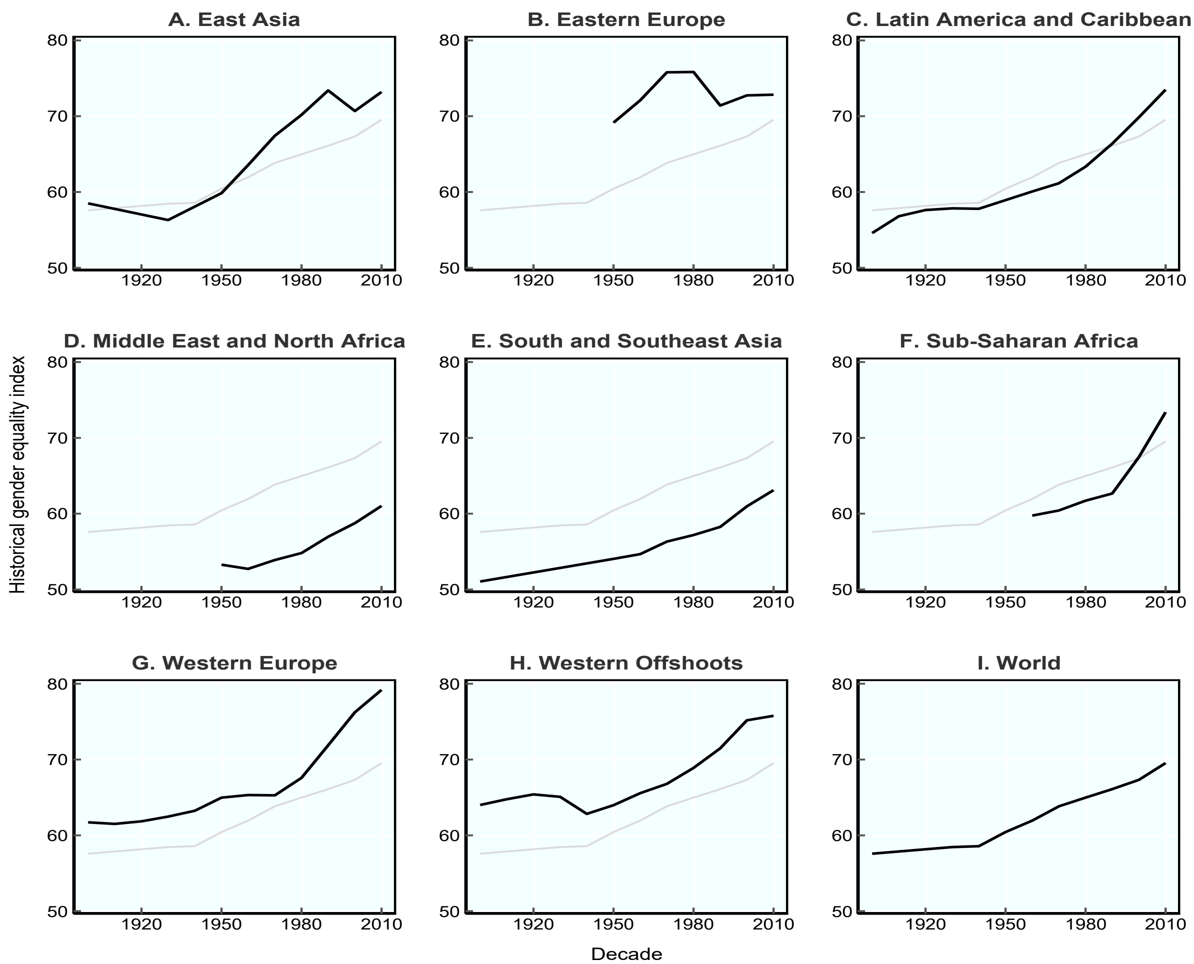
Figure 8.5 shows that Western Europe, Eastern Europe and the former Soviet Union and the Western Offshoots were outperforming the rest of the world at least since the middle of the 20th century, and were closing the gender gap. This strong performance reflects higher levels of female parliamentary representation, higher female labour force participation and a higher female-to-male life expectancy ratio. For the latter measure, this high performance partly reflects the lower life expectancy of men.

For the other indicators, an interesting question is to what extent these developments reflected the existence of socialist regimes in Eastern Europe and the former Soviet Union after the World Wars. Unfortunately, data coverage before 1950 is insufficient to assess regional averages in the pre-socialist period. At the same time, levels of gender equality in individual indicators (sex ratios, schooling ratios and labour force participation ratios) in Hungary, Bulgaria, Romania, Poland and the Czech Republic show performance equal to or above the global average performance already in the first half of the 20th century. However, the performance of these countries on the HGEI deteriorated after the 1990s, especially due to a drop in female parliamentary representation. It should also be noted that this region was historically characterised by high levels of patriarchy – sex- and age-related dominance within the household (Gruber and Szołtysek, 2015[33]), which suggests strong improvements in gender equality over the course of the 20th century.

Progress outside Europe and the Western Offshoots was limited in the first half of the 20th century, but started to accelerate from the 1940s onwards. The Middle East and North Africa and South and Southeast Asia remain the least gender egalitarian regions throughout the period. Furthermore, although Eastern Europe and the former Soviet Union made substantial progress towards gender equality, this trend reversed after the 1980s. East Asia bucks the trend observed in the rest of Asia, performing above the world average in the second half of the 20th century.

Figure 8.5. Composite Historical Gender Equality Index across world regions, 1900-2010

Decennial averages

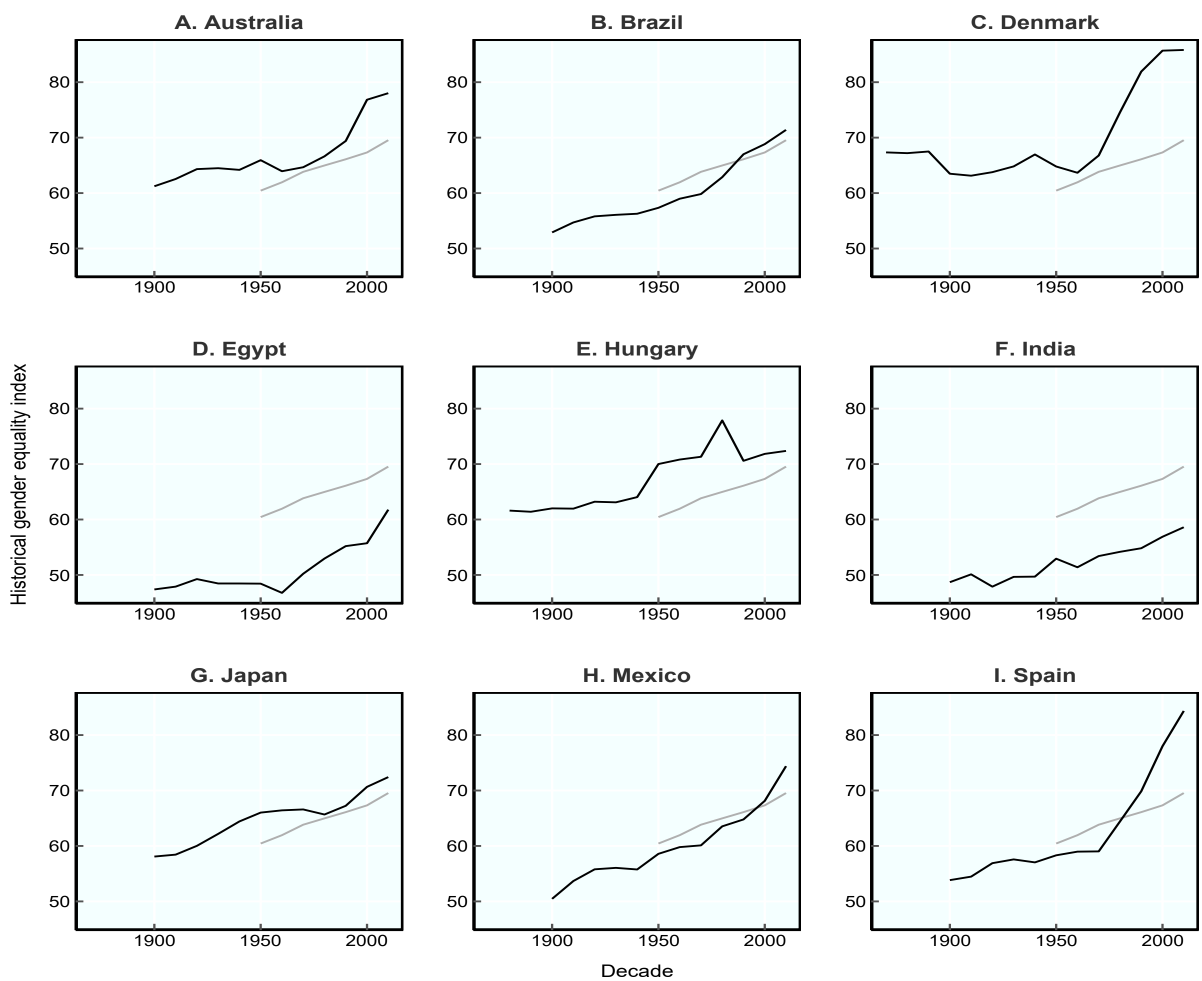


Note: Grey lines show world average.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/0q43jf

Figure 8.6 shows the diversity in countries' progress towards gender equality within regions. Mexico shows a steady increase from the 1910s onwards, which accelerates from the 1960s onwards. In Europe, Hungary and Denmark are above the world average for the entire period, while progress towards gender equality was more limited in Spain until the 1970s, after which it became among the highest-scoring countries in the 2010s. India and Egypt are some of the worst performers in the world throughout the 20th century, with very limited progress. In 2010, these two countries were still far below the world average and even below the values achieved by some countries in 1900. Mexico and Japan are both fairly close to the world average throughout the period.

Figure 8.6. Composite Historical Gender Equality Index in selected countries, 1900-2010



Note: Grey lines show world average.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/ie21lp

Figure 8.7 presents the results of the cluster analysis for the pre-1950 and post-1950 periods. These sub-periods were selected to reflect the moments when progress towards gender equality is visible in different regions (see Figure 8.6), and coincide with those identified in Goldin's (1995[6]) observation of the presence of a U-shaped relationship between economic development and women's labour force participation in the United States over the 20th century. We limit the cluster analysis to 32 countries for which we have data over the entire 20th century. Four distinct clusters emerge from the analysis, with some countries shifting between these clusters and with the difference between clusters getting larger over the 20th century.

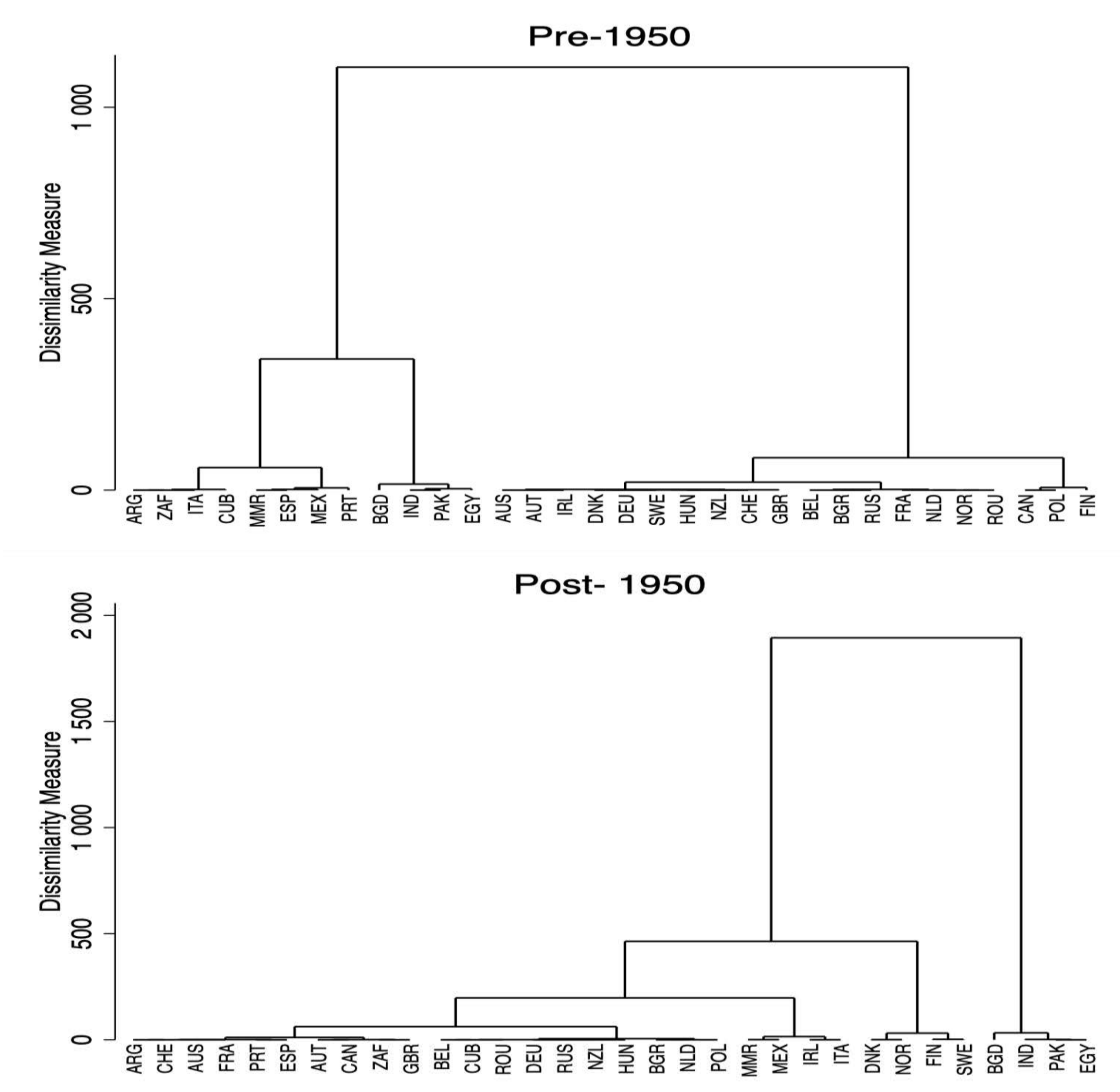
The first cluster is composed of Egypt, Pakistan, India and Bangladesh and identifies the group of countries where women are among the most disadvantaged according to the HGEI. Countries in this group have the lowest female labour force participation and educational attainment, and score the worst in health outcomes in a global comparison. This cluster also includes countries with the largest gap in marriage age between men and women. While these countries made some improvement in a few of these dimensions, such as increasing women's educational attainment and life expectancy since the 1950s, progress has been too limited for each country in this cluster to move out of it in the post-1950 period. This is no surprise given that Pakistan, for instance, scores 50 points on a 100-point scale in the pre-1950 period, and only 53 points in the post-1950 period. In both Pakistan and Bangladesh, gender inequalities continue to manifest themselves in the most extreme version of the "missing girls" phenomenon (OECD, 2019[10]). In 2010, 2 020 000 and 2 907 000 girls were estimated to be missing based on the demographic reports in Bangladesh and Pakistan respectively (United Nations Fund for Population Activities, 2013[34]).

The second cluster is made up of countries with the second-largest gender inequalities in various dimensions of daily life. An important characteristic of this cluster is that the countries move closer to those where gender inequalities are less pronounced, and further away from the first cluster where women face the largest disadvantages in all dimensions of gender inequalities. This cluster includes the Southern European countries (Spain, Italy, Portugal), Latin American countries (Mexico, Cuba, Argentina), South Africa and Myanmar in the pre-1950 period. Countries in this cluster have moderately high gender differences in education and labour force participation, which distinguishes this cluster from the better-performing countries in the third and fourth clusters in the pre-1950 period. In contrast to the first cluster, countries in this group feature relatively small differences in the health and household dimensions and show similar performance to the other countries in the third and fourth clusters. As a result of improvements in education and labour force participation, Spain and Portugal move into the third cluster in the post-1950 period. For instance, Spain, which scores on average 56 points on 100-point scale in the pre-1950 period, scores 69 points in the post-1950 period, a value similar to Australia, France and Switzerland in the same period. Argentina and Cuba also make a similar shift to the third cluster as a result of the increase in women's political representation and education.

The third cluster is composed of moderately gender-equal countries. This cluster includes most countries in continental Europe (Austria, the Netherlands, Belgium, Germany, France, Sweden, Denmark), Eastern Europe (Russia, Bulgaria, Hungary, the Czech Republic, Romania) and English-speaking countries (United Kingdom, Ireland, Australia, New Zealand) in the pre-1950 period. This cluster is characterised by smaller gender inequalities in health, marriage age and socio-economic resources than those in the second cluster. The lowest-performing country in this cluster is Belgium, which starting from a score of 62 in the pre-1950 period reached 72 in the post-1950 period. Also in the post-1950 period, Sweden, Norway and Denmark moved from this cluster to the top-performing fourth cluster because of progress they made towards gender equality in both socio-economic resources and politics.

The fourth cluster includes the most gender-equal countries. While in the pre-1950 period, it includes only Finland, Canada and Poland, Canada and Poland move to the third cluster in the post-1950 period, while Sweden, Norway and Denmark join the group of top performers on account of strong gains in women's parliamentary representation. However, none of the countries in this cluster gets close to the 100 score, with Sweden attaining the highest scores (90 in the 2000s; 92 in the 2010s).

Figure 8.7. Cluster analysis based on the HGEI in pre- and post-1950 periods



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

StatLink https://stat.link/zbmhil

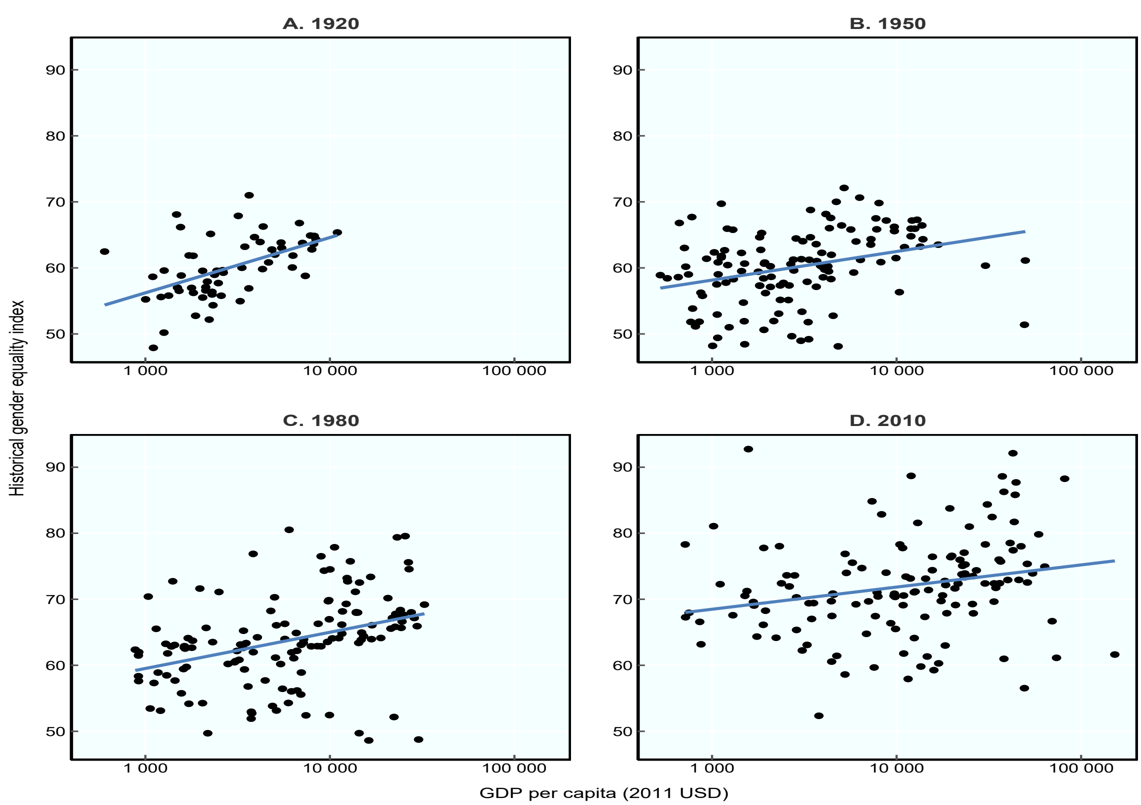
Correlation with GDP per capita

In recent decades, gender equality has received attention not only as a valuable development outcome in itself but also as a crucial contributor to long-term economic development. Figure 8.8 provides an illustration of the relation between gender equality and GDP per capita during the course of the 20th century. A positive correlation between the HGEI and GDP per capita is visible throughout the period. This positive relation is stronger in the first half of the 20th century, while it weakens thereafter. This also implies that a 10% increase in per capita GDP is associated with a 0.4 point increase in the HGEI in the earlier period, but to an increase of only 0.2 points in the latter period. It should however be remembered that many more countries are covered by the HGEI in the second half of the twentieth century than in 1900.

One interpretation of these findings is that they support the U-shaped hypothesis on the relationship between gender equality and economic development postulated by Boserup (1970[35]), who hypothesised that in the early phases of development patriarchal institutions limit both women's occupational opportunities and increase the gender gap. Goldin's U-shaped hypothesis may explain why, in Figure 8.8, a stronger positive relationship between HGEI and GDP per capita is observed in the early 20th century. However, the pattern mainly reflects higher country coverage after the 1950s.

Figure 8.8. HGEI and GDP per capita, 1920-2010

Decennial averages; logarithmic scale



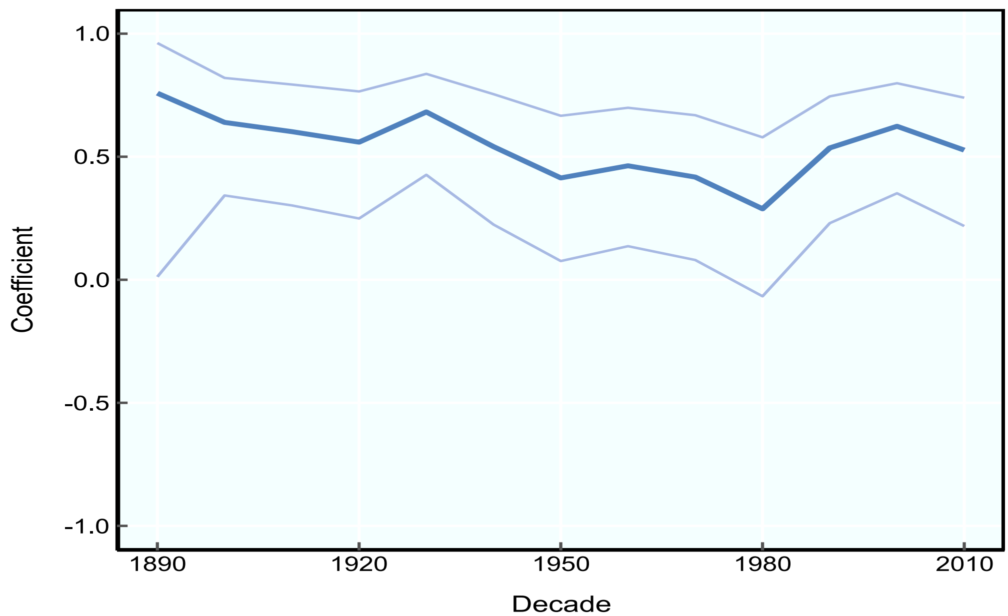
Note: Blue lines show the regression line.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/scb467

When limiting analysis to the 33 countries for which we have data over the whole 20th century, Figure 8.9 shows that the correlation between the HGEI and GDP per capita declines over time from a high of close to 0.8 in 1900 to under 0.5 in 1980 (with a slightly upward movement in the period between the two World Wars). In later years, an increase in the correlation coefficient in the 1980s and 1990s is followed by a small decline in more recent years.

Figure 8.9. Correlation coefficient between the HGEI and GDP per capita for 33 high-coverage countries, 1890-2010

Pearson correlation coefficient



Note: The thick and thin lines depict the correlation coefficient and the 95% confidence interval, respectively.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/q6inge

Priorities for future research

This chapter has provided new evidence on global developments in gender equality over the 20th century. The aim was to improve our understanding of the progress towards gender equality and its drivers. We presented additional data on women's educational attainment and labour force participation, pushing coverage back to 1900. This chapter also presented three new historical indicators of women's position: gender differences in unemployment and in wages, and the share of female-headed households in the second half of the 20th century. We also extended our composite indicator, the Historical Gender Inequality Index (HGEI), back to the beginning of the 20th century.

Despite these additions, historical data on gender inequality remain limited. With regard to the economic position of women, our indicator on wages covers only a few countries and only a selected number of occupations. Work being conducted in Utrecht by the Race to the Bottom team, under the supervision of Elise van Nederveen Meerkerk, should provide additional sources for this measure in the near future, while work by Humphries and Weisdorf (2015[36]) and de Pleijt and van Zanden (2018[37]) has extended our knowledge of wages in earlier historical periods. This is a field where additional historical research would be valuable, including on measures capturing women's income and wealth.

Moreover, coverage of the HGEI prior to the 20th century remains limited to 32 countries across the world, which means that the regional trends presented in this chapter are based on only a few countries. In addition, our country coverage for the early 20th century is larger for the European countries and Western Offshoots than for the world's other regions. To achieve a more representative coverage, the collection of women's labour force participation data in the pre-20th century should be prioritised. Additional research into marriage ages, life expectancy and infant sex-ratios for the period before 1940 for non-European countries would allow extending our information further.

References

Barro, R. and J. Lee (2015), *Education Matters: Global Schooling Gains from the 19th to the 21st Century*, Oxford University Press, New York. [21]

Barro, R. and J. Lee (2013), "A new data set of educational attainment in the world, 1950-2010", *The Journal of Development Economics*, Vol. 104, pp. 184-98, https://doi.org/10.1016/j.jdeveco.2012.10.001. [20]

Boserup, E. (1970), *Women's Role in Economic Development*, Earthscan Publications. [35]

Bradshaw, S., S. Chant and B. Linneker (2017), "Gender and poverty: What we know, don't know, and need to know for agenda 2030", *Gender, Place & Culture*, Vol. 24/12, pp. 1667-88, https://doi.org/10.1080/0966369X.2017.1395821. [15]

Carmichael, S. (2011), "Marriage and power: Age at first marriage and spousal age gap in lesser developed countries", *The History of the Family*, Vol. 16/4, pp. 416-36, https://doi.org/10.1016/j.hisfam.2011.08.002. [24]

Carmichael, S., S. Dili and A. Rijpma (2014), "Gender equality since 1820", in van Zanden, J. et al. (eds.), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://doi.org/10.1787/9789264214262-16-en. [2]

Chant, S. (1985), "Single-parent families: Choice or constraint?", *Development and Change*, Vol. 16/4, pp. 635-56. [12]

de Pleijt, A. and J. van Zanden (2018), "Two worlds of female labour: Gender wage inequality in Western Europe, 1300-1800", *European Historical Economics Society (EHES) Working Papers*, No. 138. [37]

de Zwart, P., B. van Leeuwen and J. van Leeuwen-Li (2014), "Real wages since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-8-en. [27]

Dilli, S., S. Carmichael and A. Rijpma (2019), "Introducing the historical gender equality index", *Feminist Economics*, Vol. 25/1, pp. 31-57, https://doi.org/10.1080/13545701.2018.1442582. [3]

Ferrant, G. and A. Kolev (2016), "Does gender discrimination in social institutions matter for long-term growth?: Cross-country evidence", *OECD Development Centre Working Papers*, No. 330, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jm2hz8dgls6-en. [4]

Ferro, L., Y. Fluckiger and S. Weber (2008), "A cluster analysis of multidimensional poverty in Switzerland", in *Quantitative Approaches to Multidimensional Poverty Measurement*, Palgrave-MacMillan. [31]

Goldin, C. (1995), “The U-shaped female labour force function in economic development and economic history”, in *Investment in Women’s Human Capital and Economic Development*, University of Chicago Press. [6]

Gruber, S. and M. Szołtysek (2015), “The patriarchy index: A comparative study of power relations across historical Europe”, *The History of the Family*, Vol. 0/0, pp. 1-42, https://doi.org/10.1080/1081602X.2014.1001769. [33]

Horrell, S. and J. Humphries (1997), “The origins and expansion of the male breadwinner family: The case of nineteenth-century Britain”, *International Review of Social History*, Vol. 42. [14]

Humphries, J. and J. Weisdorf (2015), “The wages of women in England, 1260-1850”, *The Journal of Economic History*, Vol. 75/2, pp. 405-47. [36]

ILO (2010), *Employment by Age and Sex*, http://www. ilo.org/ilostat. [23]

Inter-parliamentary Union (2011), *“Women in national parliaments”*, http://www.ipu.org/wmn-e/world.htm (accessed on July 2019). [26]

Lee, J. and H. Lee (2016), “Human capital in the long run”, *The Journal of Development Economics*, Vol. 122, pp. 147-69, https://doi.org/10.1016/j.jdeveco.2016.05.006. [22]

Minnesota Population Center (2018), *Integrated Public Use Microdata Series, International: Version 7.0*, https://doi.org/10.18128/D020.V7.0. [28]

Mitchell, B. (2007), *International Historical Statistics 1750-2005*, Palgrave Macmillan. [19]

Mosaic (n.d.), *Recovering Surviving Census Records to Reconstruct Population, Economic, and Cultural History*, http://www.censusmosaic.org/. [29]

Nussbaum, M. (2000), *Women and Human Development: The Capabilities Approach*, Cambridge University Press, Cambridge. [9]

OECD (2019), *SIGI 2019 Global Report: Transforming Challenges into Opportunities*, Social Institutions and Gender Index, OECD Publishing, Paris, https://dx.doi.org/10.1787/bc56d212-en. [10]

OECD (2016), *“OECD report to G7 leaders on women and entrepreneurship. A summary of recent data and policy developments in G7 countries”*, https://www.oecd.org/gender/OECD-Report%20-to-G7-Leaders-on-Women-and-Entrepreneurship.pdf. [1]

Pan, J. (2015), “Gender segregation in occupations: The role of tipping and social interactions”, *Journal of Labor Economics*, Vol. 33/2, pp. 365-408. [32]

Paxton, P., J. Green and M. Hughes (2008), “Women in parliament, 1945-2003: Cross-national dataset”, *ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research*. [25]

Preston, S. (1975), “The changing relation between mortality and level of economic development”, *Population Studies*, Vol. 29/2, pp. 231-48, https://doi.org/10.2307/2173509. [18]

Sen, A. (1999), *Development as Freedom*, Oxford: Oxford University Press. [8]

Sen, A. (1990), “Gender and cooperative conflicts”, in *Persistent Inequalities: Women and World Development*, Oxford University Press, New York. [7]

United Nations Fund for Population Activities (2013), *UNFPA annual report 2012: Promises to keep*, United Nations Fund for Population Activities, New York. [34]

United Nations, D. (2013), *World population prospects: The 2012 revision, key findings and advance tables*, http://esa.un.org/wpp/Excel-Data/population.htm. [17]

Van Driel, F. (1994), *Poor and Powerful – Female-headed households and unmarried motherhood in Botswana*, Nijmegen Studies in Development and Cultural Change. [11]

WEF (2018), *The Global Gender Gap Report 2018*, http://www3.weforum.org/docs/WEF_GGGR_2018.pdf. [5]

WEF (2013), *The Global Gender Gap Report 2013*, http://www3.weforum.org/docs/WEF_GenderGap_Report_2013.pdf. [16]

World Bank (1999), *Vietnam Development Report – Attacking Poverty*, http://documents.worldbank.org/curated/en/929551468780565502/pdf/multi0page.pdf. [13]

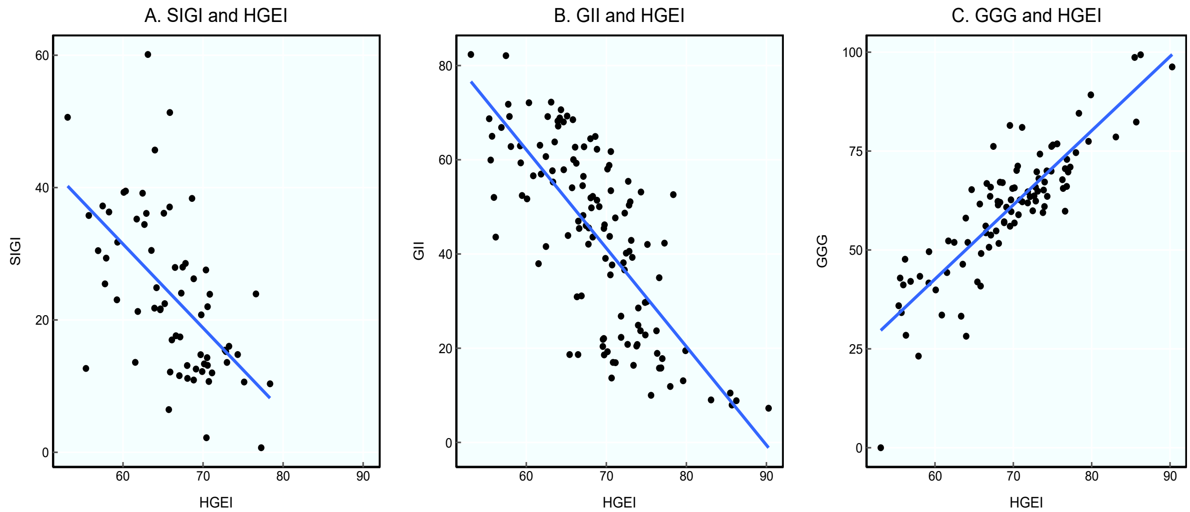
Zeumo, V., A. Tsoukiàs and B. Some (2012), “A new methodology for multidimensional poverty measurement based on the capability approach”, *Cahier du Lamsade*, Vol. 323. [30]

Annex 8.A. Additional evidence

Annex Table 8.A.1. Data quality on the other gender indicators

	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Sub-Saharan Africa	Middle East and North Africa	East Asia	East Asia	South and Southeast Asia
Wages									
1950	2	2	2	2	2	2	2	2	2
1973	2	2	2	2	2	2	2	2	2
2008									
Female-headed households									
1973	2	2	2	2	2	2	2	2	2
2008	2	2	2	2	2	2	2	2	2
Parliamentary representation									
1820	1	1	1	1	1	1	1		1
1870	1	1	1	1	1	1	1		1
1913	1	1	1	1	1	1	1		1
1950	2	2	2	2	2	2	2		2
1973	2	2	2	2	2	2	2		2
2008	2	2	2	2	2	2	2		2
Age at first marriage									
1820	3	3	1	1	1	1	1		1
1870	2	3	2						
1913	2	2	2	2					
1950	1	1	1	1	1	1	1		1
1973	1	1	1	1	1	1	1		1
2008	1	1	1	1	1	1	1		1
Sex ratios among adolescents									
1820	2	2	2	2	2	2	2		2
1870	2	2	2	2	2	2	2		2
1913	2	2	2	2	2	2	2		2
1950	1	1	1	1	1	1	1		1
1973	1	1	1	1	1	1	1		1
2008	1	1	1	1	1	1	1		1

Annex Figure 8.A.1. Relation between the Historical Gender Equality index and the SIGI, GII and GGG in 2000



Note: Blue lines indicate regression fit.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/wrtjeh

Note

[1] The two indicators are negatively correlated, as a higher score on the HGEI indicates more gender equality whereas a higher score on the SIGI refers to a context where institutions discriminate more against women.

9 Global extreme poverty: Present and past since 1820

Michail Moatsos, Utrecht University

This chapter relies on a global data set on basic commodity prices to provide first estimates of global extreme poverty in the long run using a "cost of basic needs" approach.[1] For 135 years since 1820, more than half of the global population lived in conditions of extreme poverty. It took another 46 years to cut this rate in half, which only happened as recently as 2001. In the years that followed, the reduction of extreme poverty accelerated tremendously, and in 13 more years the global poverty rate was halved again. Compared to other available estimates, the world in the 19th century was less poor than we had thought, but poorer in the more recent period. Notably, the total number of people living in conditions of extreme poverty in 1820 stands at 757 million, which is almost identical with the count two centuries later in 2018, at 764 million.

Introduction

Economic growth has spread throughout the planet over the last two hundred years with unprecedented speed, leading to improvements in many well-being indicators, albeit at variable rates (van Zanden et al., 2014[1]). However, economic growth does not necessarily improve the well-being of all people within a country at the same rate, and some may miss the tide altogether. This can be the case when economic inequality within a country rises with, for example, low incomes stagnating and top incomes increasing.

This chapter focuses on the evolution of extreme poverty during the period 1820-2018.[2] Only a handful of attempts have been made to estimate global extreme poverty in the short and medium run (Bhalla, 2002[2]; Bourguignon and Morrisson, 2002[3]; Sala-i-Martin, 2006[4]; Chen and Ravallion, 2010[5]; Ferreira et al., 2015[6]), and even fewer in the long run (Bourguignon and Morrisson, 2002[3]; Van Zanden et al., 2011[7]). All these approaches rely on a single value of the poverty line for all countries and for all years, an approach that many scholars find insufficient (Deaton, 2001[8]; Srinivasan, 2009[9]; Reddy and Pogge, 2010[10]; Atkinson, 2016[11]; Allen, 2017[12]).

The goal of this chapter is to provide global, regional and country estimates of a specific measure of extreme poverty for (almost) all present-day countries of the world as if they were sovereign in 1820. The first step in this exercise is to choose a definition of poverty. This chapter relies on the cost of basic needs (CBN) approach used by Allen (2017[12]). In this approach, poverty lines are calculated for every year and country separately, rather than using a single global line. The second step is to gather the necessary data to operationalise this approach, alongside imputation methods in cases where not all the necessary data are available. The third step is to devise a method for aggregating countries' poverty estimates on a global scale to account for countries that lack some of the relevant data.

The estimates presented in this chapter show that between 1820 and 2018 the prevalence of extreme poverty across the globe fell from 76% to 10%, the lowest level ever achieved, according to our method. This reduction, however, is not distributed evenly throughout the period. It took 136 years from 1820 for our global poverty rate to fall under 50%, then another 45 years to cut this rate in half again by 2001. In the early 21st century, global poverty reduction accelerated, and in 13 years our global measure of extreme poverty was halved again by 2014. Whether this reduction has been sufficient to meet the 50% poverty reduction target set by the first Millennium Development Goal across the developing world between 1990 and 2015 depends decisively upon the treatment of data for pre-1995 China.[3] According to the baseline calculations presented in this chapter, the MDG1 poverty reduction target has been met.

By operationalising Allen's approach on a long-run global scale, this chapter moves beyond the dollar-a-day method towards an approach that provides a set of transparent poverty lines in every year and every country. Arguably, understanding what USD 1.9 at 2011 Purchasing Power Parity (PPP) rates (hereafter simply international dollars) would buy in each country and each year is far from obvious, while the cost of basic needs approach is easily understood by scholars and laymen alike. The cost of basic needs method was recommended by the World Bank Commission on Global Poverty, presided by the late Sir Tony Atkinson, as an alternative (or complementary) method in monitoring poverty for the needs of the United Nations Sustainable Development Goals (Atkinson, 2016[11]).

The following section describes the methodology in some detail. The two sections thereafter discuss the historical sources used and assess data quality, respectively. These are followed by a description of key trends in global extreme poverty since 1820, and of the correlation between these poverty estimates with GDP per capita. The last section conclude by sketching priorities for future research.

Description of the concepts used

All poverty measures require a yardstick (a poverty line) to distinguish those living in conditions of poverty from the rest of the population. The most well-known poverty line in global poverty measurement is the USD 1 a-day line adopted by the World Bank. This measure was originally expressed in 1985 prices and has been updated since then for each new round of PPP estimates by the International Comparison Programme (ICP). Its origin can be traced to Ravallion et al. (1991[13]), who gathered data on poverty lines in national currencies for 33 developing and developed countries. In order to make these national poverty lines comparable across countries, Ravallion converted them to a common currency based on PPP exchange rates provided by Summers and Heston (1988[14]). These estimates showed that the poverty line of the group of low-income countries considered clustered around a value of USD 1 a-day.

Several scholars have expressed concerns about the soundness of these global poverty estimates (Deaton, 2001[8]; Srinivasan, 2009[9]; Atkinson, 2016[11]; Reddy and Pogge, 2010[10]) or about their large margins of error (Atkinson, 2016[11]; Moatsos, 2018[15]). These criticisms mainly focus on the use of PPP exchange rates to estimate the equivalent income of every person on the planet in a single currency and, in particular, on the method used to derive the dollar-a-day poverty line. On the first point, the problem is that a single poverty line may not represent equivalent welfare in different countries and across time (Reddy and Pogge, 2010[10]; Subramanian, 2015[16]; Moatsos, 2017[17]). Reddy and Pogge (2010[10]) argued in favour of a "basic needs approach" in order to resolve the issues raised by the application of PPPs. In response to this criticism, Ravallion (2010[18]) argued that poor people may adapt their consumption habits following changes in market price, implying that the cost of a fixed consumption basket may exceed the cost actually paid for this basket by those living in poverty. Allen (2013[19]), while recognising this substitution effect, argued that this effect is overstated, and advocated the use of linear programming techniques (to estimate the least costly basket of goods) to account for this adaptation to price changes.

Estimates of extreme poverty at the global level are generally produced using data on the distribution of household income or consumption expenditure (often adjusted using non-survey information on mean income or consumption) and a single international poverty line denominated in PPP dollars (Bourguignon and Morrisson, 2002[3]; Bhalla, 2002[2]; Sala-i-Martin, 2006[4]; Chen and Ravallion, 2010[5]; Van Zanden et al., 2011[7]; Ferreira et al., 2015[6]). Recently, Allen, (2017[12]) estimated extreme poverty using linear programming techniques to identify the cost of a diet sufficient to provide a minimum nutritional content, to which additional expenses for housing, clothing and heating are then added. Allen tests four different setups for such minima: 1) a caloric minimum of 1 700 kcal per day; 2) a caloric-protein-fat minima (dubbed CPF) targeting 2 100 kcal, 50g of protein and 34g of fat; 3) a "basic diet" that adds half the quantities of additional nutrients recommended by the Indian daily allowances (RDA), namely iron, folate, thiamine, niacin and vitamins C and B12; and 4) a full-course model that adds 6 more vitamins and minerals to the basic diet. Allen concludes that the "***basic diet***" (option 3) is the more reasonable food standard, since "people eating a CPF diet suffer many nutritional deficiencies", while the full-course model implies unreasonably high quantities of foodstuff to be consumed, and very high caloric intake.[4]

This chapter relies on a "basic diet" poverty line calculated separately for every country and year. However, as the data needed to apply linear programming to estimate a "basic diet" poverty line on a global scale and on the long run are not available,[5] The second-best approach used in this chapter is based on the price data from the ILO October Inquiry (1924-2008).[6] This approach applies linear programming on the ILO data to first calculate the CPF poverty line, which is then multiplied by the coefficient implied by Allen's results to get a proxy of the "basic diet" poverty line.[7] When price level data from the ILO are not available for a particular country and year, the chapter follows the general practice, common to the dollar-a-day approach, of using the national consumer price index (CPI) to extend over time these estimates of the basic diet poverty line.

The types of expenses considered by options 1) to 3) above refer only to the food expenses. To estimate the costs of the final "basic diet" poverty line, additional coefficients expressing the share of food in total household consumption (coefficients that change with the GDP per capita of each country) are estimated from Allen's results in order to move from a "food" to a "full-line" approximating the expenses that people living in extreme poverty in each country and year have to incur.

Historical sources

Four main types of data are required to produce global estimates of extreme poverty with a cost of basic needs approach: 1) a set of prices for different consumption goods (supplemented by general price indices when needed); 2) a set of data on the distribution of consumption or income in each country; 3) a set of population data; and 4) a set of nutrients to be assigned to the various foodstuffs in the price database. The nutrient contents are taken from the US Department of Agriculture (USDA) website (2019[20]), while the population data are from the Maddison Project (Bolt and van Zanden, 2014[21]), complemented, whenever necessary, by data from the United Nations World Population Prospects (2019[22]). The sources for the remaining data are discussed below.

Prices from the ILO October inquiry

Since 1924, the International Labor Organization (ILO) has collected prices for basic foodstuffs and necessary consumables such as bread, rice and fat products in types typically purchased by working families. This data collection started with a small set of 16 capital cities surveyed every October, which grew to cover 130 countries by 1967, dropped to 79 countries in 1983 and then increased again to about 120 countries. The product listings included 15 items in the early years, which incrementally increased to cover 39 main products by 1984. The ILO stopped the collection of prices in 2008 due to a re-organisation of its activities.[8]

A long process of digitising these price data has been undertaken by de Zwart, van Leeuwen and van Leeuwen-Li (2014[23]) using manual entries and Optical Character Recognition (OCR) techniques; in addition, the estimates in this chapter rely on an independent digitisation using a manually assisted OCR approach undertaken by the author. When the two approaches provide different values, the data were investigated for possible errors; when the two approaches gave the same result, digitisations were assumed to be valid.

The October Inquiry commodity price dataset offers data points spanning a period of at least 80 years for 18 countries, and more than 50 years for 108 countries. These data were reported to the ILO by the national statistical agencies of each country, which were instructed to report prices that were relevant for working families. This makes this dataset particularly useful for the calculation of the costs of basic needs poverty lines. As with all historical data, the ILO dataset has its problems. For example, one needs to be careful with currencies before and after a currency redenomination takes place in a country.[9] Another issue is that the October Inquiry dataset refers to prices in a single month, implying that, when inflation is high, these price data may not be representative of the full year. For high inflation years, e.g. in Brazil or Zimbabwe, the estimates in this chapter rely on changes in the average CPI to estimate the poverty line, rather than on the ILO price data. In addition, some ILO entries are clearly too low or too high, by a couple of orders of magnitude, relative to the price entries for the same products for that country in a nearby year; those prices were also removed from the final dataset.[10]

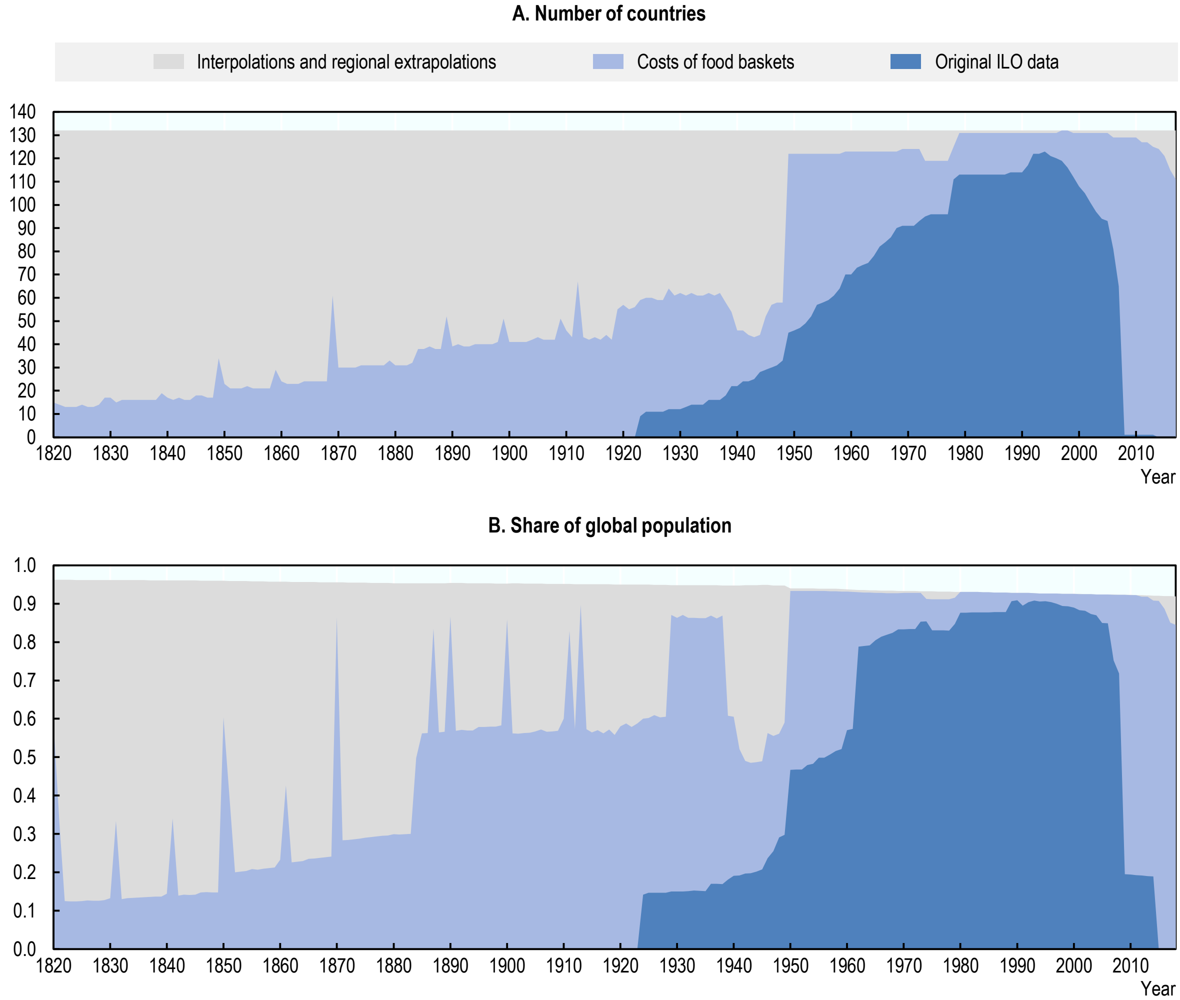
Moreover, prices for all products are not reported for all countries at all times. This means that one needs to impute these missing data using either the CPI index of that country or, when this is also not available, the evolution of similar prices in the ILO dataset.[11] Such an imputation, however, may result in prices that are distant from those actually paid by working families. For example, the prices of the cheapest product providing a given number of calories may follow a different trajectory than the average CPI of a country, and after, say, 10 years one may have a price that is much higher or much lower than the unobserved one. This is an important limitation of our approach.

The same applies for years outside the 1924-2008 period, where the cost of the food component is extrapolated using available consumer price indices. This imputation rests on the assumption that, on average, the cost of the food component does not substantially diverge from the average CPI. This is a strong assumption, but is the same used by the dollar-a-day approach. More work to gather detailed price data will be required to extend the coverage of our price dataset to the full time period covered by this chapter.

Nevertheless, the use of detailed nominal prices brings the methodology of measuring extreme poverty one step closer to the experience of people living in poverty across the world. It removes one layer of assumptions, compared to other approaches that rely on one single poverty line expressed in PPPs, or that rely on a CPI index for moving the poverty line in time for about 20, 50 or 170 years (Bourguignon and Morrisson, 2002[3]; Bhalla, 2002[2]; Sala-i-Martin, 2006[4]; Chen and Ravallion, 2010[5]; Van Zanden et al., 2011[7]; Ferreira et al., 2015[6]).

Panel A in Figure 9.1 shows the number of CPF food components calculated from the complete ILO data, with a dark colour indicating non-imputed data for the food items selected by the linear programming. Panel B shows the same information expressed as percentages of the global population covered by those baskets. Direct estimates of poverty lines based on ILO data start only in 1924, covering around 14% of the global population at the time. For the period 1950-2008, the average population coverage is 78%. Using CPI extrapolation, the average population coverage for the period 1884-2018 is also 78%, while for the period before 1884 population coverage drops to 23%.

Figure 9.1. Country and population coverage of the poverty estimates used in this chapter



Note: Panel A shows the number of food baskets calculated in this chapter based on different methods: the original ILO data (in dark blue), the costs of food baskets imputed via CPI (blue), and those based on interpolations and regional extrapolations (in grey). In both panels, the part shown in white corresponds to counties/share of population not covered by the estimates. (See the section on imputations for details.) Panel B shows the share of the global population covered by these estimates. The spikes in population coverage are due to China and India (large spikes) or India (small spikes) in the 19th century, and China in the early 20th century.

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

StatLink https://stat.link/3riw28

Consumption and income distributions

To compare with the global poverty estimates published by the World Bank, our calculations rely on the data of the distribution of consumption and income available via World Bank's *PovcalNet* (2020[24]). While the combination of consumption and income data obviously reduces cross-country comparability in the *PovcalNet dataset*, the methodology used here does not try to correct for this in order to preserve comparability between these estimates and those produced by the World Bank. The distributional data in *PovcalNet* cover the post-1980 period.

Other sources of distributional information need to be used for the years not covered by *PovcalNet*. These sources typically provide values of the Gini index of the entire distribution, rather than the distribution itself.[12] Thus, a method is required to convert this single statistic into an entire distribution, in order to estimate the share of the population with income or consumption below a poverty line. To do this, the log-normality assumption is used to convert a Gini index to a full distribution that follows a log-normal shape (Aitchison and Brown, 1957[25]; Van Zanden et al., 2011[7]).

In summary, for the most recent period, this chapter relies on the distributions provided by *PovcalNet*. For the period between 1950 and the year when *PovcalNet* coverage begins for a particular country, the chapter uses a variety of sources[13] that, for many countries, rely on household income surveys typically conducted by national authorities.[14] Preference is given to distributions of disposable income, followed by those distributions capturing gross (i.e. pre-tax) income, and finally distributions that do not specify the exact type of income they capture; in cases when several estimates of comparable quality are available, their average is used. The chapter also relies on databases providing synthetic Gini estimates[15] to provide information on the evolution of the Gini index between observations from the aforementioned sources.[16] For the period prior to 1950, inequality data are sourced from van Zanden et al. (2011[7]) and other studies, such as Bolt and Hillbo (2016[26]), Alfani and Tadei (2017[27]), Prados de la Escosura (2008[28]) and Kang (2001[29]).

Estimating unobserved household mean income

After having derived an estimate of the distribution of income or consumption in a country, one needs to "pin down" the dispersion of the distribution around a mean in all cases where this information is not available in the source used. Information about this mean can be drawn from either National Account statistics or from the same household survey that provides information about its distribution. Starting off from *PovcalNet* provides an advantage in this respect, as the values of the household survey mean are known and can be extrapolated back in time using consumption or GDP data.

Deaton (2001[8]) reports that, in the case of India, the ratio of mean household consumption measured by the System of National Accounts (SNA) and by the main survey on the distribution of household expenditures (the National Sample Survey) fell from almost unity in 1950 to about 50% by the late 1990s. Ferreira et al. (2015[6]) correct for this divergence between mean consumption growth from the surveys and from the National Accounts (by using 87% of the SNA consumption growth rate for all countries, 51% for India and 72% for China). The same approach is used here, while also taking account of the time effect noted by Deaton. To that end, the conversion factors of (Ferreira et al., 2015[6]) for 2000, linearly extrapolated to assume a value of 1 by 1950, are used; this implies that, prior to 1950, mean income from the survey is assumed to grow at the same pace as that from the SNA.

Imputations

To generate estimates of extreme poverty across the globe, imputations are needed for missing countries and years in order to avoid abrupt changes in coverage. In this chapter, imputations are based on the change in the average poverty rate of the countries in that region for which there are available data.[17] This approach avoids the downward bias associated with the greater availability of data for rich countries in historical times.[18] In addition, when enough data are available for a country at two distant points in time, the poverty rate between those two points is linearly interpolated – instead of regionally imputed. Such an approach is applied mostly to data during the 19th century.[19]

Data quality

Table 9.1 shows my assessment of the quality of poverty estimates for the eight world regions for a selection of years, in terms of credibility, accuracy and comparability. Among these dimensions, the degree of credibility is relatively high throughout the period, while the degree of accuracy is lower, especially when price data refer to non-market economies (China and the former Soviet Union among others); in these cases, no consumption goods may be available at these administrative prices, hence violating a basic assumption of the cost of basic needs approach.[20] A similar problem arises when the average consumer price index is used instead of detailed product prices. This issue, however, is easier to treat, simply with more data collection in the future. With these caveats in mind, my assessment is that the degree of comparability is more or less sufficient, in spite of differences across countries in the underlying income distributions data.

Table 9.1. Quality of data on extreme poverty by world region and year

Year	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Sub-Saharan Africa	Middle East and North Africa	East Asia	South and Southeast Asia
1820	4	4	4	4	4	4	4	4
1870	4	4	4	4	4	4	4	4
1920	3	4	3	4	4	4	4	4
1950	2	4	2	2	2	2	4	2
1980	1	4	1	1	1	1	4	1
1990	1	4	1	1	1	1	4	1
2010	1	1	1	1	1	1	1	1

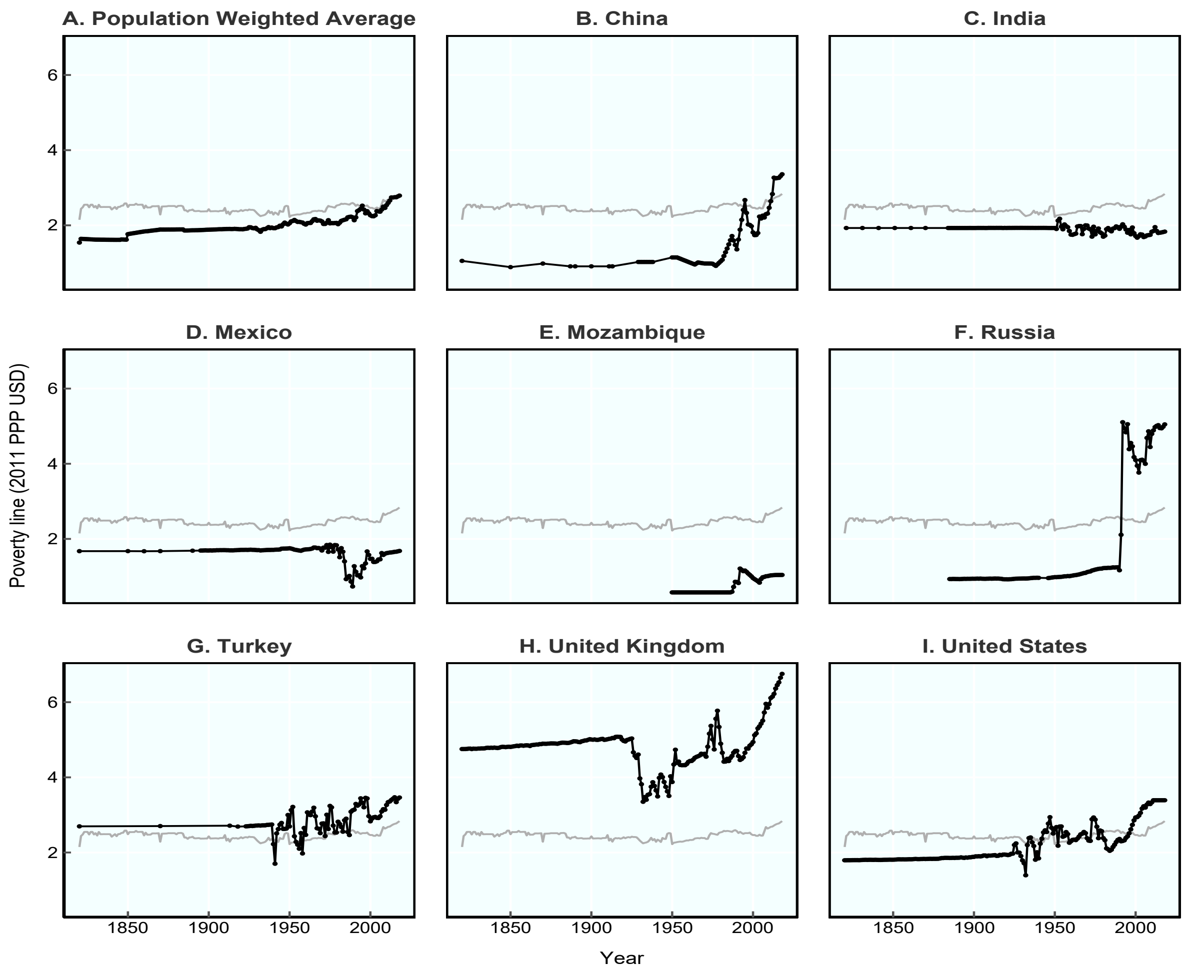
Note: 1. High quality: the product of an official statistical agency (national or international); 2. Medium quality: the product of economic-historical research using the same sources and methods as applied by official statistical agencies; 3. Moderate quality: economic historical research, but making use of indirect data and estimates; and 4. Low quality: estimates based on a range of proxy information. In case of multiple sources, the lowest quality source is given.

All the basic needs poverty lines estimated are shown in Figure 9.2. These poverty lines are based on linear programming to derive to the CPF line, to which relevant multipliers are applied for obtaining the basic diet poverty line and the non-food component. Poverty lines are estimated based on the ILO data,[21] and extrapolated based on the CPI.[22]

When looking at our poverty lines throughout historical times, the United Kingdom has the highest poverty line in PPP terms at just below USD 5/day during the long 19th century (Figure 9.2). The poverty line for the United States is very close to USD 1.9/day until World War II, increasing thereafter.[23] India has a poverty line that is very close to USD 1.9/day, for the entire period. Mexico begins with a poverty line slightly below USD 1.9/day, falling below USD 1/day in the early 1980s, but gradually returning to USD 1.9 by 2018. In the case of Russia, the poverty line shows a huge increase following the dissolution of the Soviet Union in 1991. The poverty line is very volatile in Turkey (starting at USD 2.7/day in 1924 and rising to USD 3.5 by 2018), while the lowest poverty line plotted here is Mozambique's, at around USD 0.6/day.[24]

On average, these lines for extreme poverty hover around USD 2.5/day for almost the entire period, with only the years after 2010 showing a clear increase, reaching a maximum of USD 2.8/day in 2018.[25] The population-weighted average poverty line has, however, a different trajectory. Starting from a level slightly below USD 1.9/day in the early 19th century, it is broadly constant from 1850 to the early 1940s, increasing thereafter to USD 2.8/day by 2018. It is clear that the values of these poverty lines vary considerably even for the most recent years, ranging in 2018 between USD 1 and USD 7 (Figure 9.1), which is in line with the findings of (Allen, 2017[12]; Hirvonen et al., 2019[30]).

Figure 9.2. Lines for extreme poverty, selected countries and years



Note: Grey reference line provides average of countries' poverty lines.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/g5vd4r

Main highlights on trends in extreme poverty

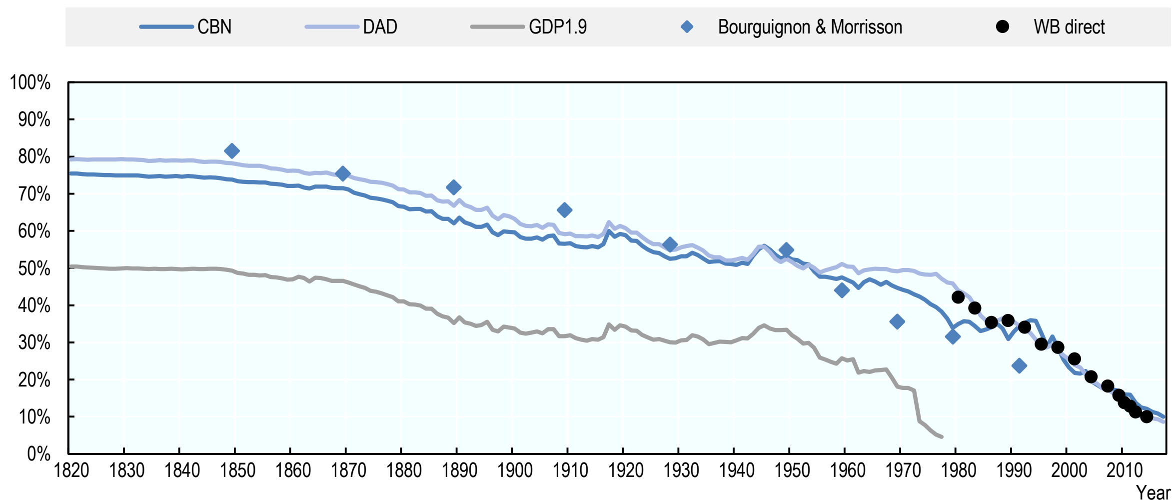
The figures in this section show the evolution of extreme poverty rates across the globe based on the cost of basic needs (CBN) method, alongside those based on other methods. Generally, three types of poverty estimates are shown. The main line, shown in dark blue, is the poverty rate based on the CBN methodology used in this chapter. The other lines correspond to the USD 1.9/day line (noted as DAD) and the USD 1.9/day poverty line applied to distribution data centred on GDP per capita as the mean value of the distribution (shown as GDP 1.9 prior to around 1979, i.e. before the *PovcalNet data*).[26] Estimates of extreme poverty at the world level from (Bourguignon and Morrisson, 2002[3]) are also included in Figure 9.3.

Global extreme poverty

Both the CBN and the DAD approaches suggest a bleak picture of extreme poverty in 1820, at 76% and 79% respectively (Figure 9.3). (Bourguignon and Morrisson, 2002[3]), with less available data and using 1985 PPP exchange rates, estimate a higher value at 84%.

Figure 9.3. Extreme poverty around the world based on different estimates

Global poverty rate



Note: Share of people living in conditions of extreme poverty globally, based on different estimates. CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution. The figure also shows the estimates reported in Bourguignon and Morrisson (2002).
Source: Clio-Infra, www.clio-infra.eu.

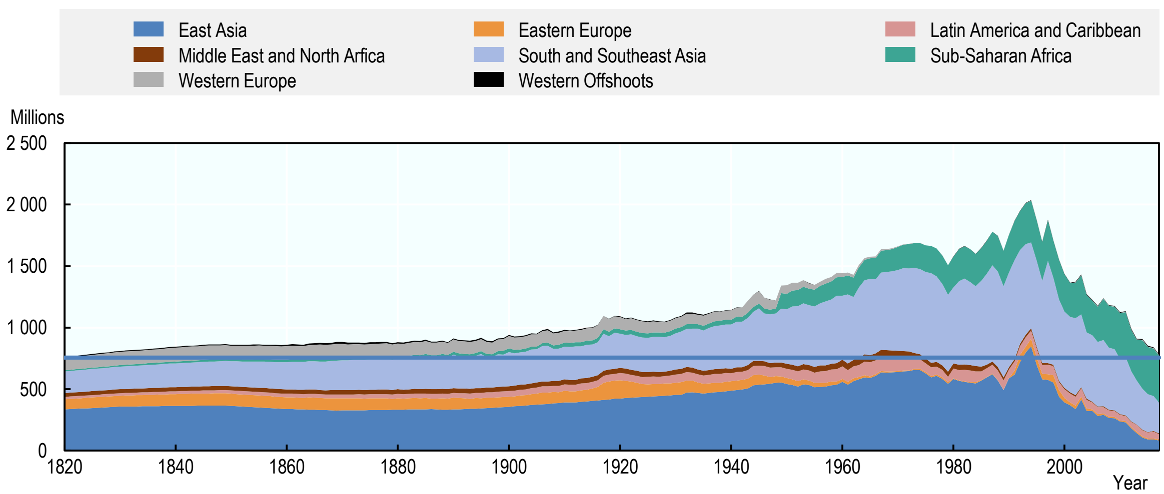
StatLink https://stat.link/6q543k

According to these measures, in 1820 roughly three-quarters of the world population, about 756 million people, could not afford a tiny space to live, food that would not induce malnutrition, and some minimum heating capacity. This total number of persons living in extreme poverty is almost identical to the estimate for 2018, which stands at 764 million people. Over the entire period between 1820 and 2018, the the number of people living in extreme poverty reached a peak of 2 billion in 1995.

By 2018, global extreme poverty dropped to 10%. Based on our methodology, the global poverty rate fell below 70% in 1873, and below 60% by 1897; after that, it takes much longer to drop below 50% by 1955, then much less time to drop below 40% by 1977. In another 20 years global extreme poverty dropped below 30% in 1997, while the 20% barrier was passed only 8 year later in 2005. The fastest drop in the entire period takes place between 1995 and 2000. During the World War years (I and II), extreme poverty rates show a small increase at the global level.

The global total count and the geographical distribution of those living in conditions of extreme poverty across the globe is shown in Figure 9.4. East Asia accounted for the largest share until 1962 but was overtaken by South and Southeast Asia thereafter. The upward trend in the global count of the population living in extreme poverty was interrupted decisively only in 1995, with a few other noticeable but transitory corrections in 1917, 1947 and 1976. At its peak in 1995, the total number of people living in extreme poverty was 2.7 times that of 1820. Both the overall increase in the total number of poor people by 1995 and the temporary decreases in the number following the two World Wars are clearly visible. The sharp increase in 1950 corresponds to the inclusion of several sub-Saharan African countries only from that year onward.

Figure 9.4. Geographical distribution of people living in extreme poverty across the globe, 1820-2018



Note: The horizontal line corresponds to the total population in extreme poverty in 1820.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/98otxe

Table 9.2 provides an overview of extreme poverty rates across world regions in some benchmark years throughout the period. Based on these measures, Eastern Europe and the former Soviet Union achieved the largest reduction in extreme poverty, from 91% in 1820 to 2% in 2018. Sub-Saharan Africa, which started from the same level in 1820, stands at 37% in 2018. Western Europe in 1820 had a higher prevalence of extreme poverty (73%) than South and Southeast Asia (69%).

Table 9.2. Extreme poverty across world regions

Percentages

Year	Western Europe	Eastern Europe	Western Offshoots	Latin America and Caribbean	Sub-Saharan Africa	Middle East and North Africa	East Asia	South and Southeast Asia
1820	73	91	55	85	91	75	78	69
1850	65	84	28	84	95	71	80	69
1870	59	68	31	80	93	68	80	74
1900	41	43	9	74	91	59	76	66
1920	38	63	3	61	87	55	77	62
1950	21	19	1	45	70	33	82	67
1960	11	8	0	45	67	34	66	69
1970	2	2	0	42	62	26	64	64
1980	0	3	1	12	63	17	46	50
1990	0	1	1	17	58	7	37	47
2000	0	13	1	12	50	3	29	35
2010	1	3	1	8	41	1	17	23
2018	1	2	1	7	37	1	5	10

StatLink https://stat.link/mpb5dn

Table 9.3. Cost of basic needs extreme poverty rates for major countries

Percentages

Year	GBR	NLD	FRA	DEU	ITA	ESP	SWE	POL	RUS	AUS	CAN	USA	MEX	BRA	ARG	EGY	TUR	KEN	NGA	ZAF	CHN	JPN	IND	IDN	THA
1820	80	75	77	66	75	83	86	64	98	87	74	52	86	91	63	60	90	92	92	86	76	95	65	96	82
1830	78	74	75	64	75	81	83	56	90	69	69	43	87	92	56	56	90	88	88	82	76	96	65	97	82
1840	78	67	72	63	74	78	82	58	90	45	58	34	87	92	48	52	89	93	93	86	77	97	65	97	82
1850	69	64	67	59	75	74	78	54	85	28	44	26	87	93	41	48	88	96	96	90	78	98	65	99	81
1860	68	65	63	55	75	67	72	36	70	15	41	28	89	90	45	44	88	89	89	83	78	97	67	98	81
1870	55	63	65	53	76	71	71	30	65	17	35	31	84	87	51	40	87	98	98	83	77	97	70	98	81
1880	48	56	61	45	77	52	68	17	49	7	35	18	77	89	43	36	86	70	70	54	76	96	68	97	75
1890	35	47	53	31	74	51	66	7	46	3	28	9	70	88	23	32	85	90	90	67	73	97	63	97	69
1900	32	47	51	25	72	53	55	3	31	5	19	9	66	95	21	27	84	100	100	77	74	92	61	95	68
1910	35	44	47	24	65	46	46	2	27	4	9	8	60	90	17	23	82	100	100	69	77	90	51	93	67
1920	33	36	45	32	45	51	34	1	70	4	10	2	56	76	18	21	92	100	100	70	76	83	57	90	69
1930	20	19	51	22	47	42	19	1	27	1	7	4	56	77	12	18	78	97	97	68	77	80	48	86	71
1940	6	12	49	14	40	42	19	0	11	0	2	1	50	89	6	16	66	100	100	52	79	64	53	85	68
1950	2	8	30	21	37	34	4	0	3	0	0	1	37	72	2	14	62	81	99	40	84	73	61	91	67
1960	4	5	16	5	20	20	2	3	0	0	0	0	30	73	1	14	51	74	94	38	73	23	66	90	60
1970	0	0	0	2	3	5	1	0	0	0	0	0	24	76	0	12	45	66	94	24	75	2	67	73	40
1980	0	0	0	0	0	0	0	0	0	0	0	1	6	11	2	1	32	62	95	15	53	0	49	40	32
1990	0	0	0	0	0	0	0	0	0	1	0	0	3	23	2	0	12	54	72	29	42	0	47	41	18
2000	0	0	0	0	0	0	0	1	18	1	1	1	5	12	9	0	8	65	71	28	33	0	32	42	10
2010	1	0	0	0	2	1	1	1	2	1	0	1	3	8	2	0	5	61	64	18	19	1	23	25	3
2018	1	0	0	0	3	2	1	0	2	2	1	1	2	7	0	0	2	61	62	21	5	1	9	11	1

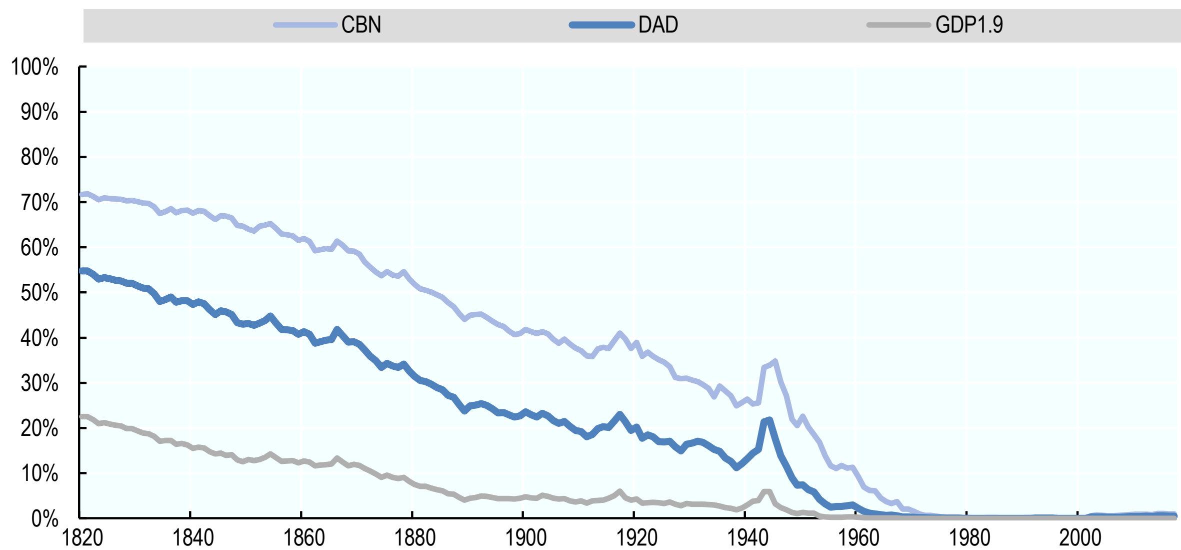
StatLink https://stat.link/lrxq9t

Extreme poverty in world regions

Extreme poverty in Western Europe (Figure 9.5) fell continuously for the entire period until the late 1970s, when it is almost zero. This trend is interrupted twice by both World Wars, with the impact of World War II being much more pronounced. At 41%, the post-WWI peak is roughly 5 percentage points higher than the minimum achieved in 1913. The post-WWII peak, at 35%, is roughly 10 percentage points higher than the minimum in 1939.[27] Since 2003, extreme poverty rates in Western Europe have experienced an upward creep, to a level slightly above 1% in 2015. The difference in extreme poverty between the CBN and DAD estimates for Western Europe is substantial, at around 15 percentage points for most of the pre-WWII period, possibly reflecting higher prices for more expensive services such as rents in the CBN estimates.

Figure 9.5. Extreme poverty in Western Europe based on different estimates

Poverty rate



Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

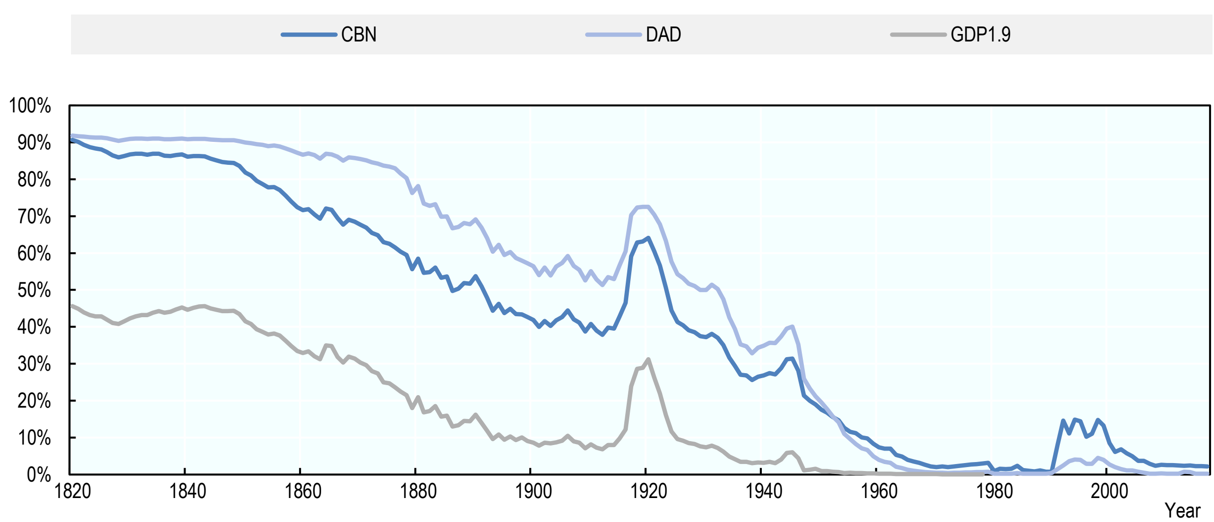
StatLink https://stat.link/5imp1q

The largest share of the total count of extreme poverty for the region in 1820 (101 million) comes from France (24 million), followed by the United Kingdom (17), Germany (16.5), Italy (15) and Spain (10). By 1900, Italy accounts for 24 million persons in extreme poverty, followed by France (20), Germany (13.7), the United Kingdom (13) and Spain (10). In 2018, Italy accounts for 1.7 million of the 4.3 million extreme poor in the region, followed by Spain (0.9), the United Kingdom (0.7), Germany (0.2) and Greece (0.2).

Figure 9.6, which describes the evolution of extreme poverty rates for Eastern Europe and the former Soviet Union, highlights three periods of abrupt changes: two after the World Wars and one after the dissolution of the Soviet Union. In 1820, this region had a poverty rate of 91%, practically the same rate as sub-Saharan Africa in the same year. A period of steady reduction starts in 1830, bringing it to 84% by 1850. In 1851 a strong – although not entirely uninterrupted – decline in poverty starts, which is faster than the one realised in Western Europe in this period, bringing the poverty rate to 38% by 1913. A remarkable increase in extreme poverty took place during WWI, which brings poverty to a high of 64% in 1921. A period of exceptional poverty decline starts from 1922, which brings extreme poverty to 1% in 1990, after a strong but temporary increase after World War II.[28] Extreme poverty increases again after the dissolution of the Soviet Union, hovering at 10-15% until 2000, and then continuously declining to 2% by 2018.

Figure 9.6. Extreme poverty in Eastern Europe and the former Soviet Union based on different estimates

Poverty rate



Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/uwf1gq

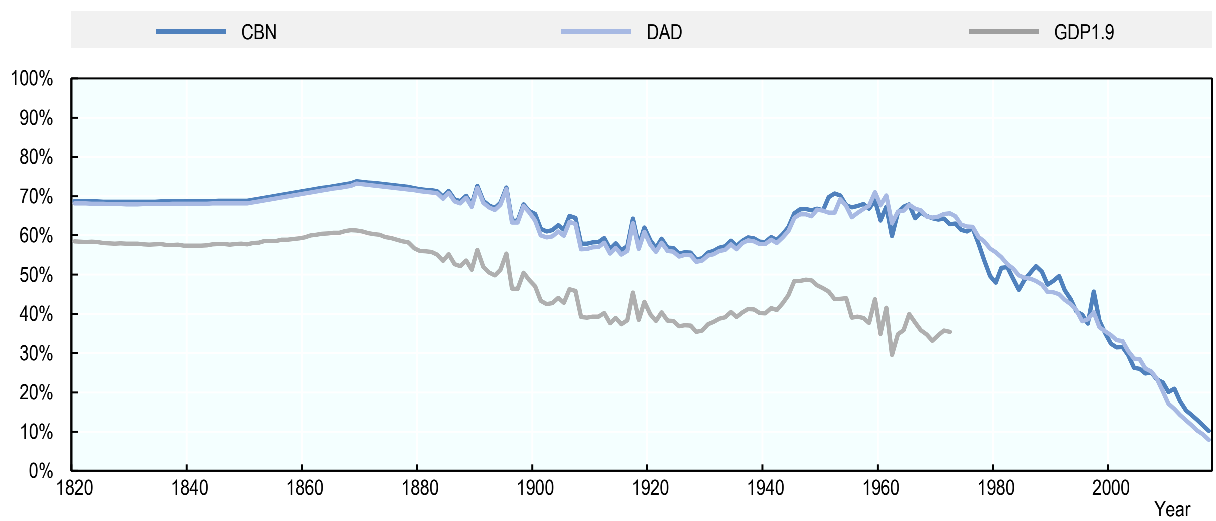
In terms of country level contributions to the world total, Russia was the biggest contributor in 1820, with approximately 31 million, followed by Ukraine (11), Poland (6.7), Romania (6.4) and the Czech Republic (5). By 1990, Ukraine contributes 24 million (almost its entire population), Russia (22), Romania (5.5), the Czech Republic (5.2) and Bulgaria (2.8).[29] In 1995, the peak year after the Soviet Union dissolution, Russia contributes 33.5, followed by Ukraine (8), Moldova (2), Armenia (1.2) and Poland (0.7).

Figure 9.7 shows the evolution of extreme poverty in South and Southeast Asia, a region that (population-wise) is dominated by India. Prior to the 1890s, data are very scarce and only available for a few countries in a few benchmark years; the linear interpolations for the best part of the 19th century are evident in the figure. In any case, the available data show a poverty rate of 69% in 1820, a level more or less unchanged until the 1890s. A volatile period between 1884 and the late 1920s brings extreme poverty to a minimum of 54%. The upward trend that follows takes the poverty rate back up to 71% in 1953, a higher level than the one recorded in the 1820s. Despite a temporary surge in 1998 to 46%, extreme poverty declined to a low of 10% in 2018.

In terms of absolute contributions, India accounts for 136.6 million people in global poverty counts in 1820, Indonesia contributes 17.3 million, followed by Thailand (3.8) and Nepal (3.2). By 1950, India contributes 217.9 million people to the global poverty counts, followed by Indonesia (75.1), Bangladesh (30.1), Pakistan (28.3) and Myanmar (17.4). In 2018, India contributes 126.9 million people, followed by Bangladesh (55.2), Indonesia (29), Pakistan (11) and the Philippines (6.4).

Figure 9.7. Extreme poverty in South and Southeast Asia based on different estimates

Poverty rate



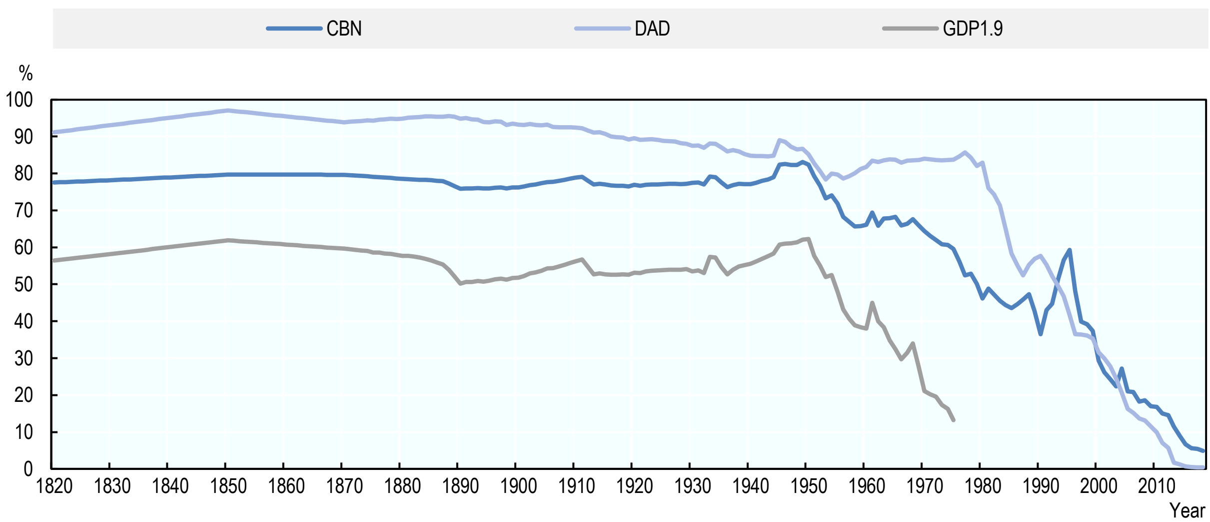
Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/3wuetv

Figure 9.8 shows the evolution of extreme poverty in East Asia.[30] East Asia is the only region where extreme poverty reached its peak in the 20th century, at 83% in 1949. The poverty reduction during the post-WWII period stalled in 1958, with a local maximum of 69% in 1961, and then resumed its fall after 1969, to 37% in 1990. Since then, extreme poverty increased to a new peak of 59% in 1995, only marginally lower than the estimate for sub-Saharan Africa in the same year, before falling to an overall low of 5% in 2018.

Figure 9.8. Extreme poverty in East Asia based on different estimates

Poverty rate



Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/0ytwvi

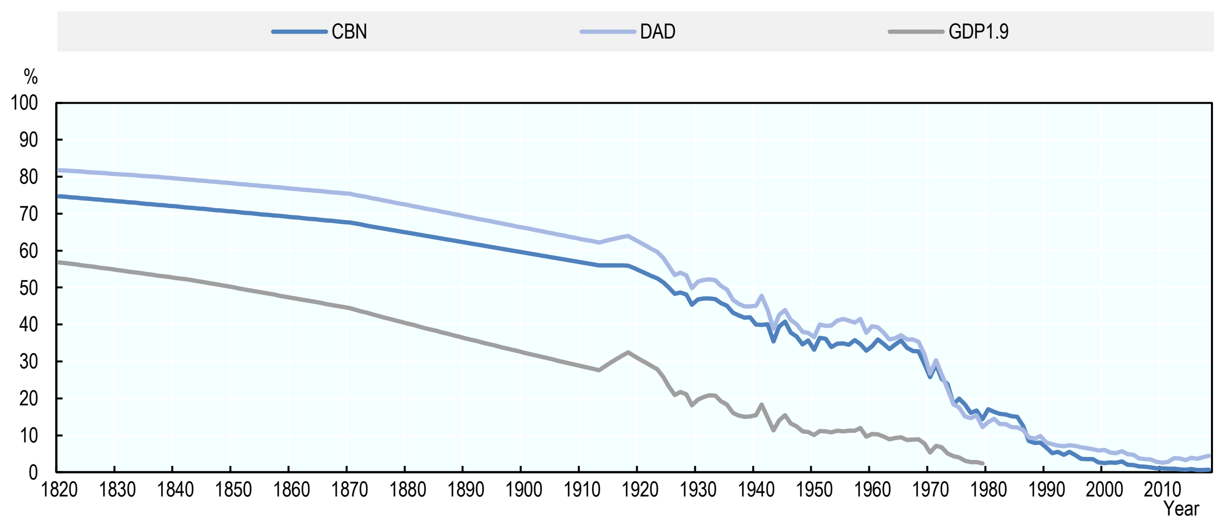
As expected, China is by far the largest contributor in absolute terms throughout the period. In 1820, China contributes 288 million people to the global extreme poverty count, followed by Japan (29.4), Korea (9.2) and Mongolia (0.6). At the local high of poverty in 1995, China contributes 815.8 million people in global poverty counts, followed by Mongolia (1.4).[31] In 2018, China contributes 77.1 million people, followed by Japan (1.4), Mongolia (0.5) and Korea (0.3).

Figure 9.9 shows the development of extreme poverty throughout the Middle East and North Africa region. For the best part of the period until 1970, both the DAD and CBN series are relatively close, and at times very close. Extreme poverty fell in the 19th century, from 75% in 1820 to 59% by 1900. The pace of reduction is slightly faster in the first half of the 20th century, to 18% by 1974. A sharp decrease follows in the 1980s, bringing extreme poverty to 7% in 1990, and to 1% in 2018.[32]

In terms of the countries' contributions to the global count of extreme poverty, Turkey is the largest contributor in 1820 with approximately 9.1 million, followed by Sudan (5.2), Iran (4.5) and Algeria (1.2). By 1974, Turkey contributes 10.2 million, Sudan (10), Morocco (5.4) and Algeria (1.6). In 2018, Turkey is again the largest contributor, but at considerably lower levels, with 1.7 million people, followed by Yemen (0.8),[33] Sudan (0.3) and Algeria (0.1).

Figure 9.9. Extreme poverty in the Middle East and North Africa based on different estimates

Poverty rate



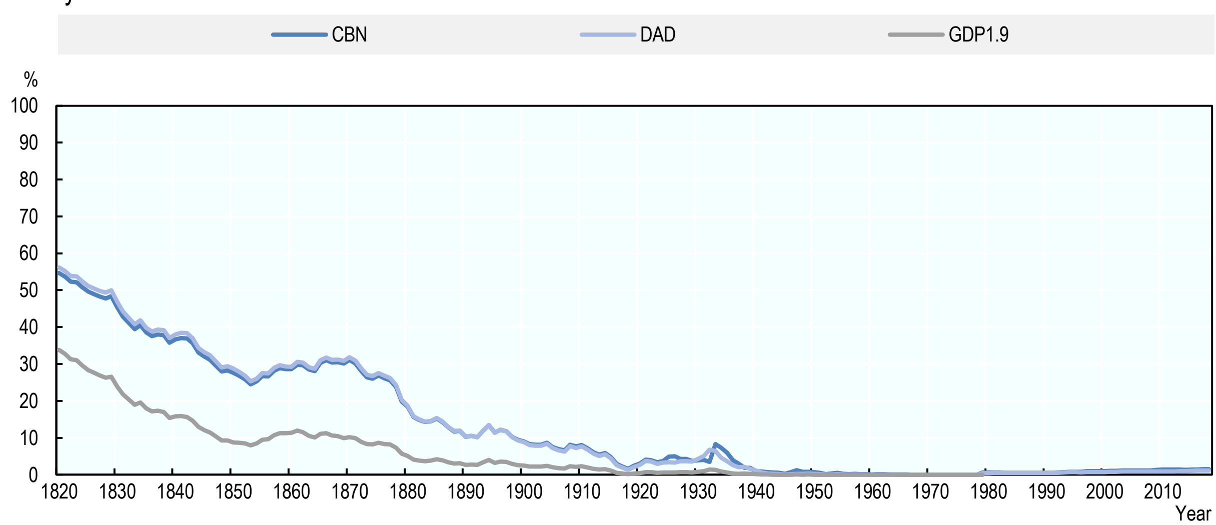
Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/vfzd8o

Figure 9.10 shows the evolution of extreme poverty in the Western Offshoots region. This region has the lowest poverty rates throughout the period. Starting from a level of 55% in 1820, the decline in poverty rate comes to a halt in 1853, and resumes after the US Civil War until 1918, when it stands at 2%. Following the 1929 financial crisis, extreme poverty reaches a peak of 8% in 1933, more than double its level in 1924. During WWII, the Western Offshoots shows no sign of a poverty increase, and its rate in 1944 (at 1%) is the world's lowest. In the years that follow, poverty rates are less than 0.5% until 1979. The period ends with a rise in extreme poverty to a rate above 1% after 2000. In terms of absolute contributions to global poverty, the United States contributes 5.2 million in 1820, followed by Canada (0.6) and Australia (0.3). By 2018, and despite the large reduction in poverty rates, the contributions are almost identical to those of 1820, with the United States contributing 4.9 million, followed by Canada (0.5) and Australia (0.4).

Figure 9.10. Extreme poverty in the Western Offshoots based on different estimates

Poverty rate



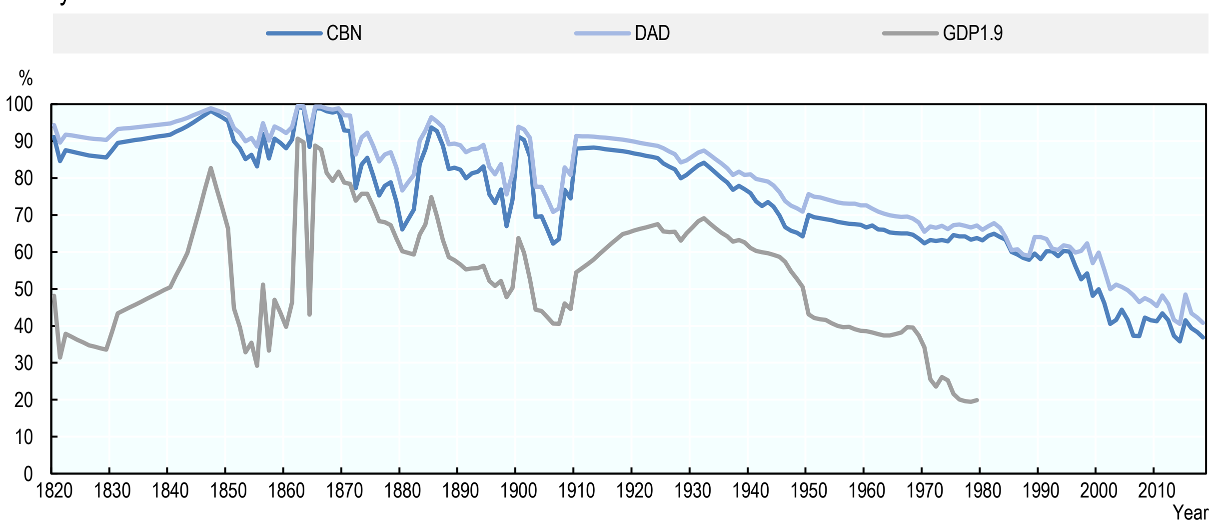
Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/avetyo

Figure 9.11 shows the evolution of extreme poverty for sub-Saharan Africa. This region featured the highest poverty rates throughout the entire period, with very few exceptions. In 1820, the regional poverty rate is 91%, which is still lower than the period's highest rate of 98% in 1847. Substantial poverty reduction starts only after 1917 and brings extreme poverty to a local minimum of 64% by 1949.[34] A slow improvement lasts until 1970, followed by slow increases until 1982. In the late 1980s, poverty rates fluctuate around 60% up until 1995, when a strong reduction is observed, bringing the rates to around 40%. By 2018 the poverty rate, at 37%, is close to its minimum reached in 2014.[35]

Figure 9.11. Extreme poverty in Sub-Saharan Africa based on different estimates

Poverty rate



Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

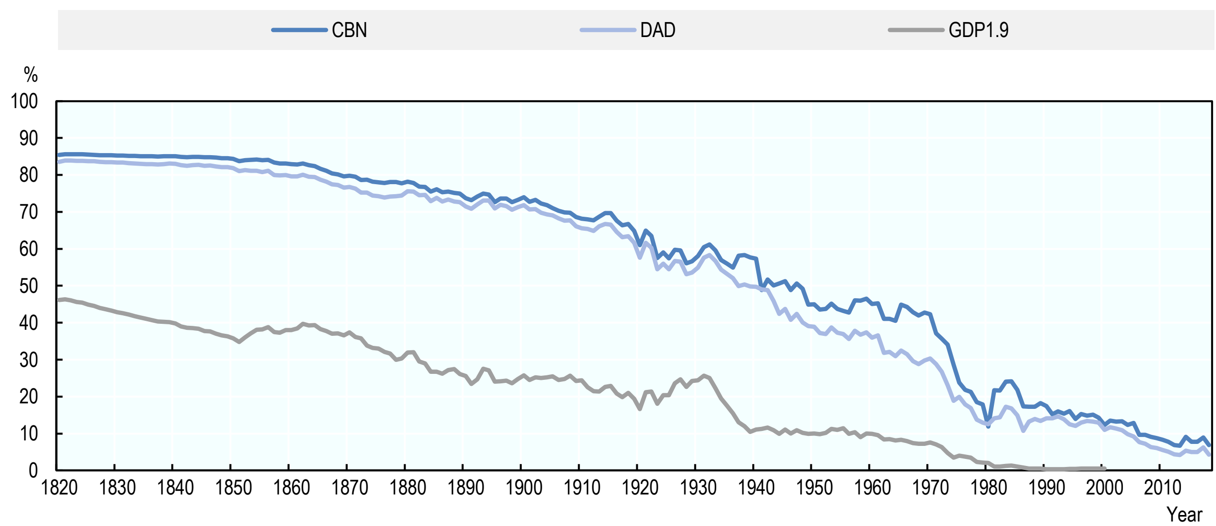
StatLink https://stat.link/5hpmg7

In terms of countries' contributions to the total count of extreme poverty, Ethiopia is the biggest contributor in 1820 with approximately 2.9 million, followed by Mozambique (1.9) and Madagascar (1.6).[36] By 1950, when data for many more countries are available, Nigeria contributes 57 million, followed by Ethiopia (22.3), Kenya (9.2) and South Africa (7.7). In 2018, Nigeria is again the largest contributor, with 121.3 million people in extreme poverty, followed by Ethiopia (40.9), Kenya (31.2), Madagascar (14.1) and South Africa (12.3).

Figure 9.12 shows the development of extreme poverty for Latin America and the Caribbean. Starting from a level of 85% in 1820, the poverty rate fell to 45% in 1950, and to 7% by 2018. A strong decline from 42% in 1970 to 12% in 1980 was partially reversed in the following year, with the poverty rate doubling to 24% by 1984. Since then, a slow reduction brought extreme poverty to a minimum by 2018. In terms of countries' contributions to the total population in extreme poverty, Mexico is the biggest contributor in 1820, with 5.7 million, followed by Brazil (4.1), Peru (1.2) and Colombia (1.1). By 1950, Brazil contributes 38.4 million, followed by Mexico (10.7), Colombia (4.8) and Peru (4.2). In 2018, Brazil is again the largest contributor with 15.6 million people in extreme poverty, followed by Haiti (7.9), Venezuela (6.7), Peru (3.4), Colombia (2.4) and Mexico (2).

Figure 9.12. Extreme poverty in Latin America and the Caribbean based on different estimates

Poverty rate



Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/wilgst

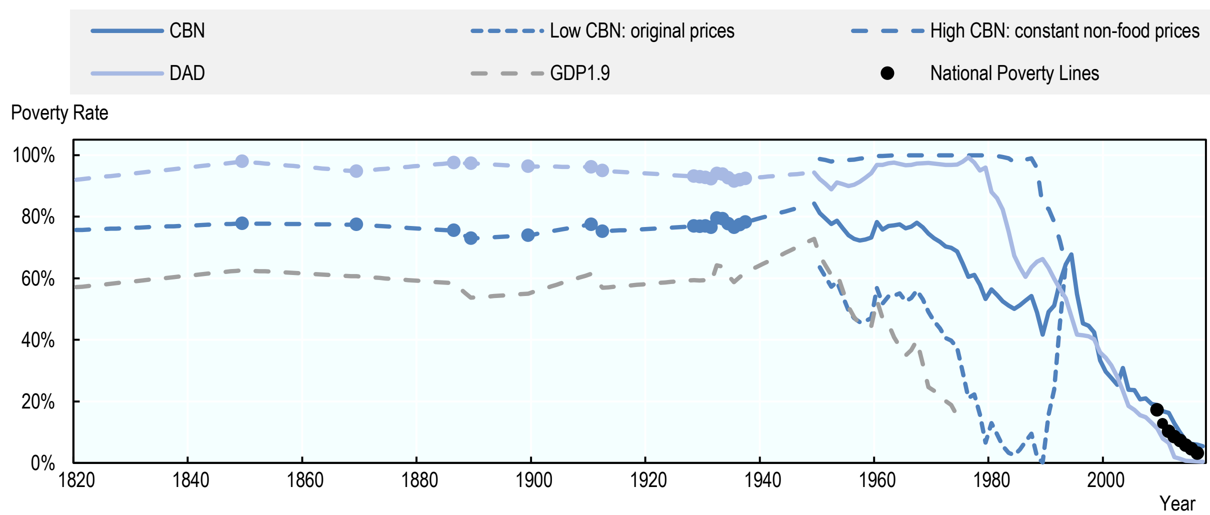
China

China requires special attention for two reasons: it was for a long period a non-market economy, and (because of its size) it has a disproportionate impact on global poverty counts. These two factors imply that uncertainty in Chinese estimates has large implications on a global scale. Figure 9.13 illustrates the magnitude of these concerns. It shows the trajectory of poverty rates based on the estimates discussed above, which result from averaging estimates based on two approaches: (a) taking price data at face value for the period 1990-95; and (b) attributing all CPI change to non-food items (explanation follows). Scenario (a) is the same approach used for all other estimates presented in this chapter, i.e. using nominal prices to estimate the CPF food poverty line, and then applying to them multipliers to obtain the Basic Diet and the non-Food poverty lines. The only difference is that the nominal prices used here are not those from the ILO data but from the Chinese Statistical Yearbooks.[37] This estimate (the lower dashed blue line at the bottom of Figure 9.13) shows an almost zero extreme poverty rate in 1990, which is unrealistic. The alternative approach is based on the idea that, since food prices have a much higher volatility in this period than the CPI, all CPI changes should reflect changes in the food component of the index, and that prices for the non-food items of the poverty basket over the period 1990-94 are at the level attained in 1995. This alternative estimate (shown by the dashed blue line at the top of Figure 9.13) is close to 100% in 1990.

Taking the average of two opposing scenarios that both seem unrealistic is far from resolving the problem, and the reader should consider the wide range of probable uncertainty as indicated in Figure 9.13, both for the 1990s and in the long run (implying possible values in the range of 60 to 95%).

Figure 9.13. Extreme poverty in China based on different estimates

Poverty rate



Note: CBN stands for the Cost of Basic Needs threshold; DAD is the standard (survey-based) USD 1.9/day poverty-line; GDP 1.9 is the USD 1.9/day poverty line adjusted using GDP per capita as the mean value of the distribution. Low CBN estimates based on nominal prices from Chinese Statistical Yearbooks; the high CBN estimates line relies on food prices and keeps non-food prices fixed at 1995 levels.
Source: Clio-Infra, www.clio-infra.eu.

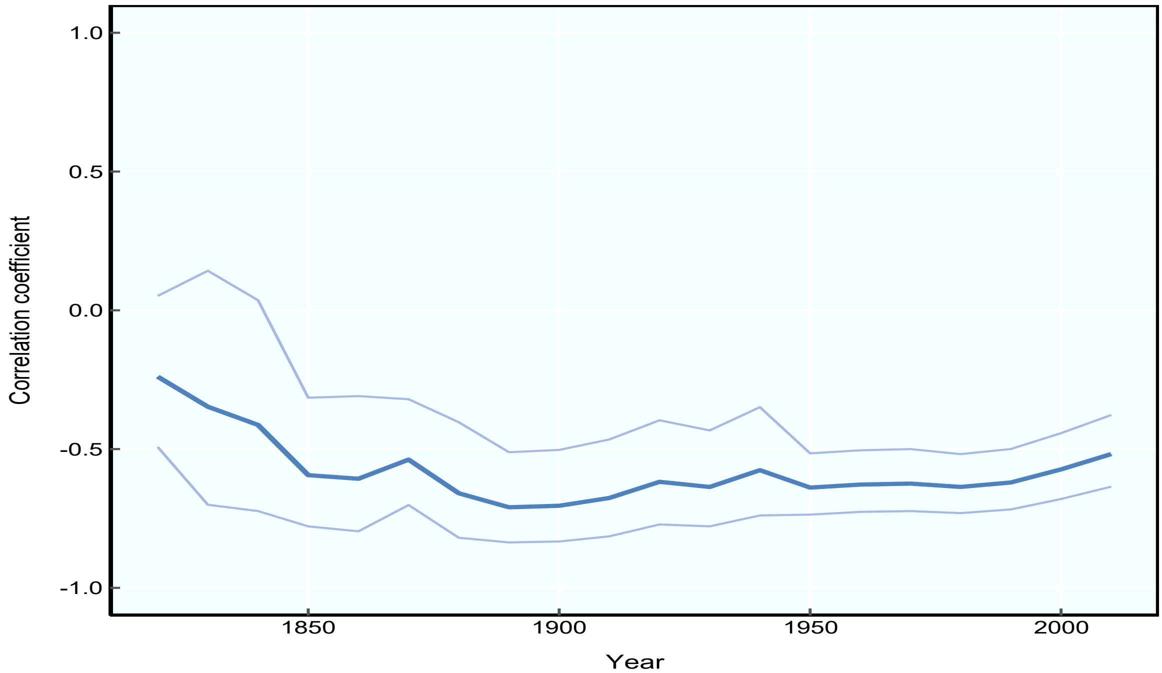
StatLink https://stat.link/v1hey0

Correlation with GDP per capita

Figure 9.14 shows the correlation in each year between the CBN-based poverty rates and the estimates of GDP per capita shown in Chapter 2 of this book. Across the full period, the correlation is -0.59, implying that countries with higher GDP per capita also feature lower rates of extreme poverty, with the largest value of about -0.77 achieved in 1919, and the smallest one in 1831 at -0.3. In the post-World War II period, the largest correlation of about -0.63 is achieved in 1978, and the smallest at -0.48 in 2014. The decline in this correlation may be explained by the increase of the within-country income inequality observed in the period between those years, albeit at different rates between countries. At the same time, the drop in global poverty after 1995 is the largest observed, despite the low correlation with GDP per capita. This implies a “lost opportunity”, for even faster poverty reduction could have been achieved if measures had been taken to contain increasing within-country income inequality. It should be noted though that the confidence intervals around these correlation coefficients are very high, implying that other factors beyond GDP per capita played an important role. The confidence interval around the correlation coefficient stretches into positive terrain in the early 19th century, implying that in the early 19th century some countries with high extreme poverty featured high GDP per capita, and vice versa.

Figure 9.14. Correlation between rates of extreme poverty and GDP per capita

Pearson correlation coefficient



Note: Thick and thin lines depict correlation coefficient and 95 % confidence intervals respectively.
Source: Clio-Infra, www.clio-infra.eu.

StatLink https://stat.link/ncvwd9

Priorities for future research

This first attempt to estimate global extreme poverty using a cost of basic needs approach over the long run highlights a number of novel empirical patterns but also the many areas where more historical research is needed. In particular, the global character of our exercise, and the still sparse availability of relevant data, has required taking shortcuts on a number of empirical issues. Most importantly, to account for a richer diet and for other non-food costs, our estimates rely on Allen's "multipliers" for 2011, instead of a direct calculation. Moving beyond this assumption will require the assembly of a broader global dataset on prices of all products and services relevant to people living in poverty. Such a dataset could then be used to provide direct estimates of poverty lines covering housing, education, health care and heating. Such a broader set of data on the prices of different consumption items would allow overcoming another limitation of the estimates presented in this chapter, i.e. the extensive use of CPI information.

An additional important consideration is the fixity across time of the welfare standard used here. The poverty line favoured by (Allen, 2017[12]) and used in this chapter is based on a requirement of 2 100 kcal of energy per person per day. Conversely, (FAO, 2001[31]; FAO, 2008[32]) provide measures of the so-called minimum dietary energy requirement as a function of the population's age and gender. Applying the approach used in this chapter to these minimal dietary energy requirements would allow providing estimates of extreme poverty that take into account changes in population structure over long periods of time. The same approach could be applied to take into account the Physical Activity Level of those living in poverty (i.e. the physical intensity of work, home, leisure and life conditions), which plays an important role in shaping the nutrient targets to be considered.

At the same time, secondary indicators are needed to address the problem of price unrepresentativeness in non-market economies, in particular for China and the former Soviet Union. Finally, a priority for future historical research should be to produce metrics of the sensitivity of the poverty rate with respect to the poverty line.

References

Aitchison, J. and J. Brown (1957), *The Lognormal Distribution*, The syndics of the Cambridge University Press, Cambridge. [25]

Alfani, G. and F. Tadei (2017), "Income inequality in colonial Africa: Building social tables for pre-Independence Central African Republic, Ivory Coast and Senegal", *African Economic History Working Paper Series*, No. 33, https://www.aehnetwork.org/wp-con. [27]

Allen, R. (2017), "Absolute poverty: When necessity displaces desire", *American Economic Review*, Vol. 107/12, pp. 3690-3721, https://doi.org/10.1257/aer.20161080. [12]

Allen, R. (2013), "Poverty lines in history, theory and current international practice", *Economics Series Working Papers*, No. 685, Department of Economics, University of Oxford, https://ora.ox.ac.uk/objects/uuid:3c9df348-d545-43dc-87f2-ae53c840f1d2. [19]

Atkinson, A. (2016), *Monitoring Global Poverty: Report of the Commission on Global Poverty*, The World Bank, Washington, DC., https://openknowledge.worldbank.org/bitstream/handle/10986/25141/9781464809613.pdf. [11]

Atkinson, A. et al. (2017), *The Chartbook of Economic Inequality*, http://www.chartbookofeconomicinequality.com. [34]

Bank of Greece, Bulgarian National Bank, National Bank of Romania, Oesterreichische Nationalbank (2014), *South-Eastern European Monetary and Economic Statistics from the Nineteenth Century to World*, Athens, Sofia, Bucharest, Vienna, http://file:///C:/Users/Faron_a/Downloads/SEEMHN%20Data%20Volume%202014%20(2).pdf. [41]

Bhalla, S. (2002), *Imagine There's No Country: Poverty, Inequality and Growth in the Era of Globalization*, Peterson Institute for International Economics. [2]

Bolt, J. and E. Hillbo (2016), "Long-term trends in economic inequality: Lessons from colonial Botswana, 1921-74", *The Economic History Review*, Vol. 69/4, pp. 1255-84, https://doi.org/10.1111/ehr.12326. [26]

Bolt, J. and J. van Zanden (2014), "The Maddison Project: Collaborative research on historical national accounts", *The Economic History Review*, Vol. 67/3, pp. 627-51, https://doi.org/10.1111/1468-0289.12032. [21]

Bourguignon, F. and C. Morrisson (2002), "Inequality Among World Citizens: 1820-1992", *The American Economic Review*, Vol. 92/4, pp. 727-44, https://www.jstor.org/stable/3083279. [3]

Chen, S. and M. Ravallion (2010), "The developing World is poorer than we thought, but no less successful in the fight against poverty", *The Quarterly Journal of Economics*, Vol. 125/4, pp. 1577-1625, http://qje.oxfordjournals.org.proxy.library.uu.nl/content. [5]

Clio Infra (2015), *Clio Infra Consumer Price Index dataset*, https://clio-infra.eu/Indicators/Inflation.html (accessed on 12 August 2015). [40]

de Zwart, P., B. van Leeuwen and J. van Leeuwen-Li (2014), “Real wages since 1820”, in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-8-en. [23]

Deaton, A. (2001), “Counting the World’s poor: Problems and possible solutions”, *The World Bank Research Observer*, Vol. 16/2, pp. 125-47, https://doi.org/10.1093/wbro/16.2.125. [8]

Eurostat (n.d.), *Eurostat Statistical Dashboard*, https://ec.europa.eu/eurostat/web/main/data/database. [51]

FAO (2008), *Updating the Minimum Dietary Energy Requirements*, Food and Agriculture Organization of the United Nations, Rome. [32]

FAO (2001), *Human Energy Requirements*, Food and Agriculture Organization of the United Nations, Rome. [31]

Ferreira, F. et al. (2015), “A global count of the extreme poor in 2012: Data issues, methodology and initial results”*, Policy Research Working Paper*, No. 7432, World Bank, Washington, http://hdl.handle.net/10986/22854. [6]

GINI (n.d.), *Growing Inequalities’ Impacts (GINI)*, http://www.gini-research.org/articles/home. [46]

Hirvonen, K. et al. (2019), “Affordability of the EAT – Lancet reference diet: A global analysis”, *The Lancet Global Health*, Vol. 8/1, pp. 1-8, https://doi.org/10.1016/S2214-109X(19)30447-4. [30]

Institute for Fiscal Studies (n.d.), *Living standards, poverty and inequality in the UK*, https://www.ifs.org.uk/tools_and_resources/incomes_in_uk. [45]

Jordà, Ò., M. Schularick and A. Taylor (2016), *Macrofinancial History and the New Business Cycle Facts*, University of Chicago Press. [33]

Kang, S. (2001), *Globalization and Income Inequality in Korea: An Overview*, OECD Technical Meeting. [29]

LIS (n.d.), *The Luxembourg Income Study Database (LIS)*, https://www.lisdatacenter.org/our-data/lis-database/. [44]

Moatsos, M. (2018), “The unbearable errorlessness of global poverty estimates”, *The Economists’ Voice*, Vol. 15/1, https://doi.org/10.1515/ev-2018-0004. [15]

Moatsos, M. (2017), “Global absolute poverty: Behind the veil of dollars”, *The Journal of Globalization and Development*, Vol. 7/2, https://doi.org/10.1515/jgd-2016-0033. [17]

Moatsos, M. (2017), *Tracing global poverty paths, 1925-2010*, Conference Proceedings of the 2017 Economic History Society Annual Conference. [35]

Moradi, A. and J. Baten (2005), “Inequality in Sub-Saharan Africa 1950-80: New estimates and new results”, *World Development*, Vol. 33/8, pp. 1233-65, http://www.eco.rug.nl/~jacobs/jjdownload/Workshop14May-Moradi Baten.pdf. [36]

OECD (n.d.), *OECD Income Distribution Database (IDD): Gini, poverty, income, Methods and Concepts*, https://www.oecd.org/social/income-distribution-database.htm (accessed on 11 February 2019). [42]

Prados de la Escosura, L. (2008), "Inequality, poverty and the Kuznets curve in Spain, 1850-2000", *The European Review of Economic History*, Vol. 12/3, pp. 287-324, https://doi.org/10.1017/S1361491608002311. [28]

R Core Team (2018), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, https://www.r-project.org/. [37]

Ravallion, M. (2010), "The debate on globalization, poverty, and inequality: Why measurement matters", in *Debates on the Measurement of Global Poverty*, Oxford Scholarship Online, https://doi.org/10.1093/acprof. [18]

Ravallion, M. et al. (1991), *Quantifying the Magnitude and a of Absolute Poverty in the Developing World in the Mid-1980s*, Policy, Research, and External Affairs, The World Bank, Washington. [13]

Reddy, S. and T. Pogge (2010), *"How not to count the poor"*, Oxford University Press. [10]

Sala-i-Martin, X. (2006), "The World distribution of income: Falling poverty and... convergence, period", *The Quarterly Journal of Economics*, Vol. 121/2, pp. 351-97, https://www.jstor.org/stable/25098796. [4]

Sandel, M. (2012), *What Money Can't Buy: The Moral Limits of Markets*, Allen Lane, London. [38]

SEDLAC (n.d.), *Socio-Economic Database for Latin America and the Caribbean*, https://www.cedlas.econo.unlp.edu.ar/wp/en/estadisticas/sedlac/. [50]

Solt, F. (2016), "The Standardized World Income Inequality Database", *Social Science Quarterly*, Vol. 97/5, pp. 1267-81. [39]

Srinivasan, T. (2009), "Irrelevance of the $1-a-day poverty line", *Initiative for Policy Dialogue Working Paper Series*, http://policydialogue.org/files/publications/Ch5_Srinivasan.pdf. [9]

Statistics Canada (n.d.), *Statistics Canada Data*, https://www150.statcan.gc.ca/n1/en/type/data?MM=1. [49]

Stone Center on Socio-Economic Inequality (2019), *All the Ginis (ALG) Dataset*, https://stonecenter.gc.cuny.edu/research/all-the-ginis-alg-dataset-version-february-2019/. [43]

Subramanian, S. (2015), "Once more unto the breach", *Economic & Political Weekly*, Vol. 45, pp. 35-40, http://ovidsp.ovid.com/ovidweb.cgi?T=JS{\&}PAGE=reference{\&}D=med5{\&}NEWS=N{\&}AN=16634276. [16]

Summers, R. and A. Heston (1988), "A new set of international comparisons of real product and price levels estimates for 130 countries, 1950-1985", *The Review of Income and Wealth*, Vol. 34/1, pp. 1-25, https://doi.org/10.1111/j.1475-4991.1988.tb00558.x. [14]

The World Bank (2020), *PovcalNet: An online analysis tool for global poverty monitoring*, http://iresearch.worldbank.org/PovcalNet/home.aspx (accessed on 19 January 2020). [24]

United Nations (2019), *World Population Prospects 2019*, UN Department of Economic and Social Affairs, https://population.un.org/wpp/ (accessed on 12 September 2019). [22]

UNU-WIDER (2018), *World Income Inequality Database - WIID*, United Nations University, https://www.wider.unu.edu/database/wiid. [47]

US CENSUS Bureau (n.d.), *US Census Data*, https://data.census.gov/cedsci/. [48]

USDA (2019), *FoodData Central*, US Department of Agriculture, https://fdc.nal.usda.gov/ (accessed on 19 September 2019). [20]

Van Zanden, J. et al. (2011), "The changing shape of global inequality 1820-2000: Exploring a new dataset", *Center for Global Economic History (CGEH) Working Paper Series*, No. 1, https://doi.org/10.1111/roiw.12014. [7]

van Zanden, J. et al. (eds.) (2014), *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-en. [1]

Notes

[1] Author's affiliation: Department of History and Art History, Economic and Social History, Utrecht University. I wish to thank Marco Mira d'Ercole, Bram van Besouw, Tim van der Valk, Gareth Austin, Alex Kolev, Wouter Ryckbosch, Mark Sanders, Jutta Bolt, Aditi Dixit, Auke Rijpma and Jan Luiten van Zanden for their suggestions, comments and remarks. I particularly thank Bas van Leeuwen for providing price data on China, as well as for sharing digitisalised ILO price data (together with Pim de Zwart); Robert Allen for discussing his method and results; Leandro Prados de la Escosura for all his valuable comments and for sharing his newly updated long-run inequality estimates for Spain along with his invaluable remarks and comments; Guido Alfani for the inequality data on sub-Saharan Africa; Christoph Lakner and Prem Sangraula from the World Bank for sharing the CPI dataset used by the Bank for global poverty measurement; and Guus Wieman for his excellent research assistantship. I want to extend my gratitude to the participants at the 2017 Posthumus Conference, the 2017 Economic History Society Annual Conference, and the workshops organised by the editors of this volume in Utrecht and at the OECD in Paris. All analysis has been conducted with R open source statistical computing software (R Core Team, 2018[37]). All remaining errors are my own.

[2] Throughout the chapter, poverty is qualified as "extreme" to distinguish it from the concept of "relative" income poverty used in most OECD publications. Relative income poverty relies on a threshold (based on a measure of household disposable income adjusted for household size) set at half of the median income of each country in each year.

[3] For details on the treatment of pre-1995 China, see the discussion later in this chapter, where two scenarios are considered on how to treat the problem posed by non-market prices. All estimates for pre-1995 China presented throughout this chapter are based on the average of the two scenarios.

[4] According to Allen's calculations, the CPF poverty line for developing countries is USD 1.84/day, on average, which is very close to the 1.9 dollar-a-day poverty line used by the World Bank and United Nations to monitor global extreme poverty. Allen comments that this line is too low, as "people eating a CPF diet suffer many nutritional deficiencies" (Allen, 2017[12]).

[5] Allen's estimates are only for 2011, and he uses the 2011 ICP price data.

[6] The data on prices from the October Inquiry of the ILO have been retrived from several ILO publications depending on the year: *International Labour Review* (1924-1935), *Yearbook of Labour Statistics* (1935-1964), *Bulletin of Labour Statistics* (1965-1991) and *Statistics on Occupational Wages and Hours of Work and on Food Prices* (1992-2009). All these publications are available at https://www.ilo.org/public/libdoc/ilo/P/.

[7] Using Table 11 from Allen (2017[12]). This is done in two steps: 1) one multiplier is used to bring the cost of the food component of the CPF poverty line to the level of the basic diet; and 2) another multiplier is used to account for non-food items, based on the share of the food component in the total consumption basket of the basic diet, see Table 12 in Allen (2017[12]). In our approach, both multipliers are a function of GDP per capita (based on the estimates by Bolt and van Zanden (2014[21]). In the calculations of those multipliers, France is excluded, as the price for wheat flour used by Allen (at less than 0.5 euro per kg in 2011) appears very low relative to the ILO price data used here (around 1.5 euro for 2011). See the online Appendix for additional information on these multipliers. This approach is not sensitive to the between-country differences in heating requirements. For a discussion of this issue see Moatsos (2017[17]), Moatsos (2017[35]) and Allen (2017[12]).

[8] The post-1984 price data were downloaded from the ILO's website (https://www.ilo.org/public/libdoc/ilo/P/), while the 1924-1984 data were scanned from various publications at the ILO document repository. The new version of ILOSTAT does not contain the price data any more, while the document repository with all available price data in pdf format can be found mainly at https://www.ilo.org/public/libdoc/ilo/P/09615/.

[9] Since price information in the ILO data are expressed using the currency denomination prevailing at the time (nominal prices), information on changes in currency denomination is required to denominate all prices in the currency used today. This was done using the dataset on the history of currencies provided by Global Financial Data, Global History of Currencies dataset; additional information on currency redenomination was used for the more recent period based on information available at the websites of national central banks, and collected on Wikipedia pages for local currencies.

[10] For more details, see the Appendix available at https://doi.org/10.33540/129.

[11] A variety of CPI sources were used, giving priority to the PovcalNet CPI data (http://iresearch.worldbank.org/PovcalNet/data.aspx), followed by other sources in order of preference: World Bank WDI data (https://datatopics.worldbank.org/world-development-indicators/), IMF CPI data (https://data.imf.org/?sk=4FFB52B2-3653-409A-B471-D47B46D904B5), Jordà, Schularick and Taylor (2016[33]), Sandel (2012[38]), Clio Infra Consumer Price Index dataset from the Clio Infra project website (Clio Infra, 2015[40]), Balkans CPI (historical CPI series from Balkan countries (Bank of Greece, Bulgarian National Bank, National Bank of Romania, Oesterreichische Nationalbank, 2014[41])) and the ILO general and food CPI index.

[12] The Gini index is a widely used measure of inequality with 0 representing total equality – everyone having the same income – and 1 representing total inequality – one person having all the income while everyone else has zero.

[13] Those sources are: (SEDLAC, n.d.[50]), STATCAN (Statistics Canada, n.d.[49]), US CENSUS data (US CENSUS Bureau, n.d.[48]), UNU-WIDER WIID (version 19 December 2018) (UNU-WIDER, 2018[47]), *OECD Income Inequality Database* (OECD, n.d.[42]), *All The Ginis Dataset* (Stone Center on Socio-Economic Inequality, 2019[43]), EUROSTAT (last update 31 January 2019) (Eurostat, n.d.[51]), GiniProject (version September 2013) (GINI, n.d.[46]), the *Luxemburg Income Studies Database* (LIS, n.d.[44]), UK

Institute for Fiscal Studies (Living Standards, Inequality and Poverty Spreadsheet) (Institute for Fiscal Studies, n.d.[45]), and Chartbook of Economic Inequality in the May 2017 edition (Atkinson et al., 2017[34]).

[14] As opposed to indirect estimates based on social tables or other methods (e.g. based on the relation between the coefficient of variation of height and income inequality (Moradi and Baten, 2005[36]; Van Zanden et al., 2011[7])). In addition to the scarce availability of data per se, the underlying methodologies and concepts used for carrying out a survey vary between countries and periods. Moreover, changes in the concepts and methodologies introduce structural breaks in the series for a given country.

[15] Solt (2016[39]) and the Estimated Household Income Inequality Data Set (EHII) V2017 v.1 from the University of Texas.

[16] The growth pattern implied by the evolution of the synthetic index is used to impute the Gini values between actual observations from other sources, with values that follow the same growth pattern as the synthetic index. When this is not possible, poverty rates are calculated using the available Gini indices at both ends; their weighted average is taken, with the estimates based on the closer Gini data point being linearly assigned a higher weight.

[17] In this calculation, I used poverty estimates only from countries that are available in both years, and that are not a result of regional imputation or interpolation.

[18] The World Bank takes a different route, assuming (regardless of the set of countries for which data are available) that countries without data have the same poverty rate as the rest of the region. This is a strong assumption that I relax here. The countries used for regional extrapolation in the 19th century going back to 1820 are: Japan for East Asia; South Africa (and Ghana up to 1870) for sub-Saharan Africa; Jordan, Lebanon, Egypt, Iran, Morocco, Tunisia, Syria and Turkey for the Middle East and North Africa; Denmark, Norway, Austria, Ireland, Italy, the Netherlands, France, the United Kingdom, Finland, Sweden and Portugal for Western Europe; the United States, Canada and Australia for the Western Offshoots; Argentina, Brazil, Chile, Colombia, Jamaica, Mexico, Peru, Uruguay and Venezuela for Latin America and the Caribbean; Indonesia, Sri Lanka, Myanmar, Malaysia, Nepal, the Philippines, Thailand and India (up to 1821) for South and Southeast Asia; and Poland for Eastern Europe and the former Soviet Union. More countries are used for imputations later in the 19th century and in the 20th and 21st centuries.

[19] These linearly interpolated estimates are not used for the regional imputation described previously. Moreover, going back in time and when no other information on inequality is available, the last and closest available Gini index is used to pin down an estimation for the country-level poverty rate.

[20] Once a cheap product is in shortage, a more expensive one should be used by the linear programming. Such information is lacking at the moment. In the case when the low-price products are indeed available and are purchased by those living in extreme poverty, those prices are used for calculating the cost of the basic needs poverty line.

[21] Thus, they include interpolations and extrapolations of prices using CPI or the evolution of prices of similar products within the ILO dataset.

[22] Both as a – generally increasing – function of real GDP per capita, see for more details the Appendix available at https://doi.org/10.33540/129.

[23] During the last few years in the period, both the Swiss and the US poverty lines flatten out at a constant PPP level. This is because the CPF cost is fixed in PPP terms, and the applied multipliers are and remain at their maximum.

[24] For these countries, the implications of using multipliers to estimate the non-food component and the richer basic diet (instead of using original data) become somehow evident. If the food component is very low, then the multiplier may underestimate the additional costs. Likewise, when the initial CPF food component is relatively high, then the multipliers most likely overestimate the additional costs.

[25] A low value of USD 2.16/day is observed in 1820, due generally to different coverage than most of the other years in the 19th century. A similar reduction is observed in 1870 for the same reason.

[26] In other words, the difference between GDP 1.9 and DAD is fully attributable to the effect of substituting GDP per capita with the estimated household mean income, and the difference between DAD and CBN lines is fully attributable to the effect of different poverty lines. More detailed estimates on a per country basis, and an exposition of some methodological details left out from the main text, can be found in the Appendix available at https://doi.org/10.33540/129.

[27] In both World Wars, a methodological problem presents itself: if and how should the death toll of the wars be counted in the poverty statistics, especially due to the acceleration in poverty reduction in the period after the peak. The observation of a positive long-run impact of war on extreme poverty may well reflect a misuse of statistical information or a demonstration of wrong accounting.

[28] As noted above, however, this period is particularly problematic due to the non-market nature of product prices; therefore caution is advised for the interpretation of these estimates.

[29] By 1900, Poland has relatively low levels of extreme poverty at around 3%. This low value is partly attributable to the very low poverty line of just above USD 1/day in that period, and partly to its GDP level and GDP/HHS mean ratio.

[30] Given the special subsection on China below, the focus of this subsection is mostly on other countries.

[31] Japan contributes only a few thousand this year, and Korea none.

[32] However, some relatively large countries are missing from this area: Libya, Iraq and Saudi Arabia. For the first two one would expect very high poverty rates.

[33] This estimate is of particular concern, as the data seem not to be representative of the dire situation for the Yemen population due to the ongoing war in its territory.

[34] Not many countries from this region have population data prior to 1950 in the sources used here (available pre-1950 population data cover: Angola, Ethiopia, Ghana, Madagascar, Mozambique, Mauritius, Somalia, Sudan and South Africa). Information from other sources could be added in later implementations of the approach used in this chapter to improve country coverage.

[35] For the sub-Saharan region in particular there is one additional methodological consideration in terms of comparability with most other regions. The *PovcalNet* distributions do not use equivalence scales when converting the “per household” income/consumption to a “per individual” basis. With this approach, larger families, who are disproportionately present in the region, are relatively penalised, with a much lower income per capita than when equivalence scales are used (as done by the OECD for example).

[36] Several countries in the region, including Nigeria, are not covered in this year by the population data.

[37] This is done as a robustness check, but the same results obtain with ILO data as well. The advantage in keeping the Yearbook statistics is that they reduce the reliance on CPI conversions, as they offer complete coverage of basic items, and cover up to, and including, 2014.

10 Biodiversity trends in a historical perspective

Thomas van Goethem and Jan Luiten van Zanden, Utrecht University

Biodiversity is important for human well-being as it provides ecosystems services such as the pollination of crops, the prevention of disease, and recreation. This chapter presents historical trends in biodiversity based on multiple indicators. Globally, the average abundance of species population has declined by 44% since 1970. Multiple indicators that cover a long-term timeframe show that biodiversity has declined for the better part of the Holocene, with this trend accelerating since 1900. In the shorter term, efforts to arrest the decline have borne some fruit, with a 36% increase in species abundance in Western Europe since 1970, which contrasts however with an 81% decline in Latin America and the Caribbean. This chapter also puts forward a framework for analysing key drivers of changes in biodiversity. An application of this analytical framework to the case of the Netherlands identifies population growth, intensification of agriculture, expansion of infrastructure and pollution as the key human drivers of biodiversity loss in the country since 1900.

Introduction

The decline of biodiversity – a term encompassing the variety and variability of life on earth – is one of the most urgent problems facing humanity. Biodiversity is a core feature of an ecosystem that produces human well-being via its positive effects on how ecosystems function. Ecosystems may provide provisioning services such as food and water; regulating services such as floods, drought, land degradation and prevention of disease; supporting services such as soil formation and nutrient cycling; and cultural services such as recreational, spiritual, religious and other non-material benefits. In recent history, economic development, often motivated by aspirations to improve human well-being, has come at the cost of biodiversity loss. Many biologists think that we are currently going through the sixth mass extinction in history, but this time with a very different cause: one species, *homo sapiens*, is largely responsible for it. Biodiversity loss is the result of complex interactions between humans and nature, and it is happening on an unprecedented scale (Díaz et al., 2019[1]; IUCN, 2008[2]; CBD, 2014[3]; Millennium Ecosystem Assessment, 2005[4]).

Global targets have been set, through the Convention on Biological Diversity (CDB), to significantly reduce the current pace of biodiversity loss (CBD, 2014[3]). Unfortunately, much remains unknown about the extent, rate and causes of the decline in global biodiversity. These knowledge gaps impede our ability to predict and mitigate its impacts. A historical perspective may help to identify events, processes and patterns of biodiversity decline. Processes directly and indirectly associated with declines in biological diversity have largely driven the development of civilisations since the Early- to Mid-Holocene about 5 000-7 000 years ago. These processes include the conversion of natural habitats to agriculture, the unsustainable exploitation of natural resources, the alteration of bio/geochemical cycles, the substitution of native and wild species by exotic and domesticated ones, the appropriation of primary production, and other human activities that generally lead to biodiversity loss (Naeem et al., 2016[5]). Additionally, a historical perspective on biodiversity may help to understand the impact of long-term processes. Short-term events such as cyclical movements of the economy and political changes may have small impacts on biodiversity, while long-term processes such as technological transitions or natural changes may have a more fundamental impact. Thus, better understanding the long-term development of biodiversity may help make better-informed decisions for protecting biodiversity. In a broader context, a long-term perspective may show the benefits of normally functioning ecosystems for human well-being in terms of the services they provide, while also revealing the boundaries of their carrying capacity (van Goethem and van Zanden, 2019[6]).

A multitude of indicators are used today to study (global) biodiversity loss, as biodiversity is a complex concept that is difficult to measure with a single metric (CBD, 2014[3]). Most of these indicators tend to be relatively data-intensive, often relying on extensive field monitoring programmes. This has resulted in two major limitations. First, because of the data availability issues, many indicators only have a species-specific or regional focus; only a few indicators have a country-specific focus. Second, again because of data limitations, these indicators generally go back only to 1970 or cover even shorter time periods. Historical indicators of changes in biodiversity, going back further in time, are scarce. Recently, however, more studies have been using interdisciplinary sources to compile longer and more complete records of developments in biodiversity. Ecologists can provide data on a micro-scale going back some 150 years, while historians can contribute data on a meso-scale up to 2 000 years ago, and archaeologists and paleo-biologists on a macro-scale, covering periods up to 10 000 years ago (Callicott, 2002[7]). Comparing measures from these partially overlapping temporal scales can help characterise historical species occurrence and biodiversity loss. This interdisciplinary approach, although relatively new, has produced some promising results, especially when combined with results from model studies (Kaplan, Krumhardt and Zimmermann, 2009[8]; Seebens et al., 2017[9]).

The set-up and content of this chapter is somewhat different from the other chapters in this book. Historical biodiversity research is a relatively new and heterogeneous field of research. In this chapter, the long-term trends in biodiversity are illustrated based on multiple indicators. Because of data limitations, no country-level data are presented, only data for larger world regions, the world as a whole, or specific ecosystems. Also, many often-used biodiversity indicators have a limited time frame; therefore we present several novel indicators relying on interdisciplinary data. Such indicators generally do not fit the "1820 to now" time frame of this book, as they tend to cover much longer time periods; also they do not refer to individual countries, as do other chapters in this book, but to the world as a whole, or specific ecosystems. Moreover, this chapter presents a case study to explore the relationship between GDP and biodiversity loss in the Netherlands and to identify potential drivers of (changes in) biodiversity. The case study relies on a novel framework that could be applied in future analysis.

The chapter has three sections. The first section presents biodiversity indicators based on the *Living Planet Index (LPI) methodology and data* for different world regions. These indicators, based on developments of vertebrate populations since 1970, show that globally the abundance of these species populations has declined by 36% between 1970 and 2010, with some regions faring much worse than others. The second section presents a suite of biodiversity indicators sourced from historical research to assess developments in global and regional biodiversity on longer timescales. These indicators, most referring to biodiversity trends well into the Holocene (up to 10 000 years BC), highlight that biodiversity has declined at least since 1500 AD, and probably for the better part of the Holocene. The last section presents a case study reconstructing biodiversity change in the Netherlands from 1900 onwards, showing a long-term decline in biodiversity between 1900 and 1970 and a partial recovery since the 1980s. By analysing specific assemblages of species, this section identifies population growth, the intensification of agriculture, the expansion of infrastructure, and pollution as the key human drivers of biodiversity loss in the Netherlands. The section also highlights an inverted U-curve link between GDP levels and biodiversity loss, which is in line with the environmental Kuznets curve hypothesis. However, the link is not an "automatic" one, suggesting the need for more explicit hypotheses about the mechanisms explaining the link between economic growth and the degradation of nature.

Description of the concepts used

Biodiversity is a complex concept, encompassing the variety and variability of all life on Earth, ranging from the genetic level (i.e. the diversity within a given species) to the species level (i.e. the number of species and the size of their population) and the ecosystem level (i.e. the diversity of ecosystems). This chapter presents evidence on the diversity of species and ecosystems as measured with several indices.

Species diversity usually refers to the number of species present in a certain territory – or in the world as a whole. It is, more or less, known how many vertebrate species have, for example, become extinct globally. Current estimates suggest that, since the year 1500, over 332 terrestrial vertebrates, 150 of them birds, have become extinct (IUCN, 2014[10]). Some authors have suggested that the current extinction rate of all species globally is over 1 000 times faster than the natural "background" rate of around 1-5 per year (Ceballos et al., 2015[11]). Species extinction is, however, only the tip of the biodiversity iceberg. Species abundance (the number of individuals belonging to each species per km^2) may be a more meaningful measure of biodiversity, as it can change dramatically over time even when no extinctions occur. Species abundance is therefore often used alongside measures of species richness.

The spatial scope of biodiversity indicators is also important (McGill et al., 2015[12]). Species may become extinct and biodiversity decline at the global level, whereas biodiversity may be stable or even rising at the local level, with certain species – perhaps those best adapted to human influences – becoming more widespread (Dornelas et al., 2014[13]; Dornelas et al., 2019[14]).

Measuring biodiversity is very difficult, especially over large spatial-temporal scales. Assessments therefore tend to look at particular components of biodiversity across terrestrial, marine and other aquatic ecosystems. This may involve species population, extinction risks, habitat extent and condition, and community composition. The suite of biodiversity indicators of the Convention on Biological Diversity (CBD) consider different components of global biodiversity (Table 10.1), divided into indicators of "state", "pressure" and "response". Among these categories, only the state indicators cover species diversity directly, albeit at different levels of detail, i.e. species populations, extinction risks and ecosystem level. The pressure indicators assess biodiversity decline only indirectly, by considering some of its key drivers, while the response indicators quantify the (policy) measures used to limit biodiversity decline. However, most indicators in the CBD framework have a relatively short time frame.

Table 10.1. Indicators of recent global biodiversity decline

	Indicator	Description	Data availability
State	Living Planet Index (LPI)	Mean population trends of vertebrates	1970-2018
	Wild Bird Index	Mean population trends of habitat specialists in EU and US	1980-2018
	Waterbird Population Status Index	Share of shorebird populations increasing, stable or decreasing	1985-2019
	Red List Index (RLI)	Extinction risk of mammals, birds, amphibians and corals	1986-2019
	Marine Trophic Index	Shifts from top predators to lower trophic levels	1990-2017
	Forest extent	Forest cover in million km^2	1990-2019
	Mangrove extent	Mangrove cover in million km^2	1980-2019
	Seagrass extent	Seagrass cover in million km^2	1930-2017
	Coral reef condition	Live hard coral cover	1980-2018
	Water Quality Index	Physical/chemical quality of freshwater	1980-2019
Pressures	Ecological footprint	Humanity's aggregate resource-consumption	1961-2019
	Nitrogen deposition rate	Annual reactive N deposited	1850-2019
	Number of alien species in Europe	Mediterranean marine, mammal and freshwater	1970-2019
	Exploitation of fish stocks	Share of stocks overexploited, fully exploited or depleted	1974-2019
	Climatic Impact Indicator	Climate change response of European bird populations	1980-2017
Responses	Extent of Protected Areas (PAs)	Cover of protected areas in million km^2	1888-2019
	Coverage bird PAs	Important Bird Areas and Alliance for Zero Extinction sites	1888-2019
	Area of FSC certified forest	Cover of FSC forest in million km^2	1995-2019
	International IAS policy adoption	Number signatories to conventions with provision for tackling IAS	1952-2019
	National IAS policy adoption	Share countries with relevant legislation	1964-2019
	Official development assistance	USD per year provided in support of CBD	2005-2019

Note: IAS stands for Invasive Alien Species. FSC is Forest Stewardship Council.
Source: Adapted from Butchart, S. et al. (2010[15]), "Global biodiversity: indicators of recent declines", *Science*, Vol. 328/5982, pp. 1164-1168.

The complexity of the concept of biodiversity, and the difficulties that arise when measuring it even today, imply that historical studies can only make use of proxies. Historical sources – in particular when they stretch far back in time – often relate to vertebrates, especially mammals and birds. Historical studies on biodiversity among plants often resort to early taxonomical works.[1] One way to overcome the measurement problem is to select "indicator species" that are believed to provide information on the overall status of the ecosystem and on the health of other species in that ecosystem. Such species include umbrella species, whose requirements for persistence include those of an array of associated species; keystone species, on which the health of the ecosystem depends; and foundation species that define much of the structure of a community by creating locally stable conditions for other species. Indicator species may also reflect the quality and changes in environmental conditions and in various aspects of community composition (Lindenmayer, Margules and Botkin, 2000[16]; Siddig et al., 2016[17]). By making a careful selection of indicator species representing the various biological characteristics of ecosystems in a region of a country, one can gain a better understanding of trends of biodiversity, in contexts where data on a broader range of species is lacking. Moreover, the impact of anthropogenic disturbances can be studied by selecting indicator species that are sensitive to environmental change. Indicator species are included in several contemporary indices of biodiversity such as the WWF *Living Planet Index*, which is based on trends in a limited range of species, most of them birds and mammals (Loh et al., 2005[18]; McRae, Price and Collen, 2012[19]). Applying such an approach in historical research is more complex, as the availability of historical sources is a limiting factor in selecting the appropriate indicator species. Several studies, however, have used indicator species in historical research.

A lively scientific debate has developed revolving around the question of whether economic growth can benefit the environment or instead typically leads to growing environmental problems. The central notion in this discussion is that of the environmental (or green) Kuznets curve, an analogy with Simon Kuznets's observation that income inequality tended to rise during the early stages of modern economic development and then to decline at later stages, due to a range of offsetting tendencies. This idea of an inverted U-curve linking GDP per capita and income inequality has since the early 1990s been applied in the field of environmental economics, suggesting a similar inverted U-curve between GDP per capita and environmental stress or loss of biodiversity. Some studies found that, during the first stages of industrialisation, pollution levels – for example, for certain emissions (SO2) – had a tendency to increase, whereas beyond a certain income threshold, decline set in – due to government policies, technological changes or changes in the structure of the economy (Grossman and Krueger, 1991[20]). The initial support for this hypothesis has, however, been replaced by a much more critical attitude, as it became clear that for many forms of pollution no "automatic" decline occurs beyond a certain income level (Dinda, 2004[21]; Stern, 2004[22]).

A similar discussion has emerged about the link between economic growth and biodiversity loss. Some authors have argued that a U-shaped relationship between GDP and biodiversity is theoretically impossible, because the extinction of species is an irreversible process and because ecosystems, once destroyed, cannot be reconstructed (Dietz and Adger, 2003[23]; Czech, 2008[24]). Others have found evidence that deforestation is negatively related to GDP levels (evidence that has been criticised by Mills and Waite (2009[25]), or that direct measures of biodiversity – for example, the share of threatened species – show a U-shaped link with GDP (Naidoo and Adamowicz, 2001[26]). McPherson and Nieswiadomy (2005[27]) documented the existence of a U-shaped link for mammals and birds, but not for plants, amphibians, reptiles and invertebrates, for which data are, however, much less satisfactory. Similarly, more recent research has produced mixed results, sometimes confirming a U-shaped relationship (De Santis, 2013[28]) and sometimes finding no evidence for it (see, for example, the comparative study of 50 US states by Tevie, Grimsrud and Berrens (2011[29])).

The limitations of this literature are, however, obvious. No satisfactory data on trends in biodiversity in the medium- and long-term are available to assess the nature of the relationship between biodiversity and GDP per capita. Data used by this research typically relate to, for example, the number of species classified as endangered (which may vary a lot from country to country), or the amount of land under conservation (De Santis, 2013[28]). Moreover, all data relate to recent years – basically the 1990s-2010s period – and lack historical depth for analysing the changing relationships between economic development and biodiversity.

Historical sources

This chapter presents evidence of biodiversity loss based on two main sources and approaches. A case study for the Netherlands is also included to introduce a framework for identifying and analysing key drivers of changes in biodiversity.

Biodiversity indices for world regions

In this section we present biodiversity indices, following the *Living Planet Index (LPI) methodology and data*, for the different world regions and on a global level. These indices differ in several respects from those included in Grooten and Almond (2018[30]). The main difference is that we derived biodiversity indices for different world regions, while the *LPI* has not published separate biodiversity trends for world regions, only for biogeographic realms (e.g. the Palearctic). The *Living Planet Index (LPI)* was developed to measure the state of the world's biodiversity from 1970 to the present time. It is one of the leading indicators used for assessing global developments in biodiversity. The index uses time-series data to calculate average rates of change in a large number of populations of terrestrial, freshwater and marine vertebrate species (Loh et al., 2005[18]; Collen et al., 2009[31]). The LPI reports how wildlife populations have changed in size – as opposed to the number of animal species that have been lost or gained. The index is not based on a census of all wildlife but rather on a variety of sources such as journals, online databases and government reports. The *Living Planet Database (LPD)* currently includes time-series data for over 20 000 populations of more than 4 200 mammal, bird, fish, reptile and amphibian species from around the world.

Population data in the *Living Planet Database* meet several requirements. They refer to single species monitored at a particular location over time. Different types of population abundance data are included in the database, such as full population counts, estimates (e.g. with population size estimated from measured parameters), densities, indices, proxies (e.g. breeding pairs, nests, tracks), measures per unit efforts (e.g. the number of fish caught per net per hour), biomass data (e.g. spawning stock biomass), sample data (e.g. where a proportion of the population is regularly monitored) and occupancy data (e.g. development in species occurrence in grid squares).[2] Abundance data feeding the *LPI* refer to a single species of vertebrate (mammals, birds, fish, reptiles and amphibians) over a period of at least two years, which do not need to be consecutive. When multiple measurements are taken over the course of a year, they were transformed into a single annual value.[3]

Here, we followed the generalised additive modelling framework to determine the trend in each population time series as described in Loh et al. (2005[18]) and Collen et al. (2009[31]). Average rates of change are then calculated and aggregated to the species level. No weighting system is applied when aggregating indices to the species level and then to the regional level. For the global measure, both weighted and unweighted aggregation methods have been applied. The weighted method gives a higher weight to trends in species-rich world regions (McRae, Price and Collen, 2012[19]). The total number of species per region was calculated from the *IUCN Red List of Threatened species database*, based on vertebrate species only, as invertebrates, plants and fungi are not sufficiently represented in the dataset. Moreover, on the global scale, disaggregated indices were computed for birds, mammals, reptiles and amphibians to assess whether trends differ among major taxonomic groups (Leung, Greenberg and Green, 2017[32]). Fish species were not included because of limited data availability.

Biodiversity state and pressure indicators based on historical research

This chapter also presents evidence on long-term trends in biodiversity based on historical research. In general, these indicators have used diverse (non-traditional) sources, including archaeological remains and palynological (i.e. plant pollen) data, but also archival records and oral histories, to reconstruct long-term trends in biodiversity: see van Goethem and van Zandem (2019a[33]) for a review of the sources. Most data are available via (disciplinary) global databases, for example the *Global Biodiversity Information Facility* (GBIF[34]), the *Living Planet Index* and the *Paleobiology Database* (PBDB[35]). These databases mainly cover survey data (>70 years), natural history collections and archaeological and paleo-ecological records, and provide information often stretching back centuries or millennia. This information typically refers to proxies of either states or drivers of biodiversity loss, for the world as a whole or for specific types of ecosystems.

Historical sources are in their infancy, and there are no global databases, only national databases, archives and repositories. An exception is the *HMAP Data Collection*, which contains marine catchment data from historical archives. Historical sources and data are often challenging to gather in major databases, as they may be reported in a variety of languages and dialects, contain disparate information (as they were created for different purposes), or be unavailable in a digital format (or without metadata). In addition to empirical data, several global models have been developed to reconstruct long-term trends in species occurrence and biodiversity. A common feature of all these models is that they are based on assumptions derived from contemporary research, i.e. the assumption that these relations also hold for the past.

Table 10.2. State and pressure biodiversity indicators from historical research

Indicator	Description	Data	Time scale	Spatial scale	Reference
Coral reef ecosystems / coral extent	Ecological condition of 7 ecological groups of reef inhabitants based on 14 regions in the tropical western Atlantic, Red Sea and northern Australia. Ecological conditions range from pristine to ecologically extinct	Paleontological, archaeological, historical, fisheries and ecological records	Pre-human to present	Global	Pandolfi et al. (2003[36])
Estuaries and coastal seas	Relative abundance of 6 taxonomic groups based on 12 temperate estuarine and coastal ecosystems in Europe, North America and Australia. Relative abundance is defined in classes ranging from pristine to extinct	Paleontological, archaeological, historical and ecological records	Pre-human to present	Global	Lotze et al. (2006[37])
Mammal distribution range	Historic and present-day distributional ranges of all the terrestrial mammals of Australia and subsets for Africa, Southeast Asia, Europe, North and South America. The subsets consist of species known to be shrinking	Historic ranges based on 19th century natural history works	1800 AD to present	Global, regional	Ceballos and Ehrlich (2002[38])
Avian extinction	Global extinctions of all bird species since 1500. Extinction data collected from various sources. Preliminary list reviewed and supplemented by regional and national experts	Birdlife International and references therein, regional and national field guides and family monographs. Older records based on historical sources	1500 AD to present	Global, regional	Szabo et al. (2012[39])
Non-native species recording	First year of record of established alien species	Online databases, scientific peer-reviewed publications, reports, books and personal collections. Older records based on historical sources	1500 AD to present	Global, regional	Seebens et al. (2017[9])
Forest extent	Deforestation modelled for prehistoric and preindustrial Europe. Models based on human population size constrained by suitability maps for agriculture and pasture. Relationship between population density and cleared area are used to determine deforested surface area	Population database and forest cover data (for population-deforestation relationship) based on historical records, statistical sources and estimations	1000 BC to 1850 AD	Europe	Kaplan, Krumhardt and Zimmermann (2009[8])
Ecological footprint	Human footprint based on expansion of cropland and pasture	Historical population, cropland and pasture statistics combined with satellite information and specific allocation algorithms (*HYDE 3.1*)	10000 BC to present	Global, regional	Goldewijk (2014[40])
Mean Species Abundance (terrestrial)	Biodiversity expressed as the remaining Mean Species Abundance relative to pristine conditions. Modelled based on historical land use change and causal relationships between land use change and biodiversity	Expansion of cropland, pasture and built-up area taken from the *HYDE 3.1 database*. Causal relationships derived from literature (meta-analyses)	1500 AD to present	Global, regional	Goldewijk (2014[40])

Case study: Biodiversity change in the Netherlands, 1900-2010

The main international reference for biodiversity indices is the *Living Planet Index (LPI).* This is based on a large number of estimates on the historical evolution of population sizes of different species in the (recent) past. The global and national biodiversity indices generally go back only to 1970. In this chapter, we present a case study where a similar approach is applied to the Netherlands going back to 1900. The basis for this national biodiversity index is the population trends of 58 individual mammal, bird and fish species collected from historical censuses, journal articles, reports and other historical sources. Moreover, by identifying – on the basis of the literature about these species – the habitats, threats and feeding types of the species concerned, it is possible to trace the main drivers of changes in the populations of these species. Changes in GDP per capita and the biodiversity index in the Netherlands (1900-2010) can also be analysed against each other.

This index is based on population data of individual mammal, bird and fish species. Population data refer to the abundance of a species for the Netherlands as a whole. Data were collected for each individual species from journal articles and reports. The dataset included not only census data but also proxies of species abundance such as fish landings and hunting records. No modelled population data were included. In total, 58 species were selected to be included in the index: 14 mammals, 14 fish and 30 birds. Of these 58 species, 55 were breeding in the Netherlands in 1900, of which four became extinct (Lesser horseshoe bat, Atlantic sturgeon, Allis shad and Twait shad); two are almost extinct (Atlantic salmon and Spotted ray); and 54 were breeding in the country in 2015, of which three resettled (Eurasian beaver, Grey seal and Wild boar) and two (Great crested grebe and Roe deer) vastly expanded since 1900.

Species representativeness was not an explicit criterion for the selection of species but was treated as a secondary criterion, because of limited data availability. However, these 58 species are spread more or less evenly over the respective species groups in terms of the number of species occurring and species listed in the Red List. Overall, 19% of the mammal species occurring in the Netherlands were selected, 13% for (breeding) birds and 11% for (freshwater) fish. Also, 57% of the selected mammals are Red List species compared to 35% for all mammals in the Netherlands (implying that our estimates for mammals may be biased downward as compared to the entire mammal group). For birds, 40% of our selected species are Red List species compared to 44% for all species. Furthermore, all major habitat types in the Netherlands (17) are represented by at least four typical species, and most habitat types (13) by eight or more species. A general picture of the number of species that have disappeared (locally extinct) and appeared (exotic species) in the 20th century provides a sense of how representative our dataset is. In total, 240 (breeding) bird, 71 mammal and 124 fish species are currently living in the Netherlands. Since 1900, nine bird, three mammal and eight fish species have disappeared from the Netherlands. On the other hand, 35 bird, eight mammal and 16 fish species have settled since 1900 as exotic species.

We collected population data for each individual species for every 5-year interval starting in 1900 until 2015. Some species (12 in total) had data for each year (or most years) in the time period considered. In this case, we computed a moving average of the 5 years around each 5-year interval. However, most species had gaps in the data, either missing data in between different 5-year intervals or missing data for the earliest time period (mostly between 1900 and 1950). In this case, data were inter- and extrapolated.[4]

For each species, a measure of relative abundance was calculated based on the maximum abundance found for that species. These biodiversity indices were calculated by taking the mean per 5-year interval of the series involved, with the index presented with a 1900 base year. The disaggregated indices (based on taxonomic class, feeding type, etc.) were calculated based on the same procedures as the national biodiversity index, but only including those species relevant for the respective categories. The details and procedures for selecting species, inter- and extrapolation and calculating the indices are explained in more detail in van Goethem and van Zanden (2019[6]).

Data quality

The data quality of the different indicators over different time periods is assessed in Table 10.3. Data quality is assessed only with respect to the performance of the chosen indicators. Other methodological issues related to the interpretation of the indicators, e.g. representativeness of indicators or the exact definition of biodiversity loss and how to quantify it, are discussed in the results section.

In general, historical data on species occurrence and abundance are often incomplete and sometimes contradictory. In most cases, the best available data are qualitative or semi-quantitative, and often originate from disparate sources. Several biases and limitations in the historical record need to be considered when studying biodiversity data (McClenachan et al., 2015[41]). Observations of species are generally biased toward those with economic or cultural value, while species that are common or have little economic or cultural value are less well recorded. The number of observations available for each species is also rarely consistent across time and space. Finally, certain original documents providing evidence on biodiversity may have been lost, or they may be fragmented or degraded. Early descriptions of the state of biodiversity should not be confused with direct observations, and a lack of written records does not imply historical absence. These biases and limitations hold true for both direct and indirect (i.e. based on proxies) indicators of biodiversity derived from historical records used in this chapter. These limitations hold true for all the indicators presented in the next section.

Table 10.3. Quality of data on historical biodiversity indicators by time period

A. Biodiversity indices for world regions							
	Pre-1500	1500	1800	1850	1900	1950	2000
World	-	-	-	-	-	2	1
Eastern. Europe	-	-	-	-	-	2	1
Middle East and North Africa	-	-	-	-	-	2	1
South and Southeast Asia	-	-	-	-	-	2	1
Western Offshoots	-	-	-	-	-	1	1
Latin America and Caribbean	-	-	-	-	-	2	1
Sub-Saharan Africa	-	-	-	-	-	2	1
East Asia	-	-	-	-	-	2	1
Western Europe	-	-	-	-	-	1	1
B. State and pressure indicators based on historical research							
	Pre-1500	1500	1800	1850	1900	1950	2000
Coral reef ecosystems	4	4	3	3	2	1	1
Estuaries and coastal seas	4	4	3	3	2	1	1
Mammal distribution range	-	-	4	3	3	2	1
Avian extinction	-	4	3	3	2	1	1
Non-native species recording	-	4	3	3	2	1	1
Forest cover	4	4	3	3	2	1	1
Coral reef ecosystems	4	4	4	3	2	2	1
Estuaries and coastal seas	4	4	4	3	2	2	1

Note:1. High quality: the product of official statistical agency (national or international); 2. Medium quality: the product of economic-historical research using the same sources and methods as applied by official statistical agencies; 3. Moderate quality: economic historical research, but making use of indirect data and estimates; and 4. Low quality: estimates based on a range of proxy information. In case of multiple sources, the lowest quality source is given.

There are also limitations in modelled indicators. In general, these models are based on (empirical) cause–effect relationships linking environmental drivers with biodiversity impact, for instance increased energy use, land-use change, forestry and climate change (Ten Brink, 2000[42]). Such an approach has two limitations. First, cause-effect relationships are based on contemporary research, assuming that these will also hold for the past, which they might not. Second, the approach is often hampered by the lack of historical data sources on environmental drivers, resulting in models that are often based on a single driver (when multiple drivers are at work). These limitations hold true for the modelled indicators described in the section on "Historical state and pressure indicators".

Main highlights on historical trends in biodiversity

Biodiversity indices for world regions

Figure 10.1 and Table 10.4 present biodiversity indices for the different world regions based on the *Living Planet Index methodology and dataset*. The global index shows that the abundance of a subset of 11 636 populations among 3 701 species has declined by 36% between 1970 and 2010. The indices further indicate that biodiversity loss for these species is much worse in Latin America and the Caribbean and in South and Southeast Asia, with declines of 81% and 75% respectively. Western Europe and Eastern Europe and the former Soviet Union regions, on the other hand, perform much better, with changes limited to 36% and 4%, respectively. Biodiversity loss in other regions is broadly in line with the global index. The disaggregated trends based on taxonomic groups show an average global decline of 21% for mammals, 29% for birds, 48% for reptiles and 44% for amphibians.

The estimates presented in Table 10.3 are average trends. This means that, in the case of the global index, some populations and species are declining by more than 36%, whereas others are not declining as much or are increasing. The average trend calculated for each species in the global biodiversity index shows that just over half of all reptile, bird and mammal species are stable or increasing. Conversely, the trends for over 50% of fish and amphibian species show a decline. As the number of species experiencing a positive or negative trend are, more or less, equal, the average decline for the global index implies that the magnitude of the losses exceeds that of the gains. This also suggests that the fall in the global index is not being driven by losses in just a few very threatened species, but that there are a large number of species in each group (almost 50%) that together produce an average decline.

Table 10.4. Summary of biodiversity indices for different world regions

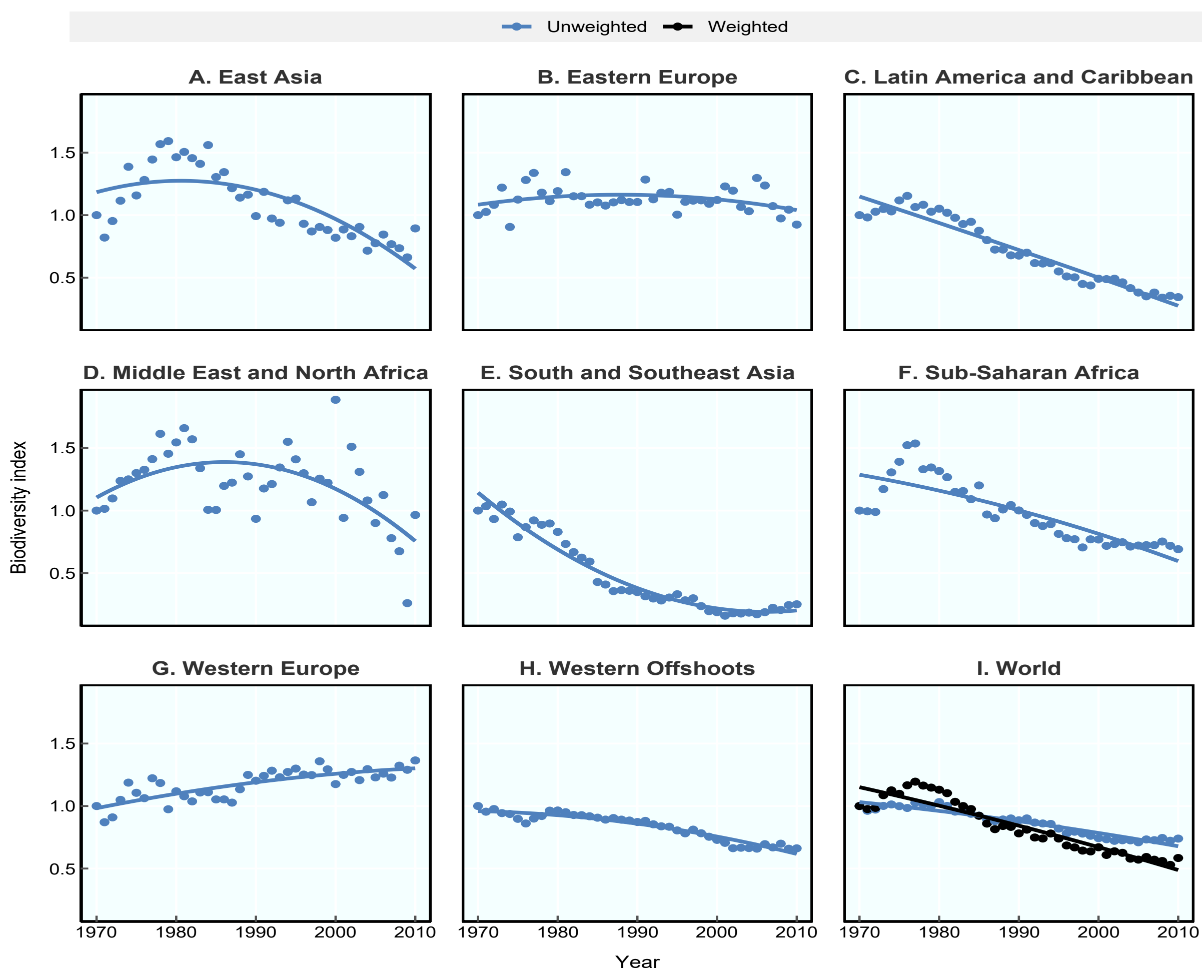
		Number of species, 2010	Number of records, 2010	Percentage change, 1970-2010
Global	Global (unweighted)	3 701	11 636	-36%
	Global (weighted)	3 701	11 636	-61%
	Birds (unweighted)	1 054	5 079	-29%
	Mammals (unweighted)	398	2 382	-21%
	Reptiles (unweighted)	130	581	-48%
	Amphibians (unweighted)	101	441	-44%
Across world regions	Eastern Europe	251	698	-4%
	Middle East and North Africa	116	255	-34%
	South and Southeast Asia	372	736	-75%
	Western Offshoots	1 645	3 806	-34%
	Latin America and Caribbean	892	1 522	-81%
	Sub-Saharan Africa	503	1 404	-62%
	East Asia	120	218	-55%
	Western Europe	618	2 644	+36%

Note: Percentage change computed based on the regression models.

StatLink https://stat.link/rihpk2

The 36% decline in biodiversity globally is much less than the 58% decline reported in the *Living Planet Index* report of 2018 (Grooten and Almond, 2018[30]) (Table 10.4). This difference is explained by the weighting system applied in the 2018 *LPI* report, which gives a greater weight to species-rich systems, realms and groups. A similar weighting system when applied to our dataset, based on the species richness across world regions, results in a decline of 61%, which is very close to that reported by the *LPI*. Moreover, the *LPI* 2012 report provided disaggregated trends based on countries' income level, ranging from an average increase of 7% for high-income countries to an average decline of 60% for low-income countries (McRae, Price and Collen, 2012[19]). While the *LPI* country-classification cannot be compared to the one used in this chapter, estimates are very similar in the case of more homogeneous regions, e.g. when comparing Western Europe and high-income countries.[5]

Figure 10.1. Biodiversity loss for different world regions based on the *Living Planet Index methodology and dataset*



Note: The blue line in the "World" panel is the population-weighted average.

StatLink https://stat.link/vag3jn

Biodiversity state and pressure indicators based on historical research

Historical research on global trends in the state of biodiversity suggest that biodiversity has declined at least since 1500, and perhaps for the better part of the Holocene (since 10 000 BC), at least in some world regions (Table 10.2 and Table 10.5). All state indicators show negative trends; three out of the four state indicators on the “pristine” to “late global” timescale show declines since the hunter-gatherer period; and both indicators on the 1500-2000 timescale have been declining since 1500. Additionally, both historical indicators of pressures on biodiversity point to an increase in pressure since the start of each respective period, i.e. since 10000 BC for the ecological footprint, and since 1500 AD for non-native species.

Historical indicators for coral reef ecosystems and for estuaries and coastal seas show similar declines (69% and 76% respectively). According to Pandolfi et al. (2003[36]), coral reef ecosystems experienced a decline in large animals before a decline in small animal species, and Atlantic reefs declined before reefs in the Red Sea and Australia, but the trajectories of decline were markedly similar worldwide. Moreover, all reefs were substantially degraded long before (modern) outbreaks of coral disease and bleaching.

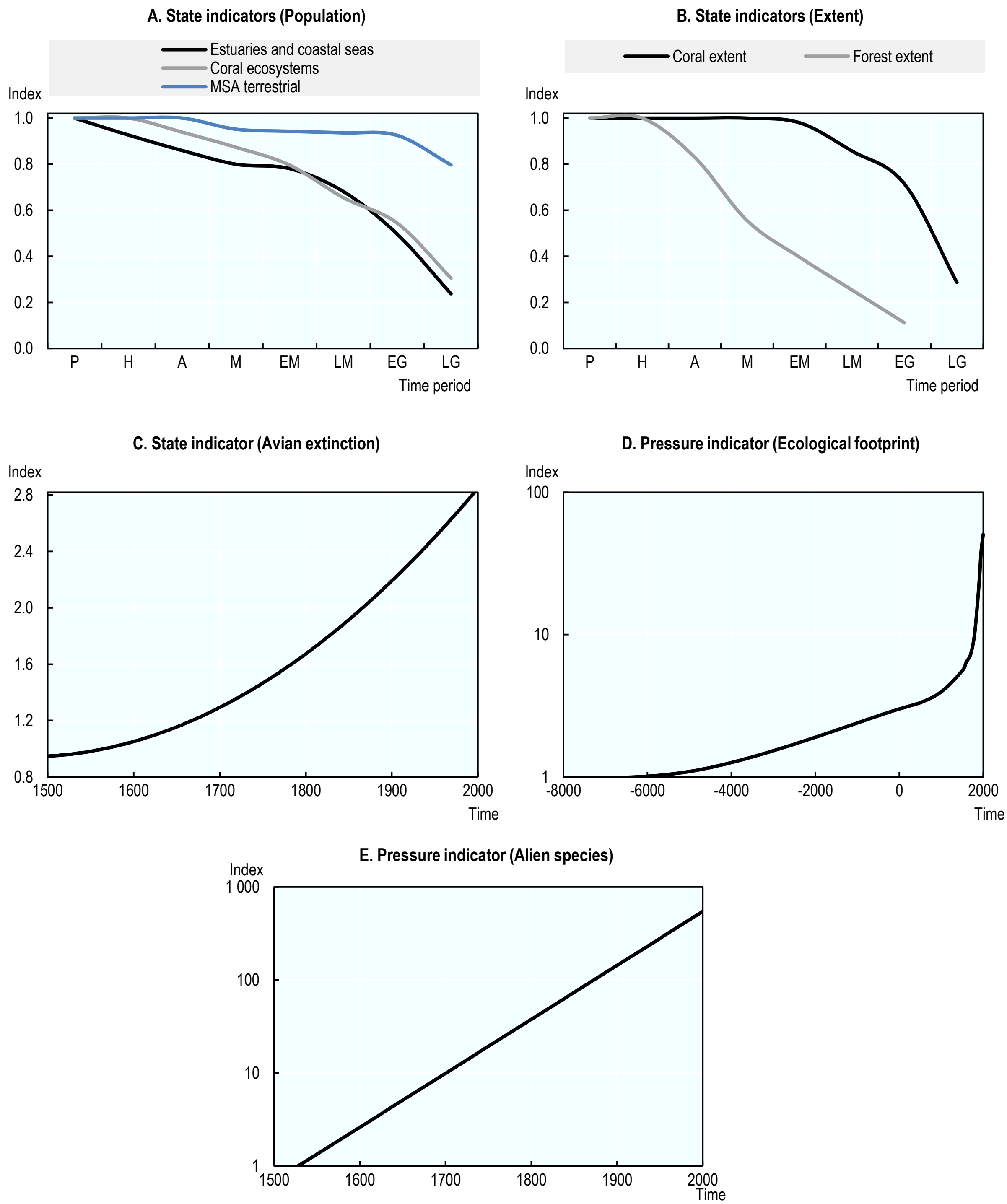
Historical research on the development of estuarine and coastal ecosystems points to similar patterns. (Lotze et al., 2006[37]) report that human impacts have depleted 90% of formerly important species in these ecosystems, destroyed 65% of seagrass and wetland habitat, degraded water quality and accelerated species invasions. All this started long before modern processes of industrialisation and globalisation.

Since 1500, 150 avian extinctions have been identified by historical research on a global scale, with the number of extinctions occurring every 25-year period increasing by 250% from 1500 to 2000. According to Szabo et al. (2012[39]) most losses (78% of species) occurred on oceanic islands, including the Hawaiian Islands, the Mascarene Islands (27 species), New Zealand (22 species) and French Polynesia (19 species), mainly driven by invasive alien species, hunting and agriculture. Ceballos and Ehrlich (2002[38]), reporting on the declining populations as a prelude to species extinction, showed that 173 mammal species across six continents lost over 68% of their historic range area. These declines mostly occurred where human activities are intensive.

The historical indicator on forest coverage in Europe showed an 89% decrease since 1000 BC. Kaplan, Krumhardt and Zimmermann (2009[8]) concluded that humans have transformed Europe’s landscapes since the first agricultural societies in the mid-Holocene. Major impacts were associated with the clearing of forests to establish cropland and pasture, and forest exploitation for fuel wood and construction materials. The human transformation of the landscape also resulted in a decrease of mean species abundance (MSA) of 20% since 1500 and a 5 100% increase of the ecological footprint since 1000 BC. Goldewijk (2014[40])showed that MSA has decreased in all regions of the world, with Europe experiencing a much earlier (and much larger) loss of biodiversity than other regions, mainly due to the loss of almost 90% of its forests. On a global scale, the decline in MSA has been only 20%, because many (larger) countries (e.g. Canada, the United States, Russia, Australia and Brazil) still have a relatively large forest area and high MSA. The historical ecological footprint indicator, on the other hand, increased remarkably in the Holocene, mainly reflecting the expansion of cropland and pasture.

Lastly, Seebens et al. (2017[9]) showed that the first records of non-native species worldwide increased during the last 200 years, with 37% of these increases reported since 1970. The increase can be largely attributed to the diaspora of European settlers in the 19th century, and to the acceleration in trade in the 20th century. For all taxonomic groups, the increase in the number of alien species does not show any sign of saturation, and most groups show increases in the rate of first records over time.

Figure 10.2. Trends in biodiversity states and pressures based on historical research



Note: The letters on the horizontal axis of panels A and C stand for: P=Pre-human, H=Hunter-gatherer societies, A=Agricultural societies, M=Medieval times, EM=Early-Modern times, LM=Late-Modern times, EG=Early-Global times, LG=Late-Global times.
Source: Data adapted from papers referenced in Table 10.3.

StatLink https://stat.link/524ybf

Table 10.5. Trends in global biodiversity based on historical research

	Indicator	Percentage change over the full-time range	Relative contribution to overall change, percentage points						
			Pre-human to Hunter-gatherer societies	Hunter-gatherer to Agricultural societies	Agriculture to Medieval times	Medieval to Early-Modern times	Early-Modern to Late-Modern times	Late-Modern to Early-Global times	Early-Global to Late-Global times
	Estuaries and coastal seas	-76	-10	-9	-10	0	-13	-24	-34
	Coral ecosystems	-69	0	-9	..	-21	-27	-9	-34
State of different ecosystems	Coral extent	-71	0	0	..	0	-20	-20	-60
	Forest extent	-89	0	-19	-31	-18	..	-32	..
			5000 BC to 0	0 to 1500	1500 to 1600	1600 to 1700	1700 to 1800	1800 to 1900	1900 to 2000
	Ecological footprint	+5 100	4	5	2	2	6	24	57
	Non-native species	+50 000	-	-	3	4	3	44	46
Pressures on biodiversity	Avian extinction	250	-	-	8	0	8	64	20
	Mean Species Abundance, terrestrial	-20	-	-24	-2	-2	-9	-24	-40
	Mammal range	-68*	-		-	-	-	..	..

Note: Values marked with an * mainly refer to the 19th century, those with a – to time periods outside the scope of this chapter. “..” means that the percentage change is not known.

StatLink https://stat.link/bo7k9m

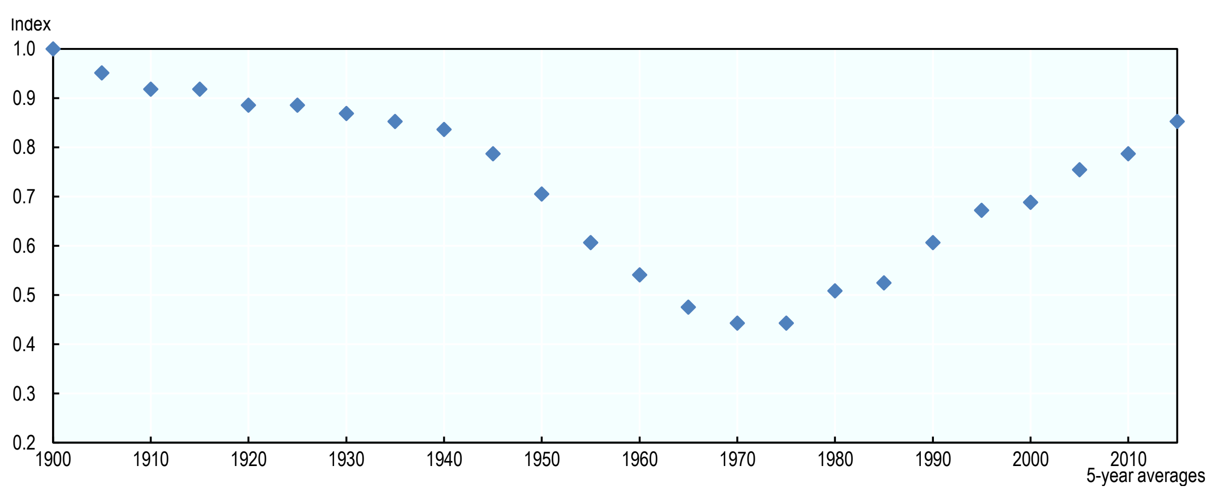
The development of historical indicators of biodiversity has progressed in recent years. However, there are still considerable gaps and variety in geographic, taxonomic and temporal coverage of the indicators presented in this section. Also, there are gaps for several key aspects of biodiversity states and pressures. Despite this, the available research on key dimensions of biodiversity suggests that, at the global scale, biodiversity has declined for the better part of the Holocene due to human impacts. Compared to Butchart et al. (2010[15]), which present an overview of 31 CBD indicators for the time period between 1970 and 2010, the directions of trends are consistent. The human impact on biodiversity seems, however, to have increased since 1900. Table 10.5 shows that from 30 to 60% of the total change reported by the historical indicators occurred in the last 100 years.

Case study: Biodiversity change in the Netherlands 1900-2010

Figure 10.3 presents a biodiversity index for the Netherlands for 1900-2015. The index shows a long-term decline in biodiversity between 1900 and 1970. The decline starts slowly with industrialisation but gains pace between the 1900s and the 1930s. Around World War II, the decline of biodiversity slows, only to speed up again after the war. The post-war recovery resulted in a strong decline in biodiversity, leading to a low point in the 1970s when biodiversity may have been less than half the level of 1900. Since the 1970s, various protective measures and nature conservation programmes (Cramer, 1989[43]; Leroy, 1994[44]) led to an increase in biodiversity in the mid-1970s that has continued up to the 2010s, while still not recovering to the same level as in 1900. The biodiversity index is based on only vertebrate species; the “state of invertebrate species” is described in Box 10.1.

Figure 10.3. Biodiversity in the Netherlands, 1900-2015

Index 1900 = 1



Note: 5-year averages. Based on abundance of 58 species (mammals, birds, fish). Unweighted index.
Source: van Goethem and van Zanden (2019[6]), “Economic development and biodiversity”, *CEPR Discussion Paper*, No. 13544, https://ssrn.com/abstract=3341351.

StatLink https://stat.link/0ncod7

Validating our estimates of biodiversity for the Netherlands is complicated, as no other indices describe biodiversity on such a long timescale. It is, however, possible to compare the contemporary part of the index to other indices of biodiversity. Here, the Netherlands national biodiversity index is compared to the *Living Planet Index* (*LPI*) for vertebrates in the Netherlands from 1990 to 2015 (van Strien et al., 2016[45]). Our national biodiversity index increases by 39% since 1990 compared to a 22% gain for the *LPI*. The indices can also be aggregated based on taxonomic class, i.e. birds, mammals, fish. Our biodiversity index shows a 70% increase for mammals as compared to a 123% increase for the *LPI* for vertebrates; a 29% increase for birds compared to 14% for the *LPI*; and 33% for fish compared to 14% for the *LPI*. The recovery in biodiversity in the more recent period is hence consistent with the *LPI* and with much of the recent research on the development of biodiversity in the Netherlands (van Strien et al., 2016[45]). Also, the data used to derive our biodiversity index are heterogeneous with regards to data type (e.g. including both census data and fish landings) and data quality (some species estimates are partly based on extrapolation). To assess the sensitivity of the results to these issues, subsets of the data were analysed. The sensitivity analysis showed that our biodiversity index (which refers only to vertebrate species) was not much affected by heterogeneity of the data, as the different subsets based on data type and quality yielded similar biodiversity curves both in shape and amplitude. More limited evidence on biodiversity trends for invertebrate species in the Netherlands also points to a continuous decline, see Box 10.1.

Box 10.1. Biodiversity loss for invertebrate species in the Netherlands

The biodiversity trends discussed above relate only to vertebrate species. Hence, they provide a partial picture of the “state of biodiversity”, as only the best-known groups are covered. What about invertebrate species groups? Many invertebrate species have disappeared from the Netherlands; these include species of water insects, bees, butterflies and beetles (Koomen, Van Nieukerken and Krikken, 1995[46]). Recently, several studies have reported on the development of insect populations in Western Europe. Hallmann et al. (2017[47]) found that insects in German nature reserves, particularly macro-moths, ground beetles and Caddisflies, appear to be in severe decline. These datasets suggest a reduction in biomass in macro-moths of approximately 61% and in ground beetles of at least 42%, over a period of 27 years. Van Strien et al. (2019[48]) found that butterflies declined by more than 80% in 1890-2017 on a national level, especially in grassland, woodland and heathland. The trend has stabilised over recent decades in grassland and woodland, but the decline continues in heathland. These studies, albeit limited to only some species groups, suggest that the declines in insects may be a widespread phenomenon.

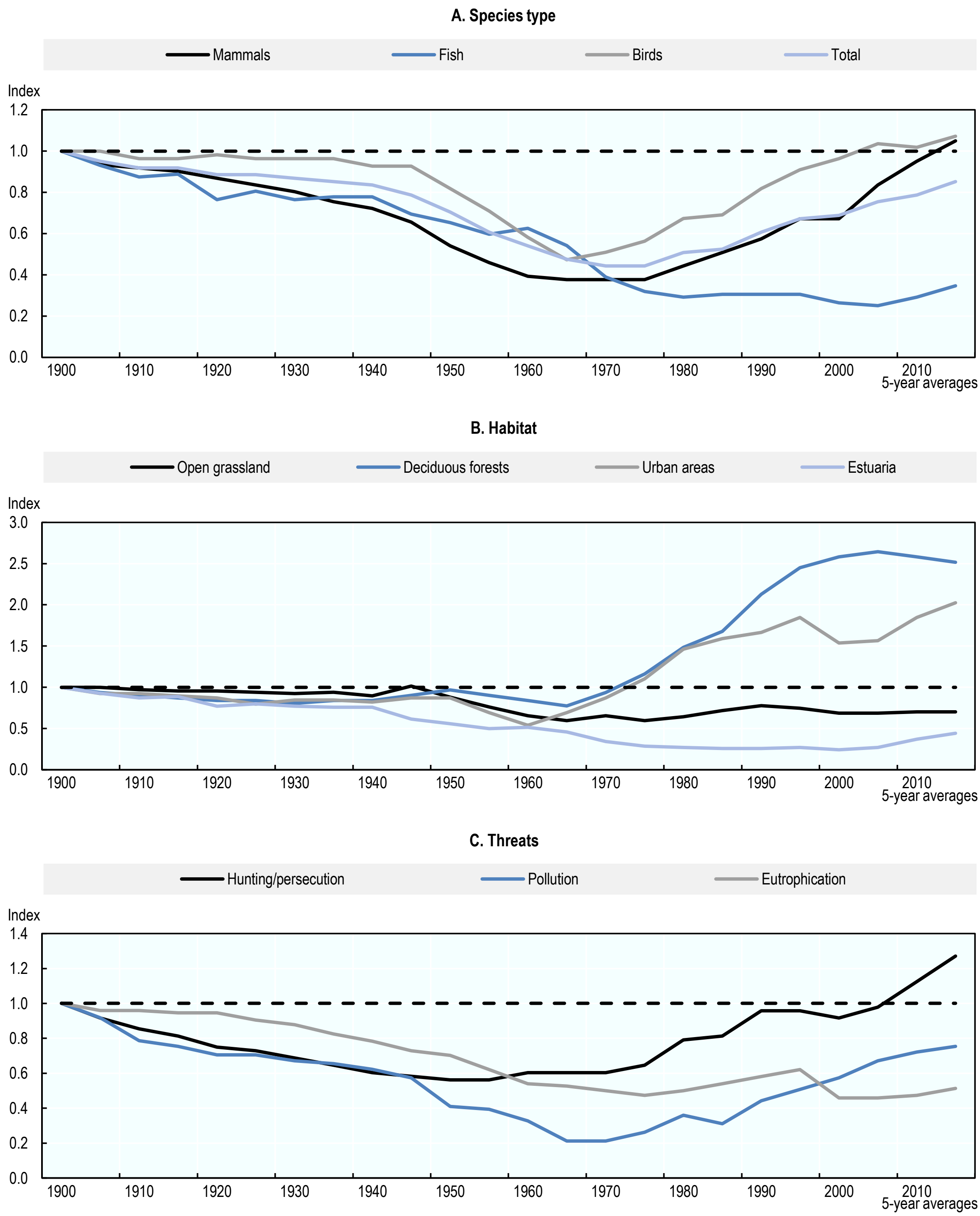
The observed decline and subsequent increase in biodiversity in the Netherlands can be interpreted by making use of “simple explanations”, such as industrialisation and the emergence of the environmental movement. While these explanations may be largely valid, a more nuanced story can be told by analysing changes in abundance for specific assemblages of species. Figure 10.4 shows aggregate indices based on species type (mammals, fish and birds), habitat preference (urban areas, deciduous forests, open grassland and estuaries) and species threats (pollution, hunting and eutrophication). The indices based on species type (Panel A) show a pattern similar to our national biodiversity index, at least for mammals and birds, and show slightly higher biodiversity in 2015 compared to 1900. Fish species, on the other hand, show a decline, and no recovery in recent years, a pattern that can be explained by the decline of migratory fishes in the second half of the 20th century as a result of damming and water-diversion systems.

Besides species type, the 58 selected species can also be grouped based on types of habitat (Panel B). The index for estuarine species shows a steady decline since 1900, with only a minor recovery since the 1970s low point. The biodiversity levels for these habitats in 2015 are only half those of 1900. Biodiversity in urban areas shows a strong increase since 1900, although this is based on only four species, while biodiversity in open grasslands shows no recovery since 1970. In part, this can be explained by developments in the landscape. Urban areas, for example, saw a six-fold increase compared to 1900. However, the surface area of a certain habitat alone might not explain everything, as habitat quality also needs to be considered, possibly in the case of open grasslands. Biodiversity in forest habitats increased sharply since the second half of the 1970s, partly reflecting a 30% increase in forest surface area.

A more experimental approach is to group species based on the reported causes of decline in the 20th century (Panel C). Qualitative statements on the causes of a decline in abundance have been collected for each species (if applicable) and aggregated into several categories of “species threats”. Species that have endured hunting and persecution show a decline since 1900, with a low point in the 1950s; the level achieved in 2015 is almost 40% higher compared to 1900. In part, this large increase can be explained by the fact that several species (e.g. beaver and grey seal) were extinct in 1900 due to hunting and persecution, but resettled successfully in the 20th century. Moreover, biodiversity levels in 1900 may have been low in general for the species included because of hunting and persecution in the 19th century and even before that time. The eutrophication index shows a steady decrease, with a steady state reached in the 1970s well below the 1900 level. This suggests that eutrophication is still a problem for these species. Here it has to be noted that the index is based on only six species. The pollution index hits a low point around the 1960s and 1970s, with increases ever since, reflecting the banning of many toxic substances since the 1960s.

Figure 10.4. Biodiversity indices by species type, habitat preference and species threats in the Netherlands, 1900-2015

Index, 1900 = 1



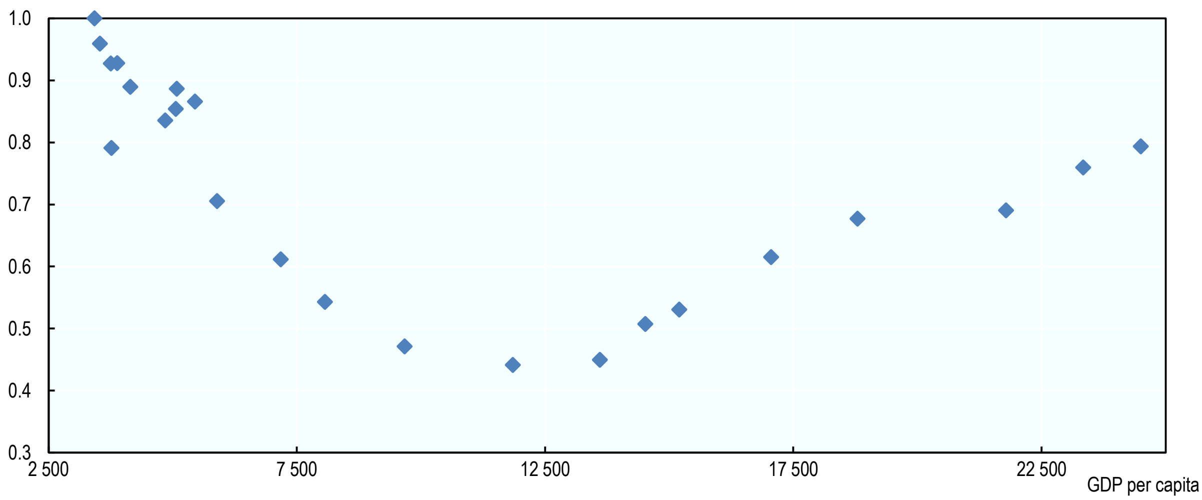
Note: The y-axis runs from 0.2 to 2.6 at 0.2 intervals, the dotted line =1.
Source: van Goethem and van Zanden (2019[6]), "Economic development and biodiversity", *CEPR Discussion Paper*, No. 13544, https://ssrn.com/abstract=3341351.

StatLink https://stat.link/mtg67f

Correlation with GDP per capita

Figure 10.5 plots the development of GDP per capita and the biodiversity index for the Netherlands from 1900 to 2010 (Bolt and van Zanden, 2014[49]). The Pearson correlation coefficients for four time periods are -0.81 (1900-1925), -0.22 (1925-1950), -0.88 (1950-1975) and +0.92 (1975-2010). The negative correlation with GDP per capita from 1900 to 1975 suggests that industrialisation in this period had a strong negative effect on biodiversity. From 1975 onwards, however, the relationship between biodiversity and GDP per capita becomes positive, a pattern that is broadly consistent with the notion of an inverted U-shaped environmental Kuznets curve.

Figure 10.5. Average species abundance and GDP per capita in the Netherlands, 1900-2010



Note: 5-year averages.
Source: van Goethem and van Zanden (2019[6]), "Economic development and biodiversity", *CEPR Discussion Paper*, No. 13544, https://ssrn.com/abstract=3341351.

StatLink https://stat.link/58z1qr

The working hypothesis underpinning Figure 10.5 is that there is a link between a country's income level and the protection of nature, but that this link is not an "automatic" one (as assumed in the "naïve" version of the environmental Kuznets curve), but the result of the growth of "green civil society" during economic development, as nature becomes more scarce and the awareness of the need to protect it increases with income growth. The environmental Kuznets curve is therefore a good starting point for this type of research, but the idea has to be supplemented with more explicit hypotheses about the mechanisms that explain the supposed U-turn in the link between GDP and the degradation of nature. There is no reason to assume that economic growth as such leads to such a change.[6]

Priorities for future research

A debated issue in biodiversity research is the representativeness of the indices developed, in terms of both indicator type (e.g. richness, diversity) and species selection. In this chapter, the most prominent question is, how representative is the species set on which the biodiversity indices are based. Three main issues can be identified: a skewed distribution of records across regions, across species groups and within-species groups. In general, most datasets are not uniformly distributed across regions and species. For example, birds and mammals are best represented in the *LPI dataset*, particularly in the Nearctic and Palearctic region, whilst data for other groups are scarce.

One way to deal with the skewed spatial distribution is a weighting approach, which has been developed for the *LPI* to make the indicator more representative of vertebrate biodiversity. The index is calculated based on a system of weighting that reflects the actual number and distribution of vertebrate species in the world, rather than the number of species present in the dataset. The under-representation of entire species groups, such as insects, is however more difficult to tackle. The IUCN recommended expanding the number of Red List assessments, especially for other species groups, as an important future goal. Besides adding more data, the problem of under-representation can also be tackled by carefully communicating the results, and especially the limitations, of biodiversity research.

The last issue with the representativeness of the data concerns the skewed distribution of records within species groups. The taxonomic catalogues of plants, terrestrial vertebrates, freshwater fish and some marine species are assumed to be sufficient to assess their status and the limitations of our knowledge (Pimm et al., 2014[50]). However, indices based on population trends (particularly at the regional scale) generally include few species that are rare, localised or difficult to survey, including those most susceptible to extinction (Butchart et al., 2004[51]). Also, most species are undescribed. The species we know best have large geographical ranges and are often common within them. Recently, however, Pimm et al. (2014[50]) showed that most newly described species have small ranges and are typically geographically concentrated and disproportionately likely to be threatened or already extinct (Pimm et al., 2014[50]). Therefore, continued attention to representation within species groups is warranted going forwards.

Qualitative and semi-quantitative historical data are essential for deriving longer trends in biodiversity (McClenachan, Ferretti and Baum, 2012[52]). However, a framework for systematically incorporating qualitative information in biodiversity indices is not yet available. This is partly caused by the disparate nature of the data (sources), making it difficult to develop a uniform methodology for matching, or mapping, qualitative information to quantitative data. In our case study for the Netherlands, qualitative data were used conservatively and on a case-by-case basis. In some cases, trends were extrapolated with the help of qualitative information. This approach also has disadvantages, for example, when dealing with historically exploited species that are experiencing recent population increases. (Branch, Matsuoka and Miyashita, 2004[53]) showed, based on data from early 20th century whaling logbooks, that although the population of Antarctic blue whales increased by around 7% per year from 1974 to 2004, it was still below 1% of pre-exploitation abundance levels. This meant that this species should remain protected. Recent population increases are promising, but without data accurately describing the period of exploitation, the magnitude of the recovery can easily be overestimated (Lotze, Coll and Dunne, 2011[54]).

The approach presented in the case study for the Netherlands can be applied to other countries. This would provide an opportunity for a comparative analysis between countries and help to identify relevant drivers of biodiversity across spatial regions and species groups. Also, this would help expand the scope, and analytical depth, of the framework by including quantitative information on drivers of biodiversity, connecting these to the disaggregated biodiversity curves. Moreover, this would provide input to further develop the Kuznets curve hypothesis as a framework, not as the primary driving mechanism, for studying the interaction between biodiversity and socio-economic developments. On a longer timescale, it would be interesting to relate the long-term biodiversity indicators to major socio-economic transitions in human development. Biodiversity research is generally concerned with the present or the recent past. Only short-term processes can be studied in these time frames, while long-term processes are considered, or observed as, boundary conditions (De Vriend, 1991[55]). Long-term biodiversity records may reveal the natural development of biodiversity over ecological timescales, including cycles of succession. The amplitude and frequency of change in these long-term records, or the lack thereof, can be a measure of the impact and scale of disturbances. For example, short-term processes or events can have long-lasting impacts manifested by state shifts of the ecosystem (Scheffer et al., 2001[56]). Perhaps even more interesting is analysing the direction of trends, which might be indicative of socio-economic transitions. These transitions can be characterised by a considerable increase in our species' impact on the biosphere (Takács-Sánta, 2004[57]). Six such transitions are identified; in chronological order, these are: 1) the use of fire, 2) language, 3) agriculture, 4) civilisation (states), 5) European conquests and 6) the technological-scientific (r)evolution and the dominance of fossil fuels as primary energy sources. Longer time frames may provide an opportunity to study the full impact of human development on biodiversity, as major socio-economic transitions and natural changes come into focus only on these scales.

References

Bolt, J. and J. van Zanden (2014), "The Maddison project: Collaborative research on historical national accounts", *The Economic History Review*, Vol. 67/3, pp. 627-651. [49]

Branch, T., K. Matsuoka and T. Miyashita (2004), "Evidence for increases in Antarctic blue whales based on Bayesian modelling", *Marine Mammal Science*, Vol. 20, pp. 726-754. [53]

Butchart, S. et al. (2004), "Measuring global trends in the status of biodiversity: Red list indices for birds", *PLoS biology*, Vol. 2/12. [51]

Butchart, S. et al. (2010), "Global biodiversity: indicators of recent declines", *Science*, Vol. 328/5982, pp. 1164-1168. [15]

Callicott, J. (2002), "Choosing appropriate temporal and spatial scales for ecological restoration", *Journal of Biosciences*, Vol. 27/2, pp. 409-420. [7]

CBD (2014), *Global Biodiversity*, Secretariat of the Convention on Biological Diversity, Montréal, https://www.cbd.int/gbo/gbo4/publication/gbo4-en-hr.pdf. [3]

Ceballos, G. and P. Ehrlich (2002), "Mammal population losses and the extinction crisis", *Science*, Vol. 296/5569, pp. 904-907. [38]

Ceballos, G. et al. (2015), "Accelerated modern human-induced species losses: Entering the sixth mass extinction", *Science advances*, Vol. 1/5, e1400253. [11]

Collen, B. et al. (2009), "Monitoring change in vertebrate abundance: the Living Planet Index", *Conservation Biology*, Vol. 23/2, pp. 317-327. [31]

Cramer, J. (1989), *"De groene golf - geschiedenis en toekomst van de Nederlandse milieubeweging"*, Utrecht. [43]

Czech, B. (2008), "Prospects for reconciling the conflict between economic growth and biodiversity conservation with technological progress", *Conservation Biology*, Vol. 22/6, pp. 1389-1398. [24]

De Santis, R. (2013), "Is there a "Biodiversity Kuznets Curve" for the main OECD countries?", *Rivista di politica economica*, Vol. 3, pp. 215-225. [28]

De Vriend, H. (1991), "Mathematical modelling and large-scale coastal behaviour: Part 1: physical processes", *Journal of hydraulic research*, Vol. 29/6, pp. 727-740. [55]

Díaz, S. et al. (2019), *"Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services"*, IPBES. [1]

Dietz, S. and W. Adger (2003), "Economic growth, biodiversity loss and conservation effort", *Journal of Environmental Management*, Vol. 68/1, pp. 23-35. [23]

Dinda, S. (2004), “Environmental Kuznets curve hypothesis: a survey”, *Ecological Economics*, Vol. 49/4, pp. 431-455. [21]

Dornelas, M. et al. (2014), “Assemblage time series reveal biodiversity change but not systematic loss”, *Science*, Vol. 344/6181, pp. 296-299. [13]

Dornelas, M. et al. (2019), “A balance of winners and losers in the Anthropocene”, *Ecology Letters*, Vol. 22/5, pp. 847-854. [14]

GBIF (n.d.), *Free and open access to biodiversity data*, Global Biodiversity Information Facility, https://www.gbif.org/. [34]

Glasbergen, P. (ed.) (1994), *“De ontwikkeling van het milieubeleid en de milieubeleidstheorie”*, VUGA Uitgeverij, Den Haag. [44]

Goldewijk, K. (2014), “Environmental quality since 1820”, in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-14-en. [40]

Grooten, M. and R. Almond (2018), *Living Planet Report - 2018: Aiming Higher*, WWF, Gland, Switzerland. [30]

Grossman, G. and A. Krueger (1991), *Environmental impacts of a North American free trade agreement*, National Bureau of Economic Research. [20]

Hallmann, C. et al. (2017), “More than 75 percent decline over 27 years in total flying insect biomass in protected areas”, *PloS one*, Vol. 12/10, e0185809. [47]

IUCN (2014), *IUCN Red List of Threatened Species. Version 2014.2*, http://www.iucnredlist.org. [10]

IUCN (2008), *IUCN Red List of Threatened Species. Version 2008*, http://www.iucnredlist.org. [2]

Kaplan, J., K. Krumhardt and N. Zimmermann (2009), “The prehistoric and preindustrial deforestation of Europe”, *Quaternary Science Reviews*, Vol. 28/27-28, pp. 3016-3034. [8]

Koomen, P., E. Van Nieukerken and J. Krikken (1995), *Biodiversiteit in Nederland*, Nationaal Natuurhistorisch Museum. [46]

Leung, B., D. Greenberg and D. Green (2017), “Trends in mean growth and stability in temperate vertebrate populations”, *Diversity and Distributions*, Vol. 23/12, pp. 1372-1380. [32]

Lindenmayer, D., C. Margules and D. Botkin (2000), “Indicators of biodiversity for ecologically sustainable forest management”, *Conservation Biology*, Vol. 14, pp. 941-950. [16]

Loh, J. et al. (2005), “The Living Planet Index: Using species population time series to track trends in biodiversity”, *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 360/1454, pp. 289-295. [18]

Lotze, H., M. Coll and J. Dunne (2011), “Historical changes in marine resources, food-web structure and ecosystem functioning in the Adriatic Sea, Mediterranean”, *Ecosystems*, Vol. 14/2, pp. 198-222. [54]

Lotze, H. et al. (2006), “Depletion, degradation, and recovery potential of estuaries and coastal seas”, *Science*, Vol. 312/5781, pp. 1806-1809. [37]

McClenachan, L. et al. (2015), "The importance of surprising results and best practices in historical ecology", *BioScience*, Vol. 65/9, pp. 932-939. [41]

McClenachan, L., F. Ferretti and J. Baum (2012), "From archives to conservation: why historical data are needed to set baselines for marine animals and ecosystems", *Conservation Letters*, Vol. 5, pp. 349-359. [52]

McGill, B. et al. (2015), "Fifteen forms of biodiversity trend in the Anthropocene", *Trends in Ecology & Evolution*, Vol. 30(2), pp. 104-113. [12]

McPherson, M. and M. Nieswiadomy (2005), "Environmental Kuznets curve: Threatened species and spatial effects", *Ecological Economics*, Vol. 55/3, pp. 395-407. [27]

McRae, L., V. Price and B. Collen (2012), *The Living Planet Index for Africa*. [19]

Millennium Ecosystem Assessment (2005), *Ecosystems and Human Well-being: Synthesis*, Island Press, Washington, DC. [4]

Mills, J. and T. Waite (2009), "Economic prosperity, biodiversity conservation, and the environmental Kuznets curve", *Ecological Economics*, Vol. 68/7, pp. 2087-2095. [25]

Naeem, S. et al. (2016), "Biodiversity and human well-being: An essential link for sustainable development", *Proceedings of the Royal Society B: Biological Sciences*, Vol. 283/1844, 20162091. [5]

Naidoo, R. and W. Adamowicz (2001), "Effects of economic prosperity on numbers of threatened species", *Conservation Biology*, Vol. 15/4, pp. 1021-1029. [26]

Pandolfi, J. et al. (2003), "Global trajectories of the long-term decline of coral reef ecosystems", *Science*, Vol. 301/5635, pp. 955-958. [36]

PBDB (n.d.), *ThePaleobiology Database*, https://paleobiodb.org/#/. [35]

Pimm, S. et al. (2014), "The biodiversity of species and their rates of extinction, distribution, and protection", *Science*, Vol. 344/6187, 1246752. [50]

Scheffer, M. et al. (2001), "Catastrophic shifts in ecosystems", *Nature*, Vol. 413/6856, pp. 591-596. [56]

Seebens, H. et al. (2017), "No saturation in the accumulation of alien species worldwide", *Nature Communications*, Vol. 8/14435. [9]

Siddig, A. et al. (2016), "How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in ecological indicators", *Ecological Indicators*, Vol. 60, pp. 223-230. [17]

Stern, D. (2004), "The rise and fall of the environmental Kuznets curve", *World Development*, Vol. 32/8, pp. 1419-1439. [22]

Szabo, J. et al. (2012), "Global patterns and drivers of avian extinctions at the species and subspecies level", *PloS one*, Vol. 7/10, e47080. [39]

Takács-Sánta, A. (2004), "The major transitions in the history of human transformation of the biosphere", *Human Ecology Review*, pp. 51-66. [57]

Ten Brink, B. (2000), *Biodiversity indicators for the OECD Environmental Outlook and Strategy*, RIVM report, Bilthoven, The Netherlands. [42]

Tevie, J., K. Grimsrud and R. Berrens (2011), "Testing the environmental Kuznets curve hypothesis for biodiversity risk in the US: A spatial econometric approach", *Sustainability*, Vol. 3/11, pp. 2182-2199. [29]

van Goethem, T. and J. van Zandem (2019a), "Who is afraid of biodiversity? Proposal for a research agenda for environmental history", *Environment and History*, Vol. 25/4, pp. 613-647. [33]

van Goethem, T. and J. van Zanden (2019), "Economic development and biodiversity", *CEPR Discussion Paper*, No. 13544, https://ssrn.com/abstract=3341351. [6]

van Strien, A. et al. (2016), "Modest recovery of biodiversity in a western European country: The Living Planet Index for the Netherlands", *Biological Conservation*, Vol. 200, pp. 44-50. [45]

van Strien, A. et al. (2019), "Over a century of data reveal more than 80% decline in butterflies in the Netherlands", *Biological Conservation*, Vol. 234, pp. 116-122. [48]

Notes

[1] Historical studies on biodiversity among plants often resort to early taxonomical works (and herbals) to get a grip on the species composition in certain areas. Information on species abundances is more difficult to infer from these types of sources. Agricultural statistics and historical maps have been used to estimate these types of indicators.

[2] Data types that are not used in the database include data from experimental observations (testing different treatments of populations over time), survival rates, recruitment data (such as number of eggs, catches or hunting data, unless they are combined with a measure of effort), or data where survey area or method of data collection has changed over time and has not been corrected for.

[3] Information on the geographic location and method are included in the LPD. To present an accurate measure of changes in species abundance, both geographical location and method of computation are kept unchanged in each year. If the method of computation changes within the period, the data are entered separately. If only part of the dataset is appropriate for entry, all other values are excluded.

[4] Data for 10 species relied on interpolation, for 24 species on extrapolation, and for 12 species on a combination of inter- and extrapolation. Linear interpolation was used when data for no more than three consecutive 5-year intervals were missing. When data gaps were larger, additional qualitative information was used. Population data were extrapolated when data were missing for the earliest time period. Extrapolation was based on qualitative and quantitative information from journals, reports and abundance data from later periods. Estimates were made conservatively, e.g. not assuming a linear trend upfront.

[5] Regional developments by income levels according to the *Living Planet Index 2012*: Globally, there was a 58% decrease, high-income countries had a 7% increase, middle-income countries had a 31% decrease, and low-income countries had a 60% decrease.

[6] There is no reason to assume there is an “automatic” environmental Kuznets curve, i.e. that there is a U-turn in the link between GDP and the degradation of nature. Simply put: higher GDP means more production and consumption, which – all other factors remaining equal – implies more pollution, less space for nature and, therefore, biodiversity loss. Changes in the structure of the economy leading to the relative decline of activities that exploit natural resources together with technological change as such (resulting in higher levels of efficiency in using resources) may to some extent help to explain this (Stern, 2004[22]). Moreover, it is important to further explore the “outsourcing” of biodiversity impacts to other countries by moving industrial production abroad. The key to understanding the U-turn, we hypothesise, is changing citizen attitudes towards the environment. The idea is that GDP growth and the increased scarcity of nature will lead to growing concerns about the environment, resulting in the emergence and growth of a “green” civil society. The development of environmental NGOs will in turn lead to increased pressure on the government to introduce and implement environmental policies, for example aimed at the reduction of pollution or the conservation of nature. The quality of the political system is, however, an important intervening variable: will the polity “translate” the bottom-up pressure to carry out green policies into effective measures for the protection of nature?

11 A composite view on inequality and well-being

Auke Rijpma, Utrecht University

This chapter provides an overall picture of well-being based on the whole body of evidence included in the two volumes of *How Was Life?* When considering average (country-level) measures, both the different indicators included in individual chapters as well as the composite indicator included in this chapter highlight great progress and convergence in well-being across the world. However, great differences between world regions continue to exist. Moreover, a slowdown in a number of indicators has become visible in recent decades. When focusing on (within-country) well-being inequalities, income inequality has increased again over the last few decades, while inequality in education and length of life have continued to decrease.

Introduction

By the end of the 2010s, the global economy was producing far more than at any time in the past. However, it is not equally evident that all the world's citizens have improved their living conditions at the same pace. Has everyone reaped the benefits of increased GDP?

While GDP and productivity are important measures of material progress, for nations as well as individuals, there is a widely felt need to look "beyond GDP" to understand people's well-being. To this end, statistical agencies and academic researchers are expanding data collection and analyses to include different types of well-being metrics. Since well-being is inherently multidimensional, this implies tracking a wide range of indicators to obtain a more complete picture of development (Stiglitz, Sen and Fitoussi, 2009[1]; Boarini and Mira D'Ercole, 2013[2]).

Economic and social historians are no exception to this trend. They had already gathered much data about development and well-being in the long-run. Thanks to the efforts to create historical national accounts, country-level series on GDP and productivity are widely available (Maddison, 2001[3]). These series are continuously being improved and are now going back further in time (Bolt and van Zanden, 2014[4]). Equally important, however, is that historical statistics have always included other well-being measures such as real wages, height and mortality (Steckel and Floud, 1997[5]; Feinstein, 1998[6]; Bengtsson and van Poppel, 2011[7]). Efforts to collect, harmonise and analyse such data are resulting in an ever-clearer picture of long-run developments. The previous *How Was Life?* volume presented the state of the art in this field (van Zanden et al., 2014[8]). Besides GDP per capita, it provided insights into the long-term development of wages, educational attainment, health (life expectancy and height), safety, political freedoms, the environment and gender inequality.

While long-run developments in well-being are clearer than ever, one key challenge is to move beyond country averages to focus on the distribution of well-being outcomes within countries. For one, this is necessary because these averages will always miss part, or sometimes even most, of the distribution of well-being. Moreover, inequality is often seen as an obstacle to achieving societal progress in general (OECD, 2015[9]). However, most of this research focuses on income inequality, while inequality in other dimensions, such as wealth, health and education, could be equally salient.

This volume therefore has a twofold contribution compared to the previous *How Was Life?* report (van Zanden et al., 2014[8]). The first is to expand the set of average (country-level) well-being indicators for which we have long-term series. These indicators are working hours, extreme poverty and improved per capita GDP series. The second contribution is to go beyond the income and gender inequalities covered in the previous volume to look at inequalities in education, length of life and wealth.

The goal of this chapter is to provide an integrated view based on this new historical data. The chapter focuses on the new indicators covered in this volume that have sufficient country coverage, such as working hours and extreme poverty, but also builds on some of evidence from the previous *How Was Life?* report. Above all, however, this chapter analyses the development of inequality of well-being at the world level. After introducing the data and concepts, the chapter describes trends in within-country inequality, and then develops a set of composite indicators to summarise progress in the full dataset of historical well-being indicators underpinning this book.

Description of the concepts used

The underlying data and concepts used are described in detail in the relevant chapters of this volume and in (van Zanden et al., 2014[8]). This section briefly summarises the most important points about the newly added data, and describes how these data are used in this chapter.

First, the measures of inequality in length of life and education are to an extent affected by a "right truncation" in the distribution as, at some point, obtaining more education or greater longevity becomes difficult. This means that, as the average level of education or life expectancy increases and comes closer to the truncation point, inequality mechanically decreases, a pattern that does not necessarily hold for income and wealth. It should therefore be remembered that inequality trends in these indicators might be driven by countries' positions relative to the maximum achievable levels.

The poverty indicator used here is extreme poverty, measured using a cost-of-basic-needs approach. It measures what share of the population cannot afford a basic consumption bundle. In this approach, the poverty metric deviates from the World Bank's "one-dollar-a-day" method, which uses an average of national poverty lines as the threshold (Chen and Ravallion, 2010[10]; Ravallion, Datt and van de Walle, 1991[11]). From a long-term perspective, however, the two indicators provide broadly consistent pictures (Ravallion, 2020[12]). Unless stated otherwise, the regional and global series of inequality indicators in this chapter are population-weighted averages of country-level Gini coefficients. Consequently, the resulting figures cannot be interpreted as a proper Gini for the world or a world region, i.e. a Gini coefficient calculated over all individuals in the world, irrespective of where they live. While the population-weighted averages can give an impression of regional developments in inequality, this limitation should be kept in mind.

More generally, all regional and global trends presented in this chapter are population-weighted averages based on all countries for which data are available in the Clio-Infra database, i.e. stretching beyond the 25 countries covered in this book. To limit the risk that the trends in the global and regional series presented in this volume simply reflect the greater availability of data over time, data for countries with missing data are imputed by (log-linear) interpolation and extrapolation.[1] Regional and global averages are reported only if at least 40% of the population in a region or the world is covered by non-imputed data.

To summarise the developments in the many well-being indicators gathered in this book, this chapter makes use of composite indicators like the one that was used in the previous *How Was Life?* volume (van Zanden et al., 2014[8]). In this chapter, this composite indicator is updated with the new indicators of the present volume. The other novelty is that two separate composite indicators are presented: one for the indicators on average well-being in countries, and one for the measures of inequality of well-being within countries.

Composite indicators can be controversial. The heart of the issue is that the indicators being aggregated are conceptually different and have different units and scales. Combining them into one number requires forcing the variables to a common scale, selecting an aggregation function and choosing weights. The trade-offs in composite indicators implied by this procedure amount to statements on the relative importance of each indicator for overall well-being (Ravallion, 2012[13]; Ravallion, 2012[14]). Since people can hold different opinions on the relative importance of each dimension of well-being, it is very difficult to devise a completely satisfactory solution to this issue.

That said, composite indicators have their own benefits. Above all, they are useful tools to summarise the large number of indicators gathered for a volume such as this. It is likely that readers will summarise the masses of data they see, either on their own account, or because the introduction, conclusion or executive summary highlight certain patterns at the expense of others. The composite indicators presented in this chapter do this in a systematic, disciplined and transparent way.

This chapter relies on the same approach to composite indicators as used in the previous *How Was Life?* volume (Rijpma, 2014[15]). A latent variable model is used to extract one or more common factors from the variables entering the model. It does this by finding the shared information between the indicators in a way that distinguishes between countries as best as possible. To this end, the procedure assigns higher weights to indicators that are highly correlated, and vice versa.

The main disadvantage to this approach is that such a statistical procedure is not guaranteed to provide correct, or even satisfying trade-offs. When each indicator captures a unique part of well-being, a latent variable model can give problematic results, because such variables might have a low correlation. There are of course also advantages to the approach. For one, the functional form of aggregation implicit in the latent variable approach comes down to a linear aggregation with minimal transformations. The indicators are standardised only to have a mean of zero and a standard deviation of one to facilitate computation. This keeps the trade-offs simple and transparent (Ravallion, 2012[13]; Chakravarty, 2003[16])). The specific model used here can moreover deal with missing data, which is an important issue for composite indicators, because they need full data for a given year and country to be calculated. This is of course difficult to achieve with historical data without an imputation procedure. Finally, the statistical approach used in this chapter can also provide estimates of uncertainty, including that caused by the imputation of missing data, and can do so at the regional level for which much of the data is reported. Details can be found in (Rijpma, 2016[17]) and (Rijpma, 2014[15]), which are in turn based on (Jackman, 2009[18]) and (Høyland, Moene and Willumsen, 2012[19]).

Historical sources and data quality

A detailed discussion of the sources and data quality of individual measures is provided in the other chapters of the current and previous volumes (van Zanden et al., 2014[8]). In this section, only a number of general features are noted.

Overall, data quality improves as we move closer to the present. For the period after World War II, the data behind most indicators are gathered either by statistical agencies or by researchers using similar methods to statistical agencies. Prior to World War II, estimates are frequently based on research using imperfect data. In the earliest decades, in particular the first half of the 19th century, the quality of the data is lower as data becomes scarcer, and guesstimates become inevitable.

The quality of data behind the new indicators that are analysed in this chapter should be summarised briefly. For working hours, data falls in the highest-quality category from the 1930s onwards. When data are available for earlier periods in Western Europe and the Western Offshoots, they are also typically of fairly high quality (research using the same methods as statistical agencies). Coverage outside these regions before the 1930s is however limited. Measures of extreme poverty are of high or fairly high quality since the 1950s. Data for the 19th century are typically worse, with guesstimates becoming common, often based on estimates of per capita GDP and income inequality. It is also important to note that price data in socialist countries can be unreliable, because they could be set by the government and goods were not always available to be bought at those prices. Because of this, guesstimates remain necessary in these countries until well into the 20th century.

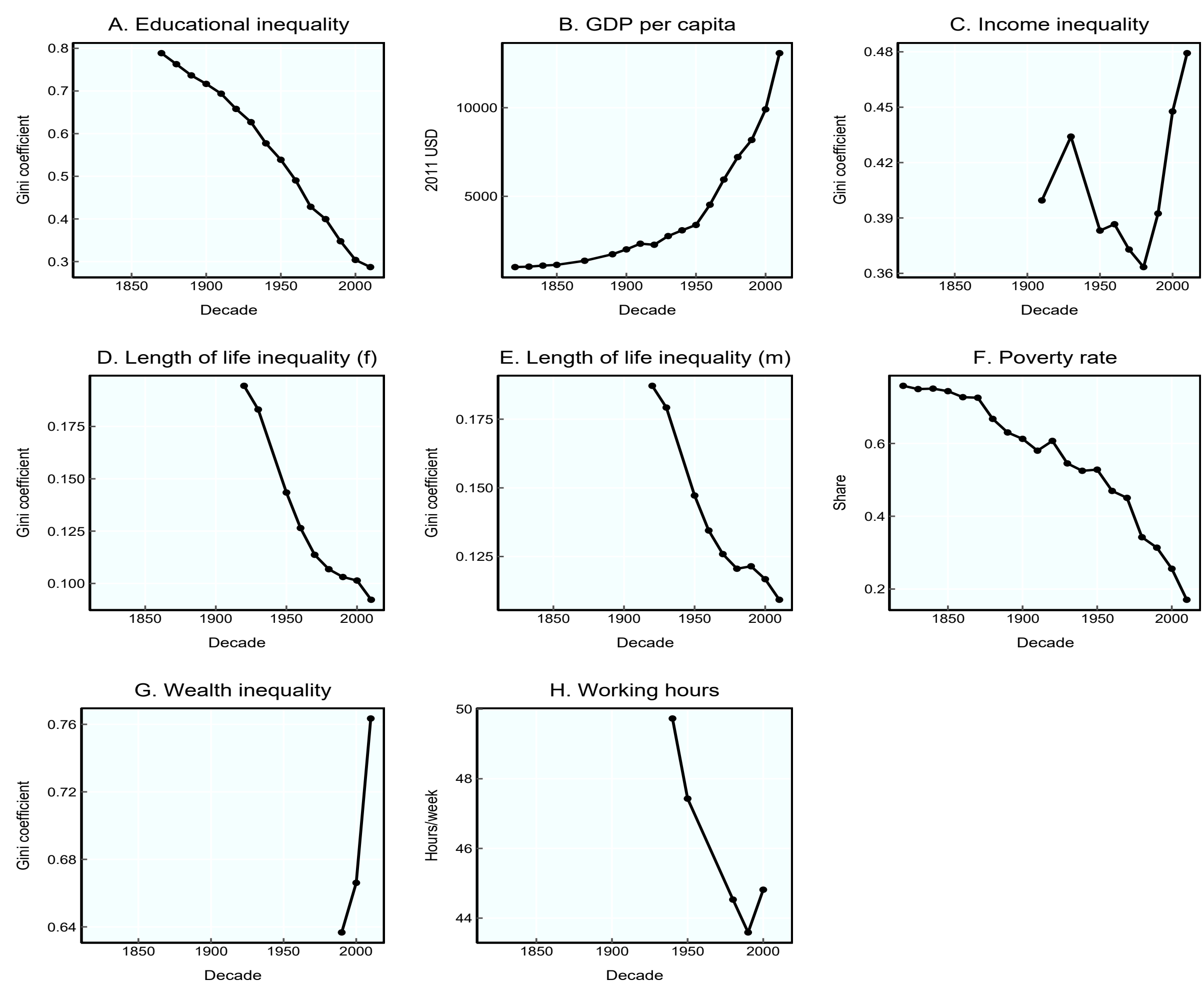
Regarding the new inequality measures, the available data on inequality in length of life are of good quality. However, coverage is limited outside the advanced economies of Western Europe and its Offshoots before the 1950s. The data underlying the educational inequality estimates are of fairly high quality from the 1950s onwards. Data are scarcer for the first half of the 20th century, and guesstimates are frequent in the 19th century. Data on wealth inequality, finally, are very scarce outside Western Europe and the Western Offshoots. Even as late as the 1990s, estimates are available for less than 10% of the countries in the Clio-Infra datasets. For this reason, this indicator is unfortunately not part of the inequality analysis of this chapter.

Main highlights

Figure 11.1 shows the world population-weighted average over time for six new indicators added in this volume, the new historical GDP per capita series described in Chapter 2, as well as the income inequality series included in van Zanden et al. (2014[8]). The developments in these indicators confirm one of the overall conclusions of the previous *How Was Life?* volume: that of improvements in well-being in the world over the past 200 years. The new GDP per capita series utilising better PPPs still shows huge progress in productive capabilities over the past 200 years. While this progress is not distributed evenly over the globe, the average economic well-being of the world's citizens has improved considerably.

Figure 11.1. World averages of selected well-being indicators, 1820-2010

Population-weighted averages



StatLink https://stat.link/pumzao

Globally, extreme poverty decreased over almost the entire period. Progress was particularly fast in the 1950s, 1970s, 1990s and 2000s, when on average, extreme poverty declined by five percentage points or more over each 10-year period. Recently, the decline has been concentrated in Asia and sub-Saharan Africa; prior to that, the declines were concentrated in the other regions.

Weekly working hours in manufacturing declined throughout the period as well. While country coverage outside Western Europe and the Western Offshoots before 1950 is limited, from that moment onwards working hours declined across the globe. In a forty-year period, they declined from 50 hours to less than 44 hours per week in the 1990s. For countries where data are available from an earlier date, working hours were higher still, commonly as high as 60 in the 19th century. The decline was set in motion in the second half of the 19th century, with the largest progress made in the first half of the 20th century. Very recently, weekly working hours have been increasing again.

The inequality indicators display a more diverse pattern. On average, income inequality decreased from the early 20th century up until the 1980s, after which it started increasing again. The coverage for wealth inequality allows us to make statements only for the 1990s and later, when it was clearly increasing. However, a slow decline in wealth inequality before this period can be observed for the countries for which data are available.

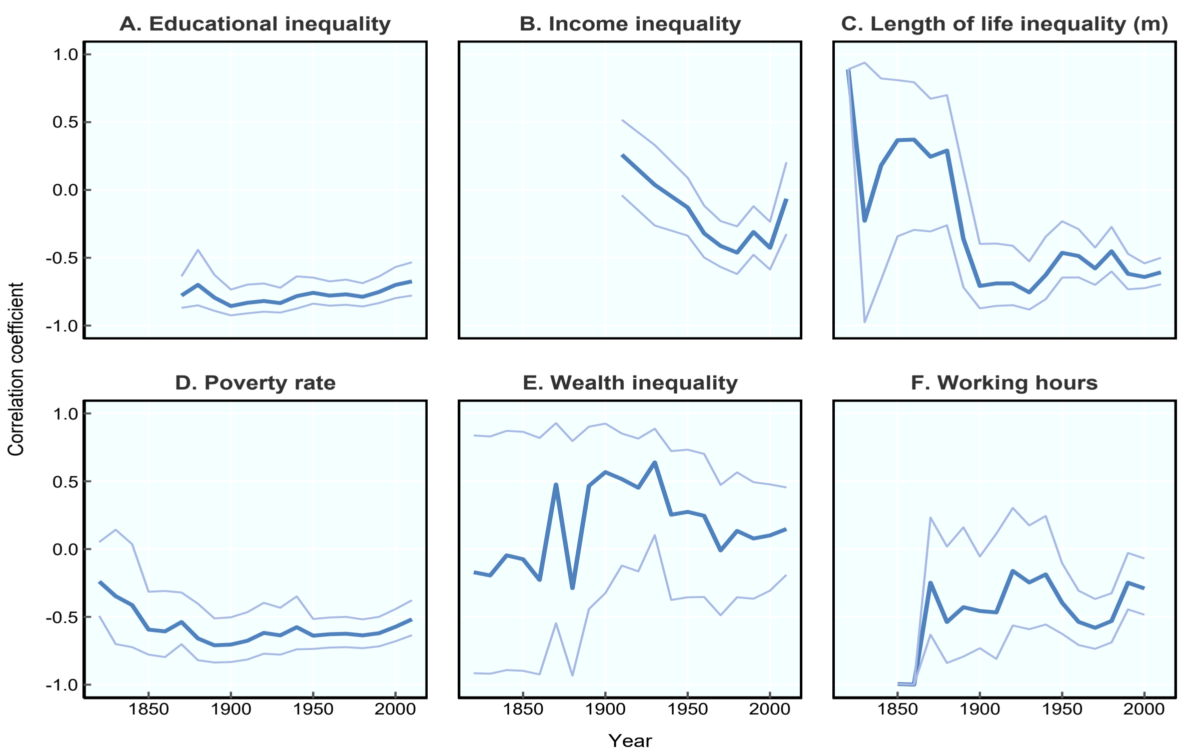
Length-of-life inequality decreased throughout the period for which we have sufficient data, for both men and women. Only in the 1990s can a slowdown be observed, mostly for men. This decrease was mostly concentrated in Eastern Europe and the former Soviet Union, where inequality increased for men. Educational inequality too has been decreasing for most of the period covered by our data, with a slowdown becoming visible only in the 2010s. These changes in inequality suggested by Figure 11.1 are investigated in more depth in the next section.

Correlation with GDP per capita

To understand these trends as well as those for the composite indicator discussed below, it is useful to look at the correlation of the indicators with per capita GDP (Figure 11.2). Compared to the core set of indicators in the previous volume, correlations with per capita GDP are somewhat weaker. Whereas correlation coefficients of 0.5 or higher were consistently reported for indicators such as (average) real wages, average years of education, or life expectancy, such large coefficients are less frequent now.

Figure 11.2. Correlation of selected well-being indicators with GDP per capita, by decade

Pearson correlation coefficient



Note: The thick and thin lines depict the correlation coefficient and 95% confidence intervals respectively.

StatLink https://stat.link/5mfjpb

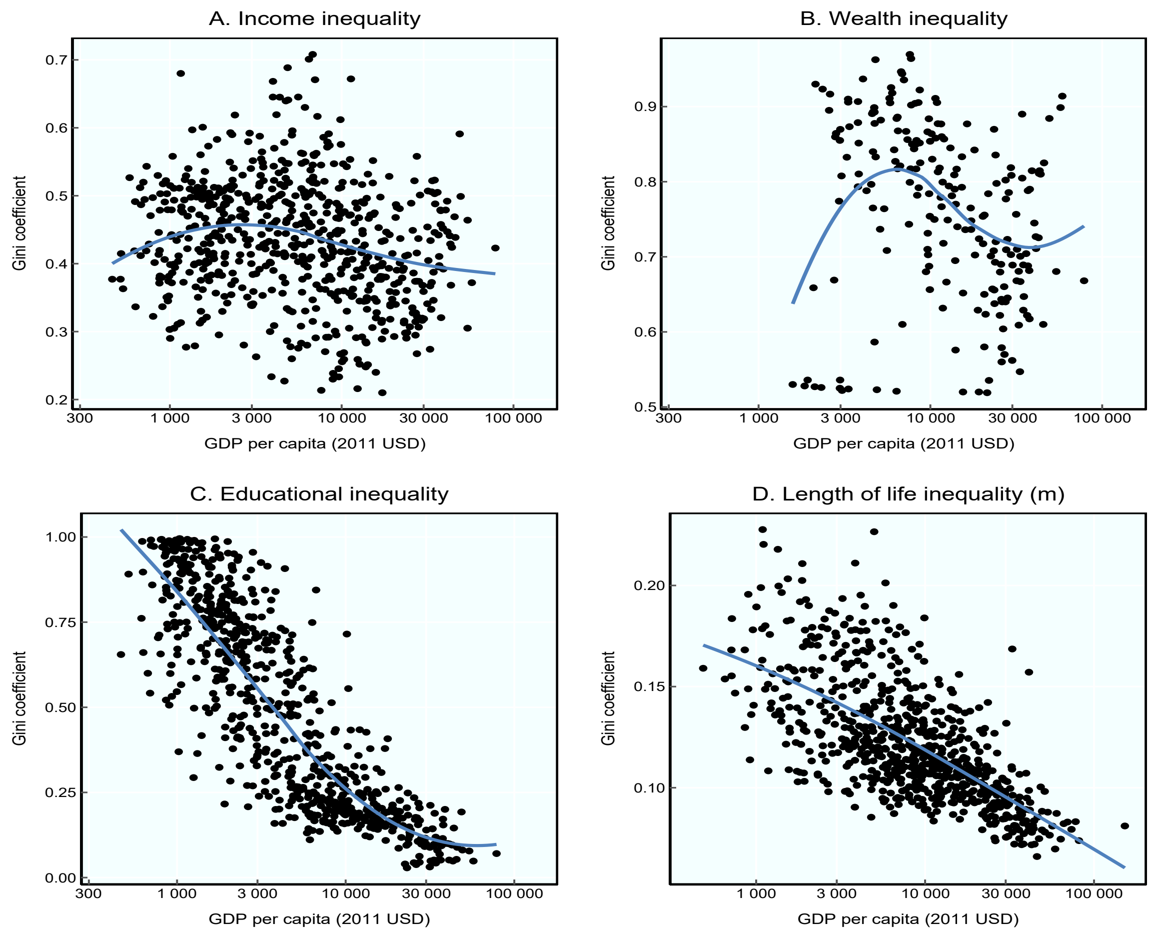
Extreme poverty does have a fairly strong negative correlation (-0.5/-0.7) with GDP per capita, although well below the strong correlation suggested by (Dollar and Kraay, 2002[20]). Figure 11.2 shows a consistent, negative correlation of working hours with GDP per capita, which is however not very strong and measured with some uncertainty. The correlation is clearly negative only from the 1960s onwards and has become weaker in recent decades, due to increases in working hours in a number of high-income regions.

Looking at the inequality indicators, the correlation of income inequality with per capita GDP has changed considerably over the course of the 20th century. In the early part of the century, the correlation was positive, meaning that high-income countries also experienced high-income inequality. This relation had reversed after World War II, so that advanced economies tended to be more equal. Today the correlation is close to zero.

Length-of-life inequality and educational inequality show a negative correlation with GDP per capita, implying that richer countries had lower inequality in health and education. In the 19th century, however, the correlation with length-of-life inequality is measured with substantial uncertainty, implying that it is probably safest to say that, in that period, length-of-life inequality did not display a strong relation with per capita GDP. Conversely, the correlation of educational inequality with per capita GDP is consistently negative. For both inequality indicators it is worth remembering that the levels of average years of education and life expectancy are strongly correlated with per capita GDP in the same period (van Leeuwen and van Leeuwen-Li, 2014[21]; Zijdeman and Ribeiro de Silva, 2014[22]).

Figure 11.3 shows the overall relation of the three inequality indicators with GDP per capita. Overall, all inequality measures display a negative correlation with GDP per capita. Strong negative relations are observed for educational inequality and length-of-life inequality. Income inequality and wealth inequality, however, show a more complex pattern, with a positive relation at low levels of GDP per capita, and a negative one at higher levels.

Figure 11.3. GDP per capita and Gini coefficient for income, length of life and education, 1870-2010



Note: The Y-axes of the different panels are on a different scale. The blue curves correspond to a locally weighted polynomial fit.

StatLink https://stat.link/mxdv8u

Trends in well-being inequality

The discussion of the main global trends in well-being inequality in the previous section suggests that the 20th century may have been characterised by a U-shape in inequality in a number of dimensions. From a high point at the start of the 20th century, income inequality declined to a low in the 1970s-80s, after which it started rising again, This U-shaped pattern matches the findings of recent research on income and wealth inequality (Piketty, 2014[23]; Scheidel, 2017[24]). While educational inequality and length-of-life inequality declined more consistently, there too a slowdown can be observed more recently.

Figure 11.4 looks at the average change in the Gini coefficient in income inequality by region to pinpoint the moment when income inequality started declining. Looking at Western Europe and its Offshoots, the 1920-29 period was, on average, still characterised by increasing inequality. From the 1930s to the 1970s in Western Europe, and until the 1960s in the Western Offshoots, income inequality declined. Therefore, according to the data presented here, the decline in income inequality in the advanced economies did not start immediately after World War I, but only began in earnest during the 1930s. Likewise, the rise of income inequality in the Western Offshoots in the 1960s seems to pre-date the breakdown of the Keynesian consensus and the adoption of neoliberal policies. More precise data than the decennial estimates presented here should however be used to determine this turning point more precisely.

Figure 11.4. Regional average change in the Gini coefficient for income inequality

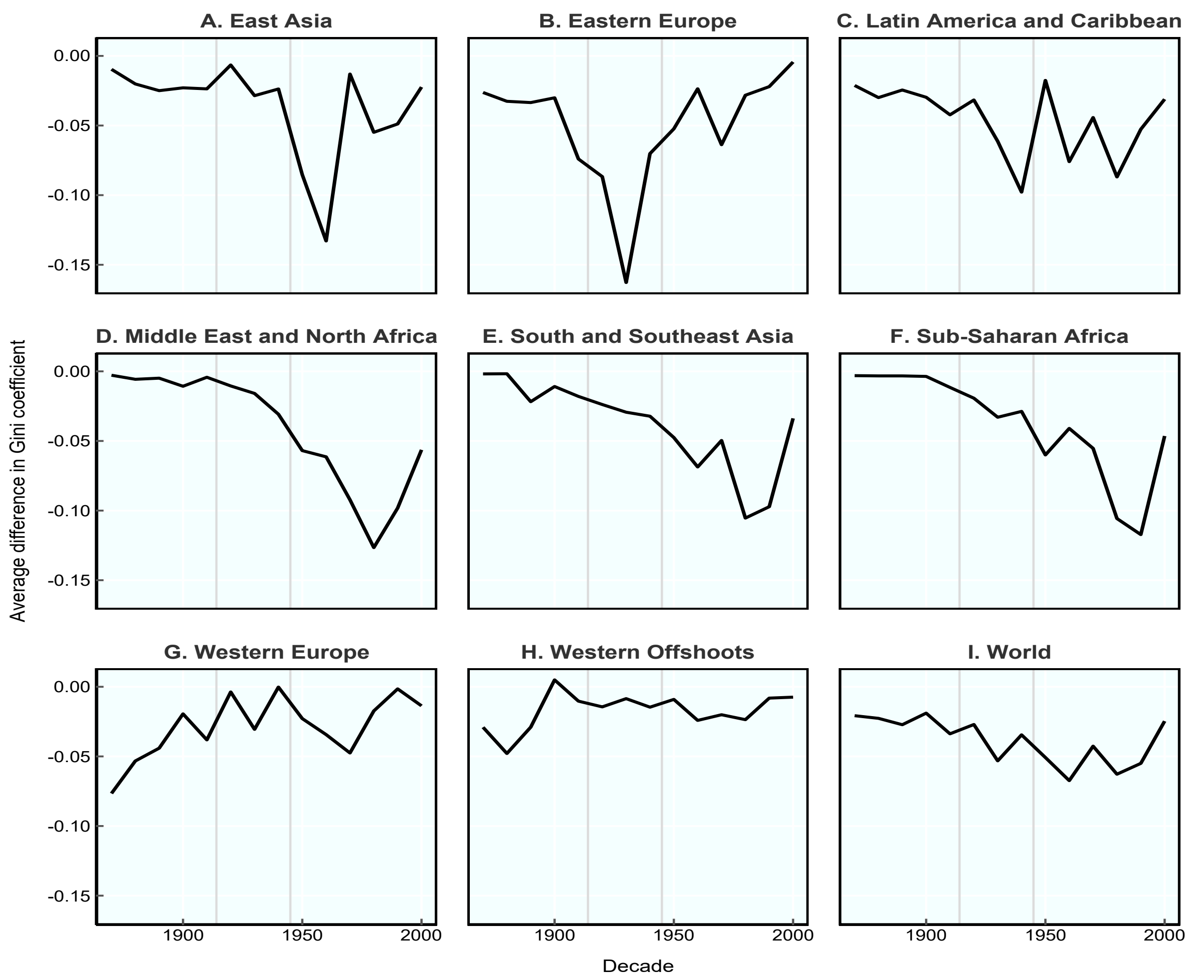
Note: The vertical reference lines correspond to 1914 and 1945.

StatLink https://stat.link/p4058o

Moving to other regions, the decline in income inequality in the 1930s and 1940s affected most regions with the exception of South and Southeast Asia, although in the Middle East and North Africa (MENA) the decline was very small. The recent increase of income inequality also extended to Asia and particularly Eastern Europe and the former Soviet Union, although not to sub-Saharan Africa, the MENA and the Latin America and the Caribbean.

Educational inequality declined throughout the period (Figure 11.5). The decline was strongest in the 1950s-80s period but has recently slowed down to a point of near-stagnation. The decline in educational inequality was less pronounced in Western Europe and the Western Offshoots, regions where educational inequality was already fairly low in the 1870s. However, the decline in educational inequality affected all regions throughout the period. The U-shaped pattern for income inequality does not apply to educational inequality.

Figure 11.5. Regional average change in Gini coefficient for educational inequality, 1870-2010

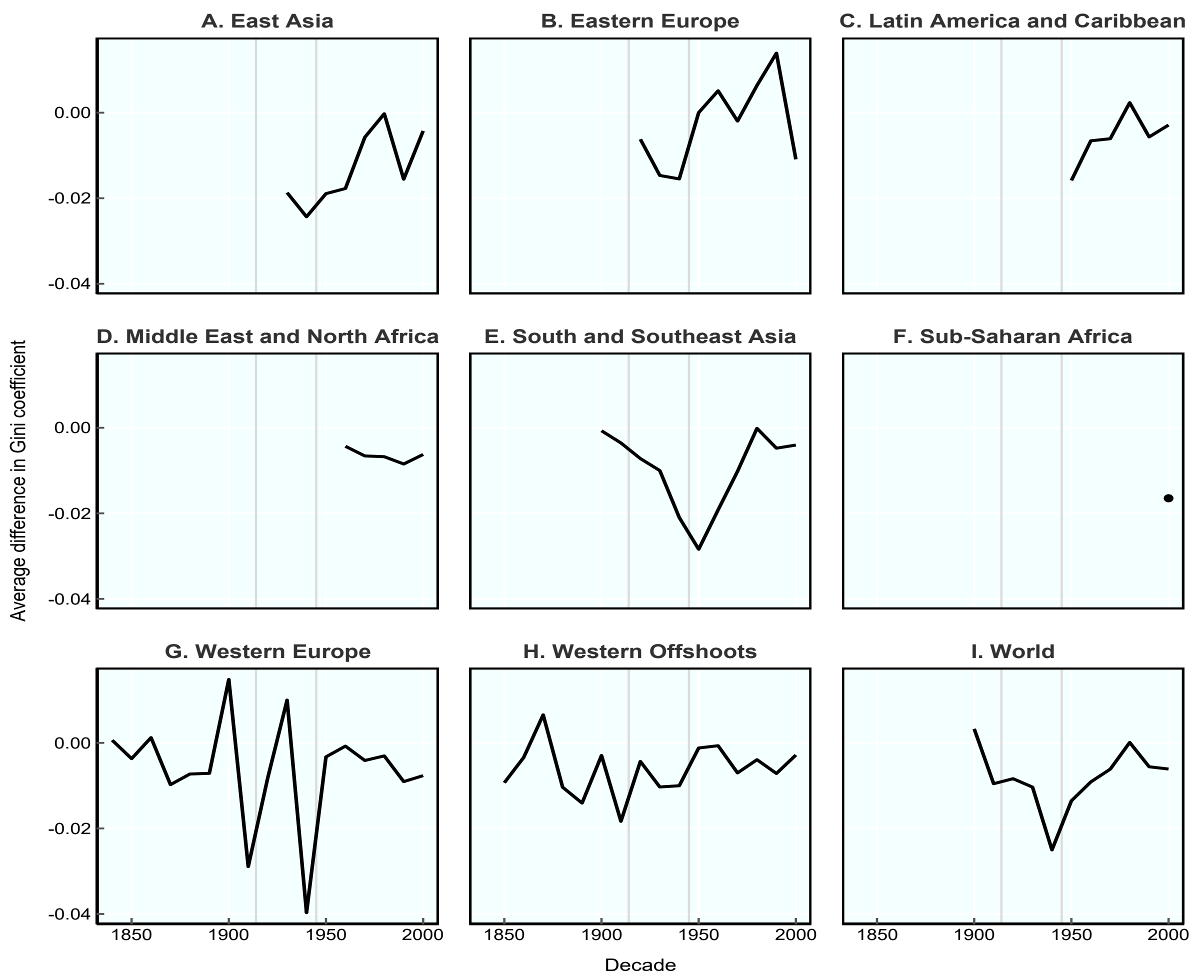


Note: The vertical reference lines correspond to 1914 and 1945.

StatLink https://stat.link/ovmst3

Finally, we look at regional changes in length-of-life inequality (Figure 11.6). Data coverage is an issue here, with few countries outside Western Europe and the Western Offshoots having enough data to allow for calculating regional averages over a long period. What stands out from Figure 11.6 is that, in most of the world's regions, the trend is towards lower inequality. This trend is common to all regions and concentrated in the middle of the 20th century. The slowdown of this decline in the late 20th century is also fairly widespread, but particularly pronounced in Eastern Europe and the former Soviet Union, where length-of-life inequality even increased.

Figure 11.6. Regional average change in Gini coefficient for length-of-life inequality, 1870-2010



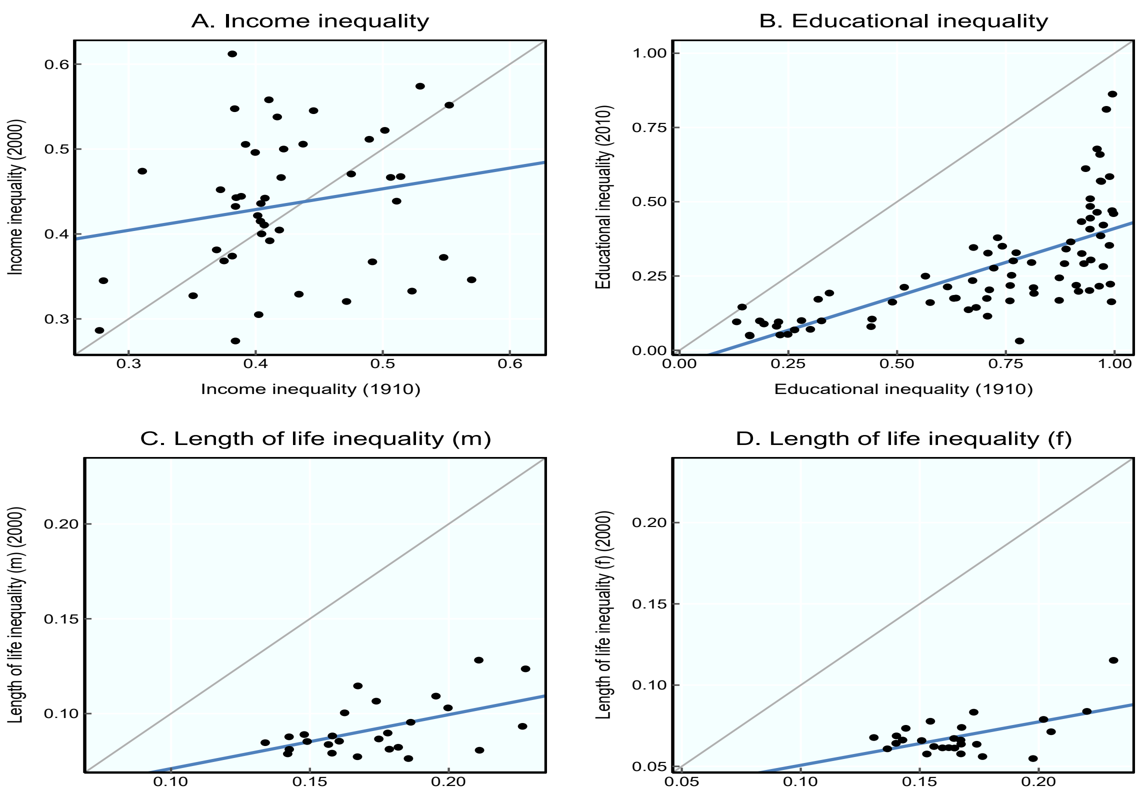
Note: The vertical reference lines correspond to 1914 and 1945.

StatLink https://stat.link/3h7ri0

Figure 11.7 compares inequality between 1910 and 2000/2010, two points in time for which data coverage is relatively good for all inequality indicators. What stands out is that, in the case of income, inequality in 1910 is very weakly correlated with inequality in 2000. This also means that that there are countries featuring high inequality in 1910 that are ranked relatively low in 2000, and vice versa. Countries such as Sweden, France, Japan, Italy and the Netherlands moved from being highly unequal countries in terms of income to being relatively equal countries in 1990. Countries like South Africa, Brazil and India show the opposite development.

The other dimensions of inequality display a different pattern. Often, inequality in 1910 is predictive of inequality in 2000/2010. Moreover, there is a degree of convergence, as very unequal countries in 1910 achieved stronger reductions in inequality than the most equal countries in the same year. What this means for multidimensional inequality is that countries that became more equal in terms of income inequality have also become more equal in other dimensions. Countries that became more unequal in terms of income over the 20th century compensated for this to some extent in other dimensions.

Figure 11.7. Gini coefficients for inequality in income, education and length of life, 1900 compared with 2000/2010

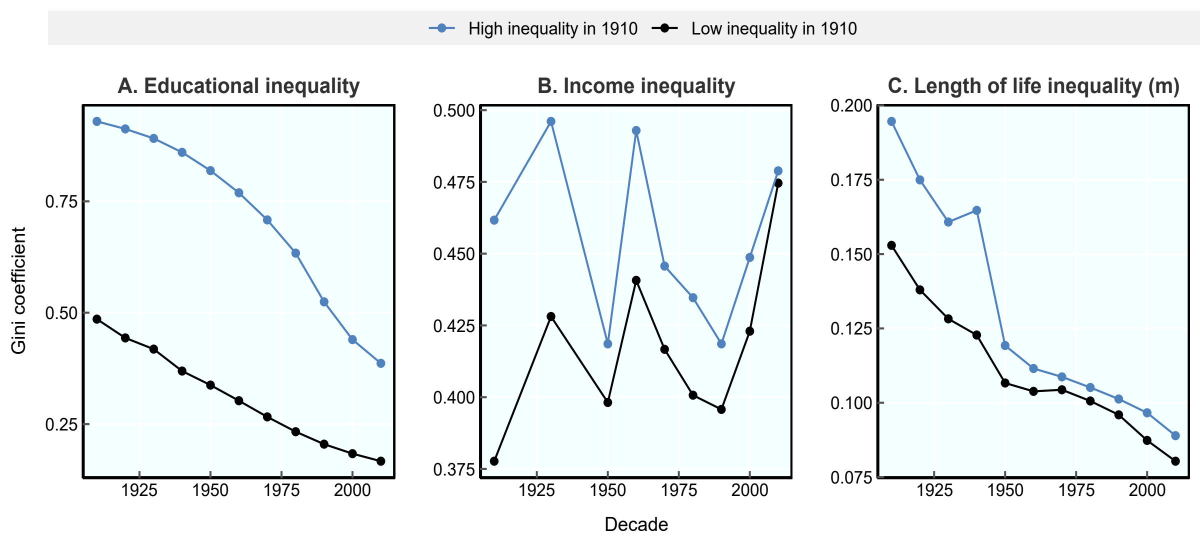


Note: Data on length of life inequality plotted separately for women and men. The grey line corresponds to the locus of points where the two indicators have the same values. The blue line gives the regression fit.

StatLink https://stat.link/wrs9xq

In Figure 11.8, these developments are analysed further by considering the entire inequality trajectory for the 25 “Clio-Infra countries” – countries with good historical data that together cover a large share of the population from various world regions.[2] By splitting the sample by the level of inequality in 1910, this figure shows whether high-inequality countries developed differently from low-inequality ones. High-inequality countries are defined as those with above median inequality in 1910. Figure 11.8 shows that in the case of educational inequality, high and low inequality countries in 1910 converged. Despite fewer observations, developments in length-of-life inequality are similar. For income inequality, the pattern is however different. Countries with low income inequality became somewhat more unequal on average, and countries with high income inequality became somewhat more equal. In the 1960s and 1970s, both high- and low-income inequality countries converged to a lower level of income inequality, after which increases were common to many countries.

Figure 11.8. Average Gini coefficient for inequality in education, income and length of life, 1910-2010, by level of inequality in 1910



Note: The y-axes have different scales.

StatLink https://stat.link/5plgoz

Composite indicators

As a final way of analysing trends in the well-being indicators covered by the two *How Was Life?* volumes, this section looks at composite indicators. Two distinct indicators are used. The first covers the country averages indicators (levels) from the two volumes. The indicators taken from van Zanden et al. (2014[8]) are average real wages, height, life expectancy, average years of education, biodiversity, democracy and homicide rates. The old GDP per capita series are replaced by the new ones that incorporate the more recent PPP estimates presented in Chapter 2. Data on working hours and extreme poverty are added from the present volume.

The second composite indicator is based on the inequality indicators: income inequality from the previous volume, and length-of-life inequality, educational inequality and gender equality from the present volume. Due to low country coverage before the 1990s, wealth inequality data are not included in this analysis. Because male and female length-of-life inequalities are very highly correlated and because gender equality is already included, only male length-of-life inequality is included (due to its higher coverage).

The weights of the composite indicators implied by the latent variable models used to construct these two composites are shown in Table 11.1. Most variables contribute as expected, the one exception being biodiversity; for a discussion, see Rijpma (2014[15]). The contribution of extreme poverty and working hours is similar in magnitude to that of the core variables from the previous volume, with the exception of average years of education and life expectancy, which make a larger contribution. The inequality measures all enter in the expected direction (for gender equality, a higher score implies less equality, so it is expected to have the opposite sign). This means that high values in the composite indicator correspond to higher inequality. The model assigns a somewhat lower weight to income inequality than to the other inequality indicators.

Table 11.1. Factor loadings for the latent variable model underlying the composite indicators used in this chapter

	Indicator	Mean	q. 05	q. 50	q. 95
A. Country-average measures	GDP per capita	0.74	0.71	0.74	0.77
	Real wage	0.75	0.71	0.75	0.80
	Height	0.76	0.72	0.76	0.80
	Life expectancy	0.99	0.96	0.99	1.02
	Average years of education	0.95	0.94	0.95	0.97
	Polity	0.73	0.70	0.73	0.77
	Biodiversity	-0.35	-0.38	-0.35	-0.33
	Homicide rate	-0.13	-0.20	-0.13	-0.06
	Working hours	-0.77	-0.83	-0.77	-0.70
	Extreme poverty	-0.77	-0.79	-0.77	-0.74
B. Inequality indicators	Income Inequality	0.17	0.12	0.17	0.23
	Length of life inequality (men)	0.73	0.67	0.73	0.79
	Education inequality	0.87	0.84	0.87	0.90
	Gender equality	-0.69	-0.72	-0.69	-0.65

Note: The mean and 5th, 50th and 95th quantile (q. 05, q. 50 and q. 95) of posterior distribution of estimates.

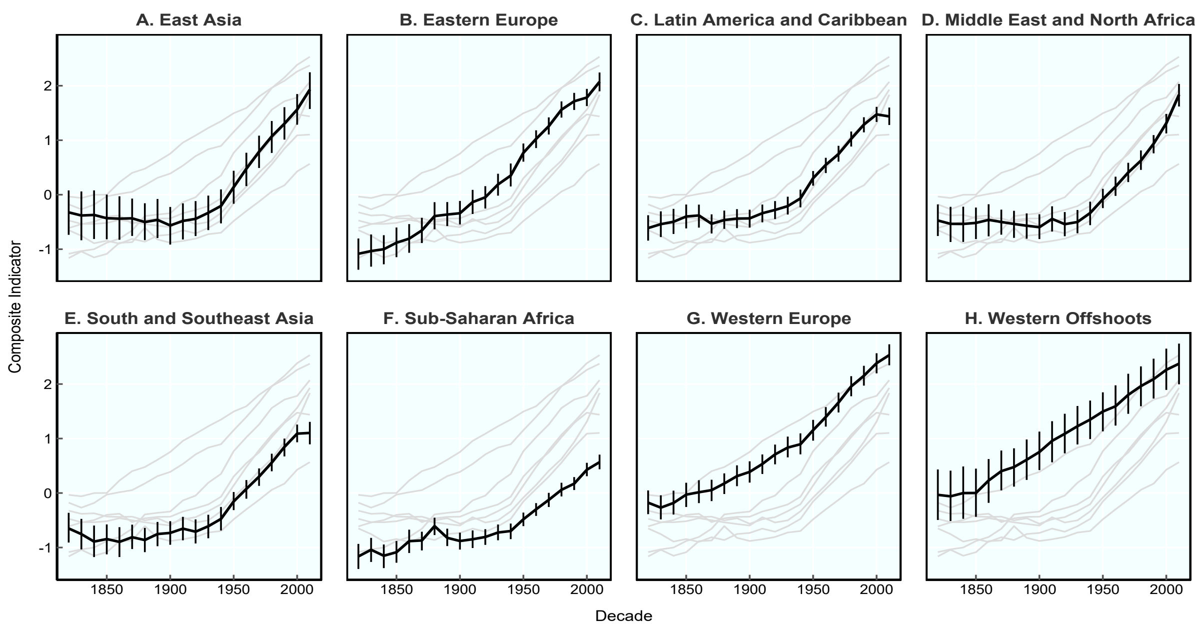
StatLink https://stat.link/49f3ej

Before discussing developments in these composite indicators, it is useful to compare the composite indicator of country averages to the one included in Rijpma (2014[15]). The overall pattern of the two indicators is very similar. For one, the factor loadings in Panel A are largely unchanged from Rijpma (2014[15]). Moreover, Annex Figure 11.A.1 shows that the regional developments implied by each composite indicator are typically very close, differences being visible only in the early period, but still well within the wide confidence intervals for this period. Larger differences exist at the country level but are typically minor.

A number of factors explain this similarity. First, the country coverage for the new indicators is usually lower than for the old set, and the model does not favour variables with a high degree of "missingness". These variables would mostly add to the uncertainty of the estimates. Furthermore, each new variable added to a composite indicator has, by definition, a lower impact than the variables previously included. In other words, since we started with nine variables, adding two additional ones was never likely to result in large changes to the composite indicator, unless they have a very large weight.

Above all, though, Panel A suggests that the indicators used in Zanden et al. (2014[8]) were also the most important ones, and the ones for which data are available in sufficient quantity and quality. Because of this, the latent variable model presented in the previous volume was already capable of capturing the shared information from the old set of variables, while the new variables added in the present volume mostly confirm this pattern. Arguably, this means that some of the variables included in the composite indicator might be redundant, a point further discussed below.

Figure 11.9. Composite indicator of average well-being by region, 1820-2010



Note: The vertical bars indicate 90% confidence intervals.

StatLink https://stat.link/0pxyat

Given the above, the story told in van Zanden et al. (2014[8]) still holds when looking at the composite indicator for country averages shown in Figure 11.9. While there is sometimes considerable uncertainty in the regional estimates, overall, the world has seen great progress, with all regions showing considerable increases in the composite well-being indicator over the 200-year period. Progress is also greater and more equally distributed than that shown by GDP per capita. In terms of the composite indicator, there are no regions or countries that are worse off today than the best-performing countries were in 1820 (Table 11.2). No country in the year 2000 does worse than the United States in 1820. If we were to make this comparison in terms of GDP per capita, some countries would fare worse (see Chapter 2).

However, strong and widespread gains do not mean that these gains proceeded at the same pace across regions. Western Europe and the Western Offshoots performed better than the other regions throughout the period. By the middle of the 19th century, they already had the highest scores on the composite well-being indicator. Both regions kept their lead over other regions throughout the period. That said, clear cases of convergence also occurred, as East Asia, Eastern Europe and the former Soviet Union, the Middle East and North Africa, and Latin America and the Caribbean began to catch up with the two leading regions. This catching-up process began roughly in the middle of the 20th century. Finally, while there has been substantial progress, South and Southeast Asia and, above all, sub-Saharan Africa have not converged to the levels achieved in the leading regions.

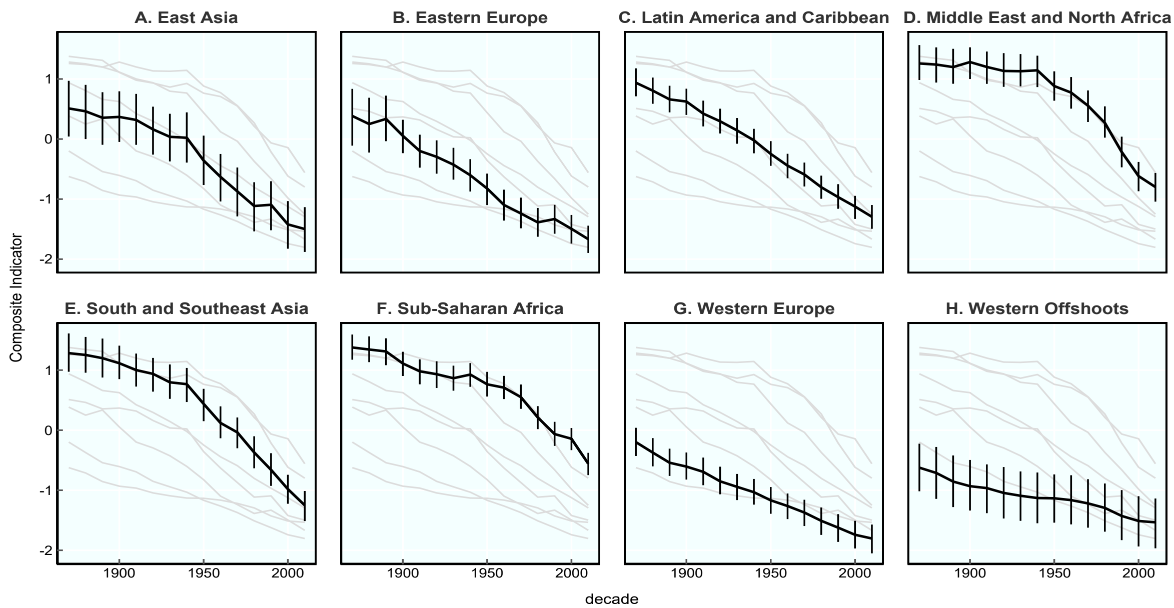
Table 11.2. Composite indicator scores of average well-being in selected countries, 1820-2000

Country	1820	1870	1910	1950	1970	2000
ARG	-0.37	-0.46	-0.02	0.76	1.12	1.65
AUS	-0.35	0.26	1.26	1.81	2.18	2.89
BRA	-0.77	-0.67	-0.65	0.12	0.41	1.49
CAN	-0.12	0.47	1.06	1.80	2.28	2.74
CHN	-0.45	-0.54	-0.52	-0.27	0.37	1.21
DEU	-0.49	0.19	0.75	1.49	2.14	2.70
EGY	-0.43	-0.60	-0.54	-0.10	0.08	1.08
ESP	-0.18	-0.29	0.19	0.75	1.31	2.24
FRA	-0.40	0.01	0.69	1.37	1.91	2.57
GBR	-0.27	0.09	0.75	1.47	2.02	2.87
IDN	-0.79	-0.94	-0.88	-0.53	0.01	1.18
IND	-0.59	-0.78	-0.84	-0.29	0.03	0.77
ITA	-0.31	-0.54	0.00	0.93	1.51	2.43
JPN	-0.61	-0.58	-0.32	0.65	1.69	2.48
KEN	..	..	..	-0.46	0.03	0.57
MEX	-0.85	-0.70	-0.65	0.02	0.64	1.52
NGA	..	..	..	-0.70	-0.35	0.27
NLD	-0.27	0.13	0.65	1.40	1.93	2.51
POL	-0.78	-0.33	0.24	0.51	1.09	1.84
RUS	..	..	..	..	1.39	1.67
SUN	..	..	-0.13	0.67	..	..
SWE	-0.45	0.07	0.69	1.46	2.10	2.52
THA	-0.76	-0.87	-0.73	-0.17	0.33	1.37
TUR	-0.67	-0.65	-0.55	-0.22	0.34	1.22
USA	0.31	0.65	1.09	1.95	2.41	2.95
ZAF	-1.06	-0.66	-0.43	0.11	0.52	0.97

StatLink https://stat.link/l8ge35

Figure 11.10 shows the regional development of the composite indicator of well-being inequalities. The composite inequality indicator is measured with a reasonable degree of certainty, though for some regions, East Asia and the Western Offshoots in particular, it is hard to make comparisons over time and with other regions with substantial certainty.

Figure 11.10. Composite indicator of well-being inequality by region, 1870-2010



Note: The vertical bars indicate 90% confidence intervals.

StatLink https://stat.link/mh4uka

The overall trend is, again, one of improvement, with inequality declining in all world regions. No substantial trends towards higher inequality can be observed when looking at multiple inequality indicators at once. Western Europe and the Western Offshoots perform best throughout the period. It is noticeable that our composite measure of inequality for Western Europe falls below the levels of the Western Offshoots around 1950. The composite indicator of multiple inequality indicators shows that this reversal had already occurred in the early 20th century.

The composite indicator of well-being inequality also displays regional convergence. Differences between regions in 2000 were substantially smaller than in 1900. Eastern Europe and the former Soviet Union in particular have become far less unequal. While sub-Saharan Africa and the Middle East and North Africa do converge with the best-performing regions, they remain the two regions with higher inequality throughout the period. For the Middle East and North Africa, this is a striking development, as the region has made much more progress, both in terms of per capita GDP and of the composite indicator of average well-being levels. Its poor performance in terms of length-of-life inequality, gender equality, and to a lesser extent income inequality, are behind this pattern.

Table 11.3 shows a trend towards declining inequality in most countries. In 1870, most of the countries shown were very unequal, with those featuring positive values of the composite indicator being in the top 50% of countries in terms of inequality. According to the composite indicator, highly developed countries like Sweden, Germany and the United Kingdom are today's low inequality countries.

Table 11.3. Composite indicator scores of well-being inequality in selected countries, 1870-2000

	1870	1910	1950	1970	2000
ARG	0.66	0.14	-0.76	-0.97	-1.41
AUS	-0.38	-1.16	-1.46	-1.55	-1.96
BRA	1.21	0.94	0.26	-0.15	-1.15
CAN	-1.19	-1.26	-1.40	-1.41	-1.80
CHN	0.70	0.52	0.07	-0.69	-1.25
DEU	-0.77	-1.12	-1.44	-1.73	-2.03
EGY	1.45	1.43	0.99	0.74	-0.72
ESP	0.47	-0.23	-1.00	-1.36	-1.58
FRA	-0.33	-0.82	-1.19	-1.54	-1.88
GBR	-0.72	-1.29	-1.42	-1.64	-2.00
IDN	1.50	1.41	0.87	-0.06	-1.20
IND	1.46	1.38	0.99	0.53	-0.30
ITA	0.64	-0.44	-1.19	-1.29	-1.74
JPN	0.74	-0.34	-1.13	-1.49	-1.81
KEN	..	..	0.64	0.19	-0.98
MEX	1.13	0.69	-0.18	-0.77	-1.26
NGA	..	..	1.02	0.71	-0.28
NLD	-0.93	-1.28	-1.40	-1.43	-1.81
POL	0.34	-0.48	-0.17	-0.96	-1.34
RUS	..	..	..	-1.24	-1.47
SUN	..	-0.15	-1.09	..	..
SWE	-0.58	-1.03	-1.36	-1.54	-2.02
THA	1.39	1.18	-0.02	-0.35	-1.30
TUR	1.42	1.34	0.73	-0.03	-0.80
USA	-0.80	-1.12	-1.38	-1.54	-1.81
ZAF	1.01	0.81	0.21	0.01	-0.83

StatLink https://stat.link/f14los

Conclusions and priorities for further research

Generally, the *How Was Life?* volumes have told a story of long-term progress. Looking at the many long-term well-being indicators produced by historical research generally shows that life has been improving throughout the world, on a range of indicators. The new and updated indicators presented in this volume have by and large confirmed this picture. Extreme poverty and working hours have declined, and the improved estimates of GDP per capita have not overturned the general conclusion drawn in the previous volume.

While large parts of the past two centuries, and the second half of the 20th century in particular, conform to this general improvement, this is not the whole story. Progress on some indicators has stalled in recent decades, the rise in working hours being one of the examples presented here. Moreover, these gains are not distributed evenly over the world, with some regions, above all Western Europe and the Western Offshoots, attaining higher well-being levels throughout the 1820-2020 period, and other regions lagging far behind.

The addition of within-country inequality indicators to the picture of well-being sketched in this volume strengthens the above observations. While within-country income inequality decreased since the 1930s in many world regions, it has increased more recently. The other inequality indicators analysed in this volume display a global trend towards more equality.

Suggestions for future research focus mostly on the inequality dimension. For one, more and better data on inequality in dimensions other than income are clearly needed, especially for countries outside Western Europe and the Western Offshoots. Currently, long-run series on inequality other than income have poorer country coverage, especially in the case of length-of-life inequality.

Another issue deserving attention is how to measure inequality when there are "soft limits" to the distribution, as in the case for length of life and years of education. These limits mean that progress in the overall level of an indicator almost automatically reduces inequality, as more and more people approach the limits of educational attainment or longevity. Further research on how to measure inequality in these circumstances may change the conclusions reached in this chapter.

To summarise trends in inequality in well-being in a composite indicator, this chapter has taken a highly practical approach. All inequality indicators were linearly combined, which is not an entirely satisfactory approach. Working with a more limited set of variables, and considering alternative aggregation functions, e.g. (Jones and Klenow, 2016[25]; UNDP, 2010[26]; Atkinson, 1983[27]), could open up new avenues here. This effort should be combined with an investigation of whether any variables in the composite indicators are redundant. This would ease the data requirements of alternative aggregation procedures.

Finally, one important reason for looking at within-country inequality is that country-level averages tell an incomplete story of well-being. Well-being is experienced by individuals. To address this, historical micro-level data, preferably in multiple dimensions for each individual, are needed. This will require substantial data collection efforts and new methods of analysis.

References

Atkinson, A. (1983), *The Economics of Inequality*, Clarendon Press, Oxford. [27]

Bengtsson, T. and F. van Poppel (2011), "Socioeconomic inequalities in death from past to present: An introduction", *Explorations in Economic History, Socioeconomic Inequalities in Death*, Vol. 48/3, pp. 343-56, https://doi.org/10.1016/j.eeh.2011.05.004. [7]

Boarini, R. and M. Mira D'Ercole (2013), "Going beyond GDP: An OECD perspective", *Fiscal Studies*, Vol. 34/3, pp. 289-314, https://doi.org/10.1111/j.1475-5890.2013.12007.x. [2]

Bolt, J. and J. van Zanden (2014), "The Maddison project: Collaborative research on historical mational accounts", *The Economic History Review*, Vol. 67/3, pp. 627-51, https://doi.org/10.1111/1468-0289.12032. [4]

Chakravarty, S. (2003), "A generalized human development index", *Review of Development Economics*, Vol. 7/1, pp. 99-114, https://doi.org/10.1111/1467-9361.00178. [16]

Chen, S. and M. Ravallion (2010), "The developing world is poorer than we thought, but no less successful in the fight against poverty", *The Quarterly Journal of Economics*, Vol. 125/4, pp. 1577-1625, https://doi.org/10.1162/qjec.2010.125.4.1577. [10]

Dollar, D. and A. Kraay (2002), "Growth is good for the poor", *Journal of Economic Growth*, Vol. 7/3, pp. 195-225, https://doi.org/10.1023/A:1020139631000. [20]

Feinstein, C. (1998), "Pessimism perpetuated: Real wages and the standard of living in Britain during and after the industrial revolution", *Journal of Economic History*, Vol. 58/3, pp. 625-58, https://doi.org/10.1017/S0022050700021100. [6]

Høyland, B., K. Moene and F. Willumsen (2012), "The tyranny of international index rankings", *Journal of Development Economics*, Vol. 97/1, pp. 1-14, https://doi.org/10.1016/j.jdeveco.2011.01.007. [19]

Jackman, S. (2009), *Bayesian Analysis for the Social Sciences. Wiley Series in Probability and Statistics*, Wiley, Chichester. [18]

Jones, C. and P. Klenow (2016), "Beyond GDP? Welfare across countries and time", *American Economic Review*, Vol. 106/9, pp. 2426-57, https://doi.org/10.1257/aer.20110236. [25]

Maddison, A. (2001), *The World Economy: A Millennial Perspective*, Development Centre Studies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264189980-en. [3]

OECD (2015), *In It Together: Why Less Inequality Benefits All*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264235120-en. [9]

Piketty, T. (2014), *Capital in the Twenty-First Century*, The Belknap Press of Harvard University Press, Cambridge Massachusetts. [23]

Ravallion, M. (2020), "On measuring global poverty", *Annual Review of Economics*, Vol. 12, pp. 167-188, https://doi.org/10.1146/annurev-economics-081919-022924. [12]

Ravallion, M. (2012), "Mashup indices of development", *The World Bank Research Observer*, Vol. 27/1, pp. 1-32, https://doi.org/10.1093/wbro/lkr009. [14]

Ravallion, M. (2012), “Troubling tradeoffs in the human development index”, *Journal of Development Economics*, Vol. 99/2, pp. 201-9, https://doi.org/10.1016/j.jdeveco.2012.01.003. [13]

Ravallion, M., G. Datt and D. van de Walle (1991), “Quantifying absolute poverty in the developing world”, *Review of Income and Wealth*, Vol. 37/4, pp. 345-61, https://doi.org/10.1111/j.1475-4991.1991.tb00378. [11]

Rijpma, A. (2016), “What Can’t Money Buy? Wellbeing and GDP Since 1820”*, Working Paper, Centre for Global Economic History*, No. 0078, Utrecht University, https://ideas.repec.org/p/ucg/wpaper/0078.html. [17]

Rijpma, A. (2014), “A composite view of well-being since 1820”, in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-17-en. [15]

Scheidel, W. (2017), *The Great Leveler: Violence and the History of Inequality from the Stone Age to the Twenty-First Century. The Princeton Economic History of the Western World*, Princeton University Press, New Jersey. [24]

Steckel, R. and R. Floud (1997), *Health and Welfare During Industrialization. A National Bureau of Economic Research Project Report*, The University of Chicago Press, Chicago, IL. [5]

Stiglitz, J., A. Sen and J. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf. [1]

UNDP (2010), *Human Development Report 2010: The Real Wealth of Nations: Pathways to Human Development*, United Nations Development Programme, New York, http://www.hdr.undp.org/sites/default/files/reports/270/hdr_2010_en_complete_reprint.pdf (accessed on 2 February 2021). [26]

van Leeuwen, B. and J. van Leeuwen-Li (2014), “Education since 1820”, in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-9-en. [21]

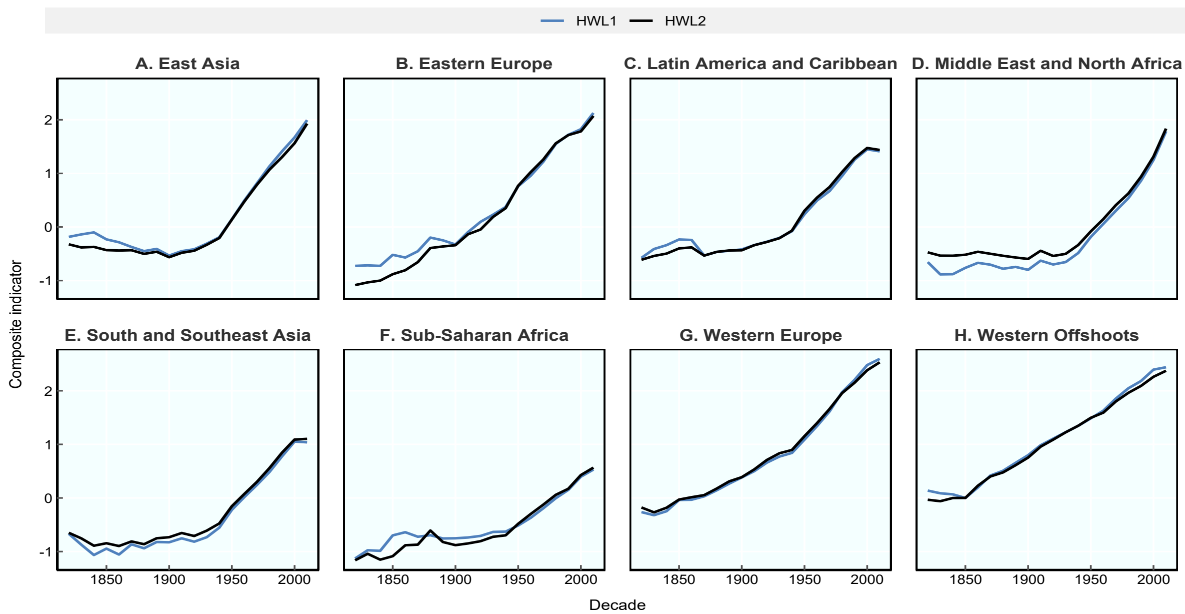
van Zanden, J. et al. (2014), *How Was Life? Global Well-Being Since 1820*, Paris: OECD Publishing, http://doi.org/10.1787/9789264214262-en. [8]

Zijdeman, R. and F. Ribeiro de Silva (2014), “Life expectancy since 1820”, in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-10-en. [22]

Annex 11.A. Supporting material

Annex Figure 11.A.1. Comparison between the expanded composite indicator of average well-being in this chapter and the one in the 2014 volume of *How was Life?*, 1820-2010

By world region



Source: Authors' calculations and Rijpma, A. (2014[15]), "A composite view of well-being since 1820", in *How Was Life?: Global Well-being since 1820*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264214262-17-en.

StatLink https://stat.link/yc1djg

Annex Table 11.A.1. Composite indicator scores and 90 percent confidence intervals in selected countries, 1820-2000

	1820	1870	1910	1950	1970	2000
ARG	-0.4±0.5	-0.5±0.3	-0.0±0.3	0.8±0.3	1.1±0.3	1.6±0.3
AUS	-0.4±0.6	0.3±0.3	1.3±0.4	1.8±0.3	2.2±0.3	2.9±0.3
BRA	-0.8±0.6	-0.7±0.4	-0.7±0.3	0.1±0.3	0.4±0.3	1.5±0.3
CAN	-0.1±0.6	0.5±0.3	1.1±0.3	1.8±0.3	2.3±0.3	2.7±0.3
CHN	-0.5±0.6	-0.5±0.4	-0.5±0.4	-0.3±0.3	0.4±0.3	1.2±0.3
DEU	-0.5±0.5	0.2±0.3	0.7±0.3	1.5±0.3	2.1±0.3	2.7±0.3
EGY	-0.4±0.6	-0.6±0.4	-0.5±0.3	-0.1±0.3	0.1±0.3	1.1±0.3
ESP	-0.2±0.5	-0.3±0.3	0.2±0.3	0.7±0.3	1.3±0.3	2.2±0.3
FRA	-0.4±0.5	0.0±0.3	0.7±0.3	1.4±0.3	1.9±0.3	2.6±0.3
GBR	-0.3±0.3	0.1±0.3	0.8±0.3	1.5±0.3	2.0±0.3	2.9±0.3
IDN	-0.8±0.5	-0.9±0.4	-0.9±0.4	-0.5±0.3	0.0±0.3	1.2±0.3
IND	-0.6±0.6	-0.8±0.3	-0.8±0.3	-0.3±0.3	0.0±0.3	0.8±0.3
ITA	-0.3±0.5	-0.5±0.3	-0.0±0.3	0.9±0.3	1.5±0.3	2.4±0.3
JPN	-0.6±0.6	-0.6±0.3	-0.3±0.3	0.7±0.3	1.7±0.3	2.5±0.3

	1820	1870	1910	1950	1970	2000
KEN				-0.5±0.3	0.0±0.3	0.6±0.3
MEX	-0.9±0.6	-0.7±0.4	-0.6±0.3	0.0±0.3	0.6±0.3	1.5±0.3
NGA				-0.7±0.3	-0.4±0.3	0.3±0.3
NLD	-0.3±0.4	0.1±0.3	0.7±0.3	1.4±0.3	1.9±0.3	2.5±0.3
POL	-0.8±0.6	-0.3±0.5	0.2±0.5	0.5±0.3	1.1±0.3	1.8±0.3
RUS					1.4±0.3	1.7±0.3
SUN			-0.1±0.7	0.7±0.4		
SWE	-0.4±0.5	0.1±0.3	0.7±0.3	1.5±0.3	2.1±0.3	2.5±0.3
THA	-0.8±0.6	-0.9±0.4	-0.7±0.4	-0.2±0.3	0.3±0.3	1.4±0.3
TUR	-0.7±0.5	-0.7±0.3	-0.6±0.3	-0.2±0.3	0.3±0.3	1.2±0.3
USA	0.3±0.6	0.6±0.4	1.1±0.3	1.9±0.3	2.4±0.3	2.9±0.3
ZAF	-1.1±0.5	-0.7±0.4	-0.4±0.4	0.1±0.3	0.5±0.3	1.0±0.3

StatLink https://stat.link/m3y1pr

Annex Table 11.A.2. Composite indicator scores and 90 percent confidence intervals of well-being inequality in selected countries, 1820-2000

	1820	1870	1910	1950	1970	2000
ARG	-0.1±1.1	0.7±0.4	0.1±0.4	-0.8±0.4	-1.0±0.4	-1.4±0.4
AUS	-0.0±1.2	-0.4±0.4	-1.2±0.4	-1.5±0.4	-1.5±0.4	-2.0±0.4
BRA	0.0±1.1	1.2±0.4	0.9±0.4	0.3±0.4	-0.2±0.4	-1.2±0.4
CAN	-0.1±1.2	-1.2±0.4	-1.3±0.4	-1.4±0.4	-1.4±0.4	-1.8±0.4
CHN	-0.1±1.2	0.7±0.4	0.5±0.4	0.1±0.4	-0.7±0.4	-1.3±0.4
DEU	0.7±1.0	-0.8±0.4	-1.1±0.4	-1.4±0.4	-1.7±0.4	-2.0±0.4
EGY	0.1±1.2	1.4±0.4	1.4±0.4	1.0±0.4	0.7±0.4	-0.7±0.4
ESP	0.7±1.0	0.5±0.4	-0.2±0.4	-1.0±0.4	-1.4±0.4	-1.6±0.4
FRA	0.9±0.9	-0.3±0.4	-0.8±0.4	-1.2±0.4	-1.5±0.4	-1.9±0.4
GBR	0.8±1.1	-0.7±0.4	-1.3±0.4	-1.4±0.4	-1.6±0.4	-2.0±0.4
IDN	-0.1±1.2	1.5±0.4	1.4±0.4	0.9±0.4	-0.1±0.4	-1.2±0.4
IND	-0.2±1.2	1.5±0.4	1.4±0.4	1.0±0.4	0.5±0.4	-0.3±0.4
ITA	0.8±1.1	0.6±0.4	-0.4±0.4	-1.2±0.4	-1.3±0.4	-1.7±0.4
JPN	0.0±1.1	0.7±0.4	-0.3±0.4	-1.1±0.4	-1.5±0.4	-1.8±0.4
KEN				0.6±0.4	0.2±0.4	-1.0±0.4
MEX	-0.1±1.2	1.1±0.4	0.7±0.4	-0.2±0.4	-0.8±0.4	-1.3±0.4
NGA				1.0±0.4	0.7±0.4	-0.3±0.4
NLD	0.8±1.0	-0.9±0.4	-1.3±0.4	-1.4±0.4	-1.4±0.4	-1.8±0.4
POL	-0.2±1.2	0.3±0.9	-0.5±0.8	-0.2±0.4	-1.0±0.4	-1.3±0.4
RUS					-1.2±0.4	-1.5±0.4
SUN			-0.2±0.8	-1.1±0.7		
SWE	0.9±0.9	-0.6±0.4	-1.0±0.4	-1.4±0.4	-1.5±0.4	-2.0±0.4
THA	-0.2±1.2	1.4±0.4	1.2±0.4	-0.0±0.4	-0.4±0.4	-1.3±0.4
TUR	0.1±1.2	1.4±0.4	1.3±0.4	0.7±0.4	-0.0±0.4	-0.8±0.4
USA	0.0±1.1	-0.8±0.4	-1.1±0.4	-1.4±0.4	-1.5±0.4	-1.8±0.4
ZAF	0.1±1.1	1.0±0.4	0.8±0.4	0.2±0.4	0.0±0.4	-0.8±0.4

StatLink https://stat.link/tei3qn

Notes

[1] Of the indicators presented here, per capita GDP and wealth are log-linearly interpolated; all other indicators are imputed with linear interpolation. The composite indicator has its own imputation procedure, which is explained in Rijpma (2016[17]).

[2] 75% of the population is covered on average, with lower percentages further back in time. The sub-Saharan Africa and Middle East and North Africa regions have lower coverage.

How Was Life? Volume II

NEW PERSPECTIVES ON WELL-BEING AND GLOBAL INEQUALITY SINCE 1820

How was life in 1820, and how has it changed since then? This question, which was at the core of *How Was Life? Global Well-being since 1820*, published by the OECD in 2014, is addressed by this second volume based on a broader perspective. *How Was Life? New Perspectives on Well-being and Global Inequality since 1820,* presents new estimates of working hours, biodiversity loss, social spending and GDP (accounting for the 2011 round on purchasing power parities) as well as measures of inequalities in wealth, longevity and educational attainment, gender disparities and extreme poverty. A final chapter synthesises the historical evidence included both in the current and previous volume of *How Was Life?* through composite measures of the average well-being performance of each country, and of different within-country inequality measures. As was the case for the previous volume, this book combines both a historical and a global perspective, presenting estimates since 1820 for 25 major countries and 8 world regions. While this evidence sometimes relies on partial and limited evidence, each chapter in this book assesses the quality of the data used and identifies areas for further historical research.

This second volume of *How Was Life?* is the product of collaboration between the OECD and the OECD Development Centre, on one side, and a group of economic historians gathered around the CLIO-INFRA and Maddison projects, on the other. The historical evidence included in the report is organised around dimensions of well-being that mirror those used by the OECD in its report *How's Life?*

PRINT ISBN 978-92-64-80029-8
PDF ISBN 978-92-64-40315-4