

Emmanuelle Marchadour

Thomas Piketty is Professor at the Paris School of Economics.

THE BELKNAP PRESS OF
HARVARD UNIVERSITY PRESS
CAMBRIDGE, MASSACHUSETTS
LONDON, ENGLAND
www.hup.harvard.edu

Jacket design: Graciela Galup

“This book is not only the definitive account of the historical evolution of inequality in advanced economies, it is also a magisterial treatise on capitalism’s inherent dynamics. Piketty ends his book with a ringing call for the global taxation of capital. Whether or not you agree with him on the solution, this book presents a stark challenge for those who would like to save capitalism from itself.”

—Dani Rodrik, Institute for Advanced Study

“A pioneering book that is at once thoughtful, measured, and provocative. It should have a major impact on our discussions of contemporary inequality and its meaning for our democratic institutions and ideals. I can only marvel at Piketty’s discipline and rigor.”

—Rakesh Khurana, Harvard Business School

“A seminal book on the economic and social evolution of the planet. . . . A masterpiece.”

—Emmanuel Todd, *Marianne*

“The book of the season.”

—*Télérama*

“Outstanding . . . A political and theoretical bulldozer.”

—*Mediapart*

“An explosive argument.”

—*Libération*



PIKETTY

CAPITAL

in the
Twenty-First
Century



Belknap
Harvard

CAPITAL

in the Twenty-First Century

THOMAS
PIKETTY

TRANSLATED BY ARTHUR GOLDHAMMER

\$39.95

What are the grand dynamics that drive the accumulation and distribution of capital? Questions about the long-term evolution of inequality, the concentration of wealth, and the prospects for economic growth lie at the heart of political economy. But satisfactory answers have been hard to find for lack of adequate data and clear guiding theories. In *Capital in the Twenty-First Century*, Thomas Piketty analyzes a unique collection of data from twenty countries, ranging as far back as the eighteenth century, to uncover key economic and social patterns. His findings will transform debate and set the agenda for the next generation of thought about wealth and inequality.

Piketty shows that modern economic growth and the diffusion of knowledge have allowed us to avoid inequalities on the apocalyptic scale predicted by Karl Marx. But we have not modified the deep structures of capital and inequality as much as we thought in the optimistic decades following World War II. The main driver of inequality—the tendency of returns on capital to exceed the rate of economic growth—today threatens to generate extreme inequalities that stir discontent and undermine democratic values. But economic trends are not acts of God. Political action has curbed dangerous inequalities in the past, Piketty says, and may do so again.

A work of extraordinary ambition, originality, and rigor, *Capital in the Twenty-First Century* reorients our understanding of economic history and confronts us with sobering lessons for today.

Capital in the Twenty-First Century

—-1
—0
—+1

-1—
0—
+1—

CAPITAL IN THE TWENTY-FIRST CENTURY

Thomas Piketty

Translated by Arthur Goldhammer

The Belknap Press of Harvard University Press

CAMBRIDGE, MASSACHUSETTS

LONDON, ENGLAND

2014

—-1
—0
—+1

Copyright © 2014 by the President and Fellows of Harvard College

All rights reserved

Printed in the United States of America

First published as *Le capital au XXI siècle*,

copyright © 2013 Éditions du Seuil

Design by Dean Bornstein

Library of Congress Cataloging-in-Publication Data

Piketty, Thomas, 1971–

[Capital au XXIe siècle. English]

Capital in the twenty-first century / Thomas Piketty ; translated by Arthur Goldhammer.

pages cm

Translation of the author's *Le capital au XXIe siècle*.

Includes bibliographical references and index.

ISBN 978-0-674-43000-6 (alk. paper)

1. Capital. 2. Income distribution. 3. Wealth. 4. Labor economics.

I. Goldhammer, Arthur, translator. II. Title.

HB501.P43613 2014

332'.041—dc23

2013036024

-1—

0—

+1—

Contents

Acknowledgments · vii

Introduction · 1

Part One: Income and Capital

1. Income and Output · 39
2. Growth: Illusions and Realities · 72

Part Two: The Dynamics of the Capital/Income Ratio

3. The Metamorphoses of Capital · 113
4. From Old Europe to the New World · 140
5. The Capital/Income Ratio over the Long Run · 164
6. The Capital-Labor Split in the Twenty-First Century · 199

Part Three: The Structure of Inequality

7. Inequality and Concentration: Preliminary Bearings · 237
8. Two Worlds · 271
9. Inequality of Labor Income · 304
10. Inequality of Capital Ownership · 336
11. Merit and Inheritance in the Long Run · 377
12. Global Inequality of Wealth in the Twenty-First Century · 430

Part Four: Regulating Capital in the Twenty-First Century

13. A Social State for the Twenty-First Century · 471
14. Rethinking the Progressive Income Tax · 493
15. A Global Tax on Capital · 515
16. The Question of the Public Debt · 540

Conclusion · 571

Notes · 579

Contents in Detail · 657

List of Tables and Illustrations · 665

Index · 671

—-1

—0

—+1

-1—
0—
+1—

Acknowledgments

This book is based on fifteen years of research (1998–2013) devoted essentially to understanding the historical dynamics of wealth and income. Much of this research was done in collaboration with other scholars.

My earlier work on high-income earners in France, *Les hauts revenus en France au 20e siècle* (2001), had the extremely good fortune to win the enthusiastic support of Anthony Atkinson and Emmanuel Saez. Without them, my modest Francocentric project would surely never have achieved the international scope it has today. Tony, who was a model for me during my graduate school days, was the first reader of my historical work on inequality in France and immediately took up the British case as well as a number of other countries. Together, we edited two thick volumes that came out in 2007 and 2010, covering twenty countries in all and constituting the most extensive database available in regard to the historical evolution of income inequality. Emmanuel and I dealt with the US case. We discovered the vertiginous growth of income of the top 1 percent since the 1970s and 1980s, and our work enjoyed a certain influence in US political debate. We also worked together on a number of theoretical papers dealing with the optimal taxation of capital and income. This book owes a great deal to these collaborative efforts.

The book was also deeply influenced by my historical work with Gilles Postel-Vinay and Jean-Laurent Rosenthal on Parisian estate records from the French Revolution to the present. This work helped me to understand in a more intimate and vivid way the significance of wealth and capital and the problems associated with measuring them. Above all, Gilles and Jean-Laurent taught me to appreciate the many similarities, as well as differences, between the structure of property in the period 1900–1910 and the structure of property now.

All of this work is deeply indebted to the doctoral students and young scholars with whom I have been privileged to work over the past fifteen years. Beyond their direct contribution to the research on which this book draws, their enthusiasm and energy fueled the climate of intellectual excitement in which the work matured. I am thinking in particular of Facundo Alvaredo, Laurent Bach, Antoine Bozio, Clément Carbonnier, Fabien Dell, Gabrielle

ACKNOWLEDGMENTS

Fack, Nicolas Frémeaux, Lucie Gadenne, Julien Grenet, Elise Huilery, Camille Landais, Ioana Marinescu, Elodie Morival, Nancy Qian, Dorothée Rouzet, Stefanie Stantcheva, Juliana Londono Velez, Guillaume Saint-Jacques, Christoph Schinke, Aurélie Sotura, Mathieu Valdenaire, and Gabriel Zucman. More specifically, without the efficiency, rigor, and talents of Facundo Alvaredo, the World Top Incomes Database, to which I frequently refer in these pages, would not exist. Without the enthusiasm and insistence of Camille Landais, our collaborative project on “the fiscal revolution” would never have been written. Without the careful attention to detail and impressive capacity for work of Gabriel Zucman, I would never have completed the work on the historical evolution of the capital/income ratio in wealthy countries, which plays a key role in this book.

I also want to thank the institutions that made this project possible, starting with the *École des Hautes Études en Sciences Sociales*, where I have served on the faculty since 2000, as well as the *École Normale Supérieure* and all the other institutions that contributed to the creation of the *École d'économie* of Paris, where I have been a professor since it was founded, and of which I served as founding director from 2005 to 2007. By agreeing to join forces and become minority partners in a project that transcended the sum of their private interests, these institutions helped to create a modest public good, which I hope will continue to contribute to the development of a pluralistic political economy in the twenty-first century.

Finally, thanks to Juliette, Déborah, and Hélène, my three precious daughters, for all the love and strength they give me. And thanks to Julia, who shares my life and is also my best reader. Her influence and support at every stage in the writing of this book have been essential. Without them, I would not have had the energy to see this project through to completion.

Capital in the Twenty-First Century

—-1
—0
—+1

-1—
0—
+1—

Introduction

“Social distinctions can be based only on common utility.”
—Declaration of the Rights of Man and the Citizen, article 1, 1789

The distribution of wealth is one of today’s most widely discussed and controversial issues. But what do we really know about its evolution over the long term? Do the dynamics of private capital accumulation inevitably lead to the concentration of wealth in ever fewer hands, as Karl Marx believed in the nineteenth century? Or do the balancing forces of growth, competition, and technological progress lead in later stages of development to reduced inequality and greater harmony among the classes, as Simon Kuznets thought in the twentieth century? What do we really know about how wealth and income have evolved since the eighteenth century, and what lessons can we derive from that knowledge for the century now under way?

These are the questions I attempt to answer in this book. Let me say at once that the answers contained herein are imperfect and incomplete. But they are based on much more extensive historical and comparative data than were available to previous researchers, data covering three centuries and more than twenty countries, as well as on a new theoretical framework that affords a deeper understanding of the underlying mechanisms. Modern economic growth and the diffusion of knowledge have made it possible to avoid the Marxist apocalypse but have not modified the deep structures of capital and inequality—or in any case not as much as one might have imagined in the optimistic decades following World War II. When the rate of return on capital exceeds the rate of growth of output and income, as it did in the nineteenth century and seems quite likely to do again in the twenty-first, capitalism automatically generates arbitrary and unsustainable inequalities that radically undermine the meritocratic values on which democratic societies are based. There are nevertheless ways democracy can regain control over capitalism and ensure that the general interest takes precedence over private interests, while preserving economic openness and avoiding protectionist and nationalist reactions. The policy recommendations I propose later in the book tend in this

—-1
—0
—+1

direction. They are based on lessons derived from historical experience, of which what follows is essentially a narrative.

A Debate without Data?

Intellectual and political debate about the distribution of wealth has long been based on an abundance of prejudice and a paucity of fact.

To be sure, it would be a mistake to underestimate the importance of the intuitive knowledge that everyone acquires about contemporary wealth and income levels, even in the absence of any theoretical framework or statistical analysis. Film and literature, nineteenth-century novels especially, are full of detailed information about the relative wealth and living standards of different social groups, and especially about the deep structure of inequality, the way it is justified, and its impact on individual lives. Indeed, the novels of Jane Austen and Honoré de Balzac paint striking portraits of the distribution of wealth in Britain and France between 1790 and 1830. Both novelists were intimately acquainted with the hierarchy of wealth in their respective societies. They grasped the hidden contours of wealth and its inevitable implications for the lives of men and women, including their marital strategies and personal hopes and disappointments. These and other novelists depicted the effects of inequality with a verisimilitude and evocative power that no statistical or theoretical analysis can match.

Indeed, the distribution of wealth is too important an issue to be left to economists, sociologists, historians, and philosophers. It is of interest to everyone, and that is a good thing. The concrete, physical reality of inequality is visible to the naked eye and naturally inspires sharp but contradictory political judgments. Peasant and noble, worker and factory owner, waiter and banker: each has his or her own unique vantage point and sees important aspects of how other people live and what relations of power and domination exist between social groups, and these observations shape each person's judgment of what is and is not just. Hence there will always be a fundamentally subjective and psychological dimension to inequality, which inevitably gives rise to political conflict that no purportedly scientific analysis can alleviate. Democracy will never be supplanted by a republic of experts—and that is a very good thing.

Nevertheless, the distribution question also deserves to be studied in a systematic and methodical fashion. Without precisely defined sources, meth-

INTRODUCTION

ods, and concepts, it is possible to see everything and its opposite. Some people believe that inequality is always increasing and that the world is by definition always becoming more unjust. Others believe that inequality is naturally decreasing, or that harmony comes about automatically, and that in any case nothing should be done that might risk disturbing this happy equilibrium. Given this dialogue of the deaf, in which each camp justifies its own intellectual laziness by pointing to the laziness of the other, there is a role for research that is at least systematic and methodical if not fully scientific. Expert analysis will never put an end to the violent political conflict that inequality inevitably instigates. Social scientific research is and always will be tentative and imperfect. It does not claim to transform economics, sociology, and history into exact sciences. But by patiently searching for facts and patterns and calmly analyzing the economic, social, and political mechanisms that might explain them, it can inform democratic debate and focus attention on the right questions. It can help to redefine the terms of debate, unmask certain preconceived or fraudulent notions, and subject all positions to constant critical scrutiny. In my view, this is the role that intellectuals, including social scientists, should play, as citizens like any other but with the good fortune to have more time than others to devote themselves to study (and even to be paid for it—a signal privilege).

There is no escaping the fact, however, that social science research on the distribution of wealth was for a long time based on a relatively limited set of firmly established facts together with a wide variety of purely theoretical speculations. Before turning in greater detail to the sources I tried to assemble in preparation for writing this book, I want to give a quick historical overview of previous thinking about these issues.

Malthus, Young, and the French Revolution

When classical political economy was born in England and France in the late eighteenth and early nineteenth century, the issue of distribution was already one of the key questions. Everyone realized that radical transformations were under way, precipitated by sustained demographic growth—a previously unknown phenomenon—coupled with a rural exodus and the advent of the Industrial Revolution. How would these upheavals affect the distribution of wealth, the social structure, and the political equilibrium of European society?

For Thomas Malthus, who in 1798 published his *Essay on the Principle of Population*, there could be no doubt: the primary threat was overpopulation.¹ Although his sources were thin, he made the best he could of them. One particularly important influence was the travel diary published by Arthur Young, an English agronomist who traveled extensively in France, from Calais to the Pyrenees and from Brittany to Franche-Comté, in 1787–1788, on the eve of the Revolution. Young wrote of the poverty of the French countryside.

His vivid essay was by no means totally inaccurate. France at that time was by far the most populous country in Europe and therefore an ideal place to observe. The kingdom could already boast of a population of 20 million in 1700, compared to only 8 million for Great Britain (and 5 million for England alone). The French population increased steadily throughout the eighteenth century, from the end of Louis XIV's reign to the demise of Louis XVI, and by 1780 was close to 30 million. There is every reason to believe that this unprecedentedly rapid population growth contributed to a stagnation of agricultural wages and an increase in land rents in the decades prior to the explosion of 1789. Although this demographic shift was not the sole cause of the French Revolution, it clearly contributed to the growing unpopularity of the aristocracy and the existing political regime.

Nevertheless, Young's account, published in 1792, also bears the traces of nationalist prejudice and misleading comparison. The great agronomist found the inns in which he stayed thoroughly disagreeable and disliked the manners of the women who waited on him. Although many of his observations were banal and anecdotal, he believed he could derive universal consequences from them. He was mainly worried that the mass poverty he witnessed would lead to political upheaval. In particular, he was convinced that only the English political system, with separate houses of Parliament for aristocrats and commoners and veto power for the nobility, could allow for harmonious and peaceful development led by responsible people. He was convinced that France was headed for ruin when it decided in 1789–1790 to allow both aristocrats and commoners to sit in a single legislative body. It is no exaggeration to say that his whole account was overdetermined by his fear of revolution in France. Whenever one speaks about the distribution of wealth, politics is never very far behind, and it is difficult for anyone to escape contemporary class prejudices and interests.

INTRODUCTION

When Reverend Malthus published his famous *Essay* in 1798, he reached conclusions even more radical than Young's. Like his compatriot, he was very afraid of the new political ideas emanating from France, and to reassure himself that there would be no comparable upheaval in Great Britain he argued that all welfare assistance to the poor must be halted at once and that reproduction by the poor should be severely scrutinized lest the world succumb to overpopulation leading to chaos and misery. It is impossible to understand Malthus's exaggeratedly somber predictions without recognizing the way fear gripped much of the European elite in the 1790s.

Ricardo: The Principle of Scarcity

In retrospect, it is obviously easy to make fun of these prophecies of doom. It is important to realize, however, that the economic and social transformations of the late eighteenth and early nineteenth centuries were objectively quite impressive, not to say traumatic, for those who witnessed them. Indeed, most contemporary observers—and not only Malthus and Young—shared relatively dark or even apocalyptic views of the long-run evolution of the distribution of wealth and class structure of society. This was true in particular of David Ricardo and Karl Marx, who were surely the two most influential economists of the nineteenth century and who both believed that a small social group—landowners for Ricardo, industrial capitalists for Marx—would inevitably claim a steadily increasing share of output and income.²

For Ricardo, who published his *Principles of Political Economy and Taxation* in 1817, the chief concern was the long-term evolution of land prices and land rents. Like Malthus, he had virtually no genuine statistics at his disposal. He nevertheless had intimate knowledge of the capitalism of his time. Born into a family of Jewish financiers with Portuguese roots, he also seems to have had fewer political prejudices than Malthus, Young, or Smith. He was influenced by the Malthusian model but pushed the argument farther. He was above all interested in the following logical paradox. Once both population and output begin to grow steadily, land tends to become increasingly scarce relative to other goods. The law of supply and demand then implies that the price of land will rise continuously, as will the rents paid to landlords. The landlords will therefore claim a growing share of national income, as the share available to the rest of the population decreases, thus upsetting the social

—-1
—0
—+1

equilibrium. For Ricardo, the only logically and politically acceptable answer was to impose a steadily increasing tax on land rents.

This somber prediction proved wrong: land rents did remain high for an extended period, but in the end the value of farm land inexorably declined relative to other forms of wealth as the share of agriculture in national income decreased. Writing in the 1810s, Ricardo had no way of anticipating the importance of technological progress or industrial growth in the years ahead. Like Malthus and Young, he could not imagine that humankind would ever be totally freed from the alimentary imperative.

His insight into the price of land is nevertheless interesting: the “scarcity principle” on which he relied meant that certain prices might rise to very high levels over many decades. This could well be enough to destabilize entire societies. The price system plays a key role in coordinating the activities of millions of individuals—indeed, today, billions of individuals in the new global economy. The problem is that the price system knows neither limits nor morality.

It would be a serious mistake to neglect the importance of the scarcity principle for understanding the global distribution of wealth in the twenty-first century. To convince oneself of this, it is enough to replace the price of farmland in Ricardo’s model by the price of urban real estate in major world capitals, or, alternatively, by the price of oil. In both cases, if the trend over the period 1970–2010 is extrapolated to the period 2010–2050 or 2010–2100, the result is economic, social, and political disequilibria of considerable magnitude, not only between but within countries—disequilibria that inevitably call to mind the Ricardian apocalypse.

To be sure, there exists in principle a quite simple economic mechanism that should restore equilibrium to the process: the mechanism of supply and demand. If the supply of any good is insufficient, and its price is too high, then demand for that good should decrease, which should lead to a decline in its price. In other words, if real estate and oil prices rise, then people should move to the country or take to traveling about by bicycle (or both). Never mind that such adjustments might be unpleasant or complicated; they might also take decades, during which landlords and oil well owners might well accumulate claims on the rest of the population so extensive that they could easily come to own everything that can be owned, including rural real estate and bicycles, once and for all.³ As always, the worst is never certain to arrive.

INTRODUCTION

It is much too soon to warn readers that by 2050 they may be paying rent to the emir of Qatar. I will consider the matter in due course, and my answer will be more nuanced, albeit only moderately reassuring. But it is important for now to understand that the interplay of supply and demand in no way rules out the possibility of a large and lasting divergence in the distribution of wealth linked to extreme changes in certain relative prices. This is the principal implication of Ricardo's scarcity principle. But nothing obliges us to roll the dice.

Marx: The Principle of Infinite Accumulation

By the time Marx published the first volume of *Capital* in 1867, exactly one-half century after the publication of Ricardo's *Principles*, economic and social realities had changed profoundly: the question was no longer whether farmers could feed a growing population or land prices would rise sky high but rather how to understand the dynamics of industrial capitalism, now in full blossom.

The most striking fact of the day was the misery of the industrial proletariat. Despite the growth of the economy, or perhaps in part because of it, and because, as well, of the vast rural exodus owing to both population growth and increasing agricultural productivity, workers crowded into urban slums. The working day was long, and wages were very low. A new urban misery emerged, more visible, more shocking, and in some respects even more extreme than the rural misery of the Old Regime. *Germinal*, *Oliver Twist*, and *Les Misérables* did not spring from the imaginations of their authors, any more than did laws limiting child labor in factories to children older than eight (in France in 1841) or ten in the mines (in Britain in 1842). Dr. Villermé's *Tableau de l'état physique et moral des ouvriers employés dans les manufactures*, published in France in 1840 (leading to the passage of a timid new child labor law in 1841), described the same sordid reality as *The Condition of the Working Class in England*, which Friedrich Engels published in 1845.⁴

In fact, all the historical data at our disposal today indicate that it was not until the second half—or even the final third—of the nineteenth century that a significant rise in the purchasing power of wages occurred. From the first to the sixth decade of the nineteenth century, workers' wages stagnated at very low levels—close or even inferior to the levels of the eighteenth and

—-1
—0
—+1

previous centuries. This long phase of wage stagnation, which we observe in Britain as well as France, stands out all the more because economic growth was accelerating in this period. The capital share of national income—industrial profits, land rents, and building rents—insofar as can be estimated with the imperfect sources available today, increased considerably in both countries in the first half of the nineteenth century.⁵ It would decrease slightly in the final decades of the nineteenth century, as wages partly caught up with growth. The data we have assembled nevertheless reveal no structural decrease in inequality prior to World War I. What we see in the period 1870–1914 is at best a stabilization of inequality at an extremely high level, and in certain respects an endless inegalitarian spiral, marked in particular by increasing concentration of wealth. It is quite difficult to say where this trajectory would have led without the major economic and political shocks initiated by the war. With the aid of historical analysis and a little perspective, we can now see those shocks as the only forces since the Industrial Revolution powerful enough to reduce inequality.

In any case, capital prospered in the 1840s and industrial profits grew, while labor incomes stagnated. This was obvious to everyone, even though in those days aggregate national statistics did not yet exist. It was in this context that the first communist and socialist movements developed. The central argument was simple: What was the good of industrial development, what was the good of all the technological innovations, toil, and population movements if, after half a century of industrial growth, the condition of the masses was still just as miserable as before, and all lawmakers could do was prohibit factory labor by children under the age of eight? The bankruptcy of the existing economic and political system seemed obvious. People therefore wondered about its long-term evolution: what could one say about it?

This was the task Marx set himself. In 1848, on the eve of the “spring of nations” (that is, the revolutions that broke out across Europe that spring), he published *The Communist Manifesto*, a short, hard-hitting text whose first chapter began with the famous words “A specter is haunting Europe—the specter of communism.”⁶ The text ended with the equally famous prediction of revolution: “The development of Modern Industry, therefore, cuts from under its feet the very foundation on which the bourgeoisie produces and appropriates products. What the bourgeoisie therefore produces, above all, are

INTRODUCTION

its own gravediggers. Its fall and the victory of the proletariat are equally inevitable.”

Over the next two decades, Marx labored over the voluminous treatise that would justify this conclusion and propose the first scientific analysis of capitalism and its collapse. This work would remain unfinished: the first volume of *Capital* was published in 1867, but Marx died in 1883 without having completed the two subsequent volumes. His friend Engels published them posthumously after piecing together a text from the sometimes obscure fragments of manuscript Marx had left behind.

Like Ricardo, Marx based his work on an analysis of the internal logical contradictions of the capitalist system. He therefore sought to distinguish himself from both bourgeois economists (who saw the market as a self-regulated system, that is, a system capable of achieving equilibrium on its own without major deviations, in accordance with Adam Smith’s image of “the invisible hand” and Jean-Baptiste Say’s “law” that production creates its own demand), and utopian socialists and Proudhonians, who in Marx’s view were content to denounce the misery of the working class without proposing a truly scientific analysis of the economic processes responsible for it.⁷ In short, Marx took the Ricardian model of the price of capital and the principle of scarcity as the basis of a more thorough analysis of the dynamics of capitalism in a world where capital was primarily industrial (machinery, plants, etc.) rather than landed property, so that in principle there was no limit to the amount of capital that could be accumulated. In fact, his principal conclusion was what one might call the “principle of infinite accumulation,” that is, the inexorable tendency for capital to accumulate and become concentrated in ever fewer hands, with no natural limit to the process. This is the basis of Marx’s prediction of an apocalyptic end to capitalism: either the rate of return on capital would steadily diminish (thereby killing the engine of accumulation and leading to violent conflict among capitalists), or capital’s share of national income would increase indefinitely (which sooner or later would unite the workers in revolt). In either case, no stable socioeconomic or political equilibrium was possible.

Marx’s dark prophecy came no closer to being realized than Ricardo’s. In the last third of the nineteenth century, wages finally began to increase: the improvement in the purchasing power of workers spread everywhere, and this changed the situation radically, even if extreme inequalities persisted and in

some respects continued to increase until World War I. The communist revolution did indeed take place, but in the most backward country in Europe, Russia, where the Industrial Revolution had scarcely begun, whereas the most advanced European countries explored other, social democratic avenues—fortunately for their citizens. Like his predecessors, Marx totally neglected the possibility of durable technological progress and steadily increasing productivity, which is a force that can to some extent serve as a counterweight to the process of accumulation and concentration of private capital. He no doubt lacked the statistical data needed to refine his predictions. He probably suffered as well from having decided on his conclusions in 1848, before embarking on the research needed to justify them. Marx evidently wrote in great political fervor, which at times led him to issue hasty pronouncements from which it was difficult to escape. That is why economic theory needs to be rooted in historical sources that are as complete as possible, and in this respect Marx did not exploit all the possibilities available to him.⁸ What is more, he devoted little thought to the question of how a society in which private capital had been totally abolished would be organized politically and economically—a complex issue if ever there was one, as shown by the tragic totalitarian experiments undertaken in states where private capital was abolished.

Despite these limitations, Marx's analysis remains relevant in several respects. First, he began with an important question (concerning the unprecedented concentration of wealth during the Industrial Revolution) and tried to answer it with the means at his disposal: economists today would do well to take inspiration from his example. Even more important, the principle of infinite accumulation that Marx proposed contains a key insight, as valid for the study of the twenty-first century as it was for the nineteenth and in some respects more worrisome than Ricardo's principle of scarcity. If the rates of population and productivity growth are relatively low, then accumulated wealth naturally takes on considerable importance, especially if it grows to extreme proportions and becomes socially destabilizing. In other words, low growth cannot adequately counterbalance the Marxist principle of infinite accumulation: the resulting equilibrium is not as apocalyptic as the one predicted by Marx but is nevertheless quite disturbing. Accumulation ends at a finite level, but that level may be high enough to be destabilizing. In particular, the very high level of private wealth that has been attained since the 1980s

INTRODUCTION

and 1990s in the wealthy countries of Europe and in Japan, measured in years of national income, directly reflects the Marxian logic.

From Marx to Kuznets, or Apocalypse to Fairy Tale

Turning from the nineteenth-century analyses of Ricardo and Marx to the twentieth-century analyses of Simon Kuznets, we might say that economists' no doubt overly developed taste for apocalyptic predictions gave way to a similarly excessive fondness for fairy tales, or at any rate happy endings. According to Kuznets's theory, income inequality would automatically decrease in advanced phases of capitalist development, regardless of economic policy choices or other differences between countries, until eventually it stabilized at an acceptable level. Proposed in 1955, this was really a theory of the magical postwar years referred to in France as the "Trente Glorieuses," the thirty glorious years from 1945 to 1975.⁹ For Kuznets, it was enough to be patient, and before long growth would benefit everyone. The philosophy of the moment was summed up in a single sentence: "Growth is a rising tide that lifts all boats." A similar optimism can also be seen in Robert Solow's 1956 analysis of the conditions necessary for an economy to achieve a "balanced growth path," that is, a growth trajectory along which all variables—output, incomes, profits, wages, capital, asset prices, and so on—would progress at the same pace, so that every social group would benefit from growth to the same degree, with no major deviations from the norm.¹⁰ Kuznets's position was thus diametrically opposed to the Ricardian and Marxist idea of an inegalitarian spiral and antithetical to the apocalyptic predictions of the nineteenth century.

In order to properly convey the considerable influence that Kuznets's theory enjoyed in the 1980s and 1990s and to a certain extent still enjoys today, it is important to emphasize that it was the first theory of this sort to rely on a formidable statistical apparatus. It was not until the middle of the twentieth century, in fact, that the first historical series of income distribution statistics became available with the publication in 1953 of Kuznets's monumental *Shares of Upper Income Groups in Income and Savings*. Kuznets's series dealt with only one country (the United States) over a period of thirty-five years (1913–48). It was nevertheless a major contribution, which drew on two sources of data totally unavailable to nineteenth-century authors: US federal income tax returns (which did not exist before the creation of the income tax

in 1913) and Kuznets's own estimates of US national income from a few years earlier. This was the very first attempt to measure social inequality on such an ambitious scale.¹¹

It is important to realize that without these two complementary and indispensable datasets, it is simply impossible to measure inequality in the income distribution or to gauge its evolution over time. To be sure, the first attempts to estimate national income in Britain and France date back to the late seventeenth and early eighteenth century, and there would be many more such attempts over the course of the nineteenth century. But these were isolated estimates. It was not until the twentieth century, in the years between the two world wars, that the first yearly series of national income data were developed by economists such as Kuznets and John W. Kendrick in the United States, Arthur Bowley and Colin Clark in Britain, and L. Dugé de Bernonville in France. This type of data allows us to measure a country's total income. In order to gauge the share of high incomes in national income, we also need statements of income. Such information became available when many countries adopted a progressive income tax around the time of World War I (1913 in the United States, 1914 in France, 1909 in Britain, 1922 in India, 1932 in Argentina).¹²

It is crucial to recognize that even where there is no income tax, there are still all sorts of statistics concerning whatever tax basis exists at a given point in time (for example, the distribution of the number of doors and windows by *département* in nineteenth-century France, which is not without interest), but these data tell us nothing about incomes. What is more, before the requirement to declare one's income to the tax authorities was enacted in law, people were often unaware of the amount of their own income. The same is true of the corporate tax and wealth tax. Taxation is not only a way of requiring all citizens to contribute to the financing of public expenditures and projects and to distribute the tax burden as fairly as possible; it is also useful for establishing classifications and promoting knowledge as well as democratic transparency.

In any event, the data that Kuznets collected allowed him to calculate the evolution of the share of each decile, as well as of the upper centiles, of the income hierarchy in total US national income. What did he find? He noted a sharp reduction in income inequality in the United States between 1913 and 1948. More specifically, at the beginning of this period, the upper decile of the

INTRODUCTION

income distribution (that is, the top 10 percent of US earners) claimed 45–50 percent of annual national income. By the late 1940s, the share of the top decile had decreased to roughly 30–35 percent of national income. This decrease of nearly 10 percentage points was considerable: for example, it was equal to half the income of the poorest 50 percent of Americans.¹³ The reduction of inequality was clear and incontrovertible. This was news of considerable importance, and it had an enormous impact on economic debate in the postwar era in both universities and international organizations.

Malthus, Ricardo, Marx, and many others had been talking about inequalities for decades without citing any sources whatsoever or any methods for comparing one era with another or deciding between competing hypotheses. Now, for the first time, objective data were available. Although the information was not perfect, it had the merit of existing. What is more, the work of compilation was extremely well documented: the weighty volume that Kuznets published in 1953 revealed his sources and methods in the most minute detail, so that every calculation could be reproduced. And besides that, Kuznets was the bearer of good news: inequality was shrinking.

The Kuznets Curve: Good News in the Midst of the Cold War

In fact, Kuznets himself was well aware that the compression of high US incomes between 1913 and 1948 was largely accidental. It stemmed in large part from multiple shocks triggered by the Great Depression and World War II and had little to do with any natural or automatic process. In his 1953 work, he analyzed his series in detail and warned readers not to make hasty generalizations. But in December 1954, at the Detroit meeting of the American Economic Association, of which he was president, he offered a far more optimistic interpretation of his results than he had given in 1953. It was this lecture, published in 1955 under the title “Economic Growth and Income Inequality,” that gave rise to the theory of the “Kuznets curve.”

According to this theory, inequality everywhere can be expected to follow a “bell curve.” In other words, it should first increase and then decrease over the course of industrialization and economic development. According to Kuznets, a first phase of naturally increasing inequality associated with the early stages of industrialization, which in the United States meant, broadly speaking, the nineteenth century, would be followed by a phase of sharply

—-1
—0
—+1

decreasing inequality, which in the United States allegedly began in the first half of the twentieth century.

Kuznets's 1955 paper is enlightening. After reminding readers of all the reasons for interpreting the data cautiously and noting the obvious importance of exogenous shocks in the recent reduction of inequality in the United States, Kuznets suggests, almost innocently in passing, that the internal logic of economic development might also yield the same result, quite apart from any policy intervention or external shock. The idea was that inequalities increase in the early phases of industrialization, because only a minority is prepared to benefit from the new wealth that industrialization brings. Later, in more advanced phases of development, inequality automatically decreases as a larger and larger fraction of the population partakes of the fruits of economic growth.¹⁴

The "advanced phase" of industrial development is supposed to have begun toward the end of the nineteenth or the beginning of the twentieth century in the industrialized countries, and the reduction of inequality observed in the United States between 1913 and 1948 could therefore be portrayed as one instance of a more general phenomenon, which should theoretically reproduce itself everywhere, including underdeveloped countries then mired in postcolonial poverty. The data Kuznets had presented in his 1953 book suddenly became a powerful political weapon.¹⁵ He was well aware of the highly speculative nature of his theorizing.¹⁶ Nevertheless, by presenting such an optimistic theory in the context of a "presidential address" to the main professional association of US economists, an audience that was inclined to believe and disseminate the good news delivered by their prestigious leader, he knew that he would wield considerable influence: thus the "Kuznets curve" was born. In order to make sure that everyone understood what was at stake, he took care to remind his listeners that the intent of his optimistic predictions was quite simply to maintain the underdeveloped countries "within the orbit of the free world."¹⁷ In large part, then, the theory of the Kuznets curve was a product of the Cold War.

To avoid any misunderstanding, let me say that Kuznets's work in establishing the first US national accounts data and the first historical series of inequality measures was of the utmost importance, and it is clear from reading his books (as opposed to his papers) that he shared the true scientific ethic. In addition, the high growth rates observed in all the developed coun-

INTRODUCTION

tries in the post–World War II period were a phenomenon of great significance, as was the still more significant fact that all social groups shared in the fruits of growth. It is quite understandable that the Trente Glorieuses fostered a certain degree of optimism and that the apocalyptic predictions of the nineteenth century concerning the distribution of wealth forfeited some of their popularity.

Nevertheless, the magical Kuznets curve theory was formulated in large part for the wrong reasons, and its empirical underpinnings were extremely fragile. The sharp reduction in income inequality that we observe in almost all the rich countries between 1914 and 1945 was due above all to the world wars and the violent economic and political shocks they entailed (especially for people with large fortunes). It had little to do with the tranquil process of intersectoral mobility described by Kuznets.

Putting the Distributional Question Back at the Heart of Economic Analysis

The question is important, and not just for historical reasons. Since the 1970s, income inequality has increased significantly in the rich countries, especially the United States, where the concentration of income in the first decade of the twenty-first century regained—indeed, slightly exceeded—the level attained in the second decade of the previous century. It is therefore crucial to understand clearly why and how inequality decreased in the interim. To be sure, the very rapid growth of poor and emerging countries, especially China, may well prove to be a potent force for reducing inequalities at the global level, just as the growth of the rich countries did during the period 1945–1975. But this process has generated deep anxiety in the emerging countries and even deeper anxiety in the rich countries. Furthermore, the impressive disequilibria observed in recent decades in the financial, oil, and real estate markets have naturally aroused doubts as to the inevitability of the “balanced growth path” described by Solow and Kuznets, according to whom all key economic variables are supposed to move at the same pace. Will the world in 2050 or 2100 be owned by traders, top managers, and the superrich, or will it belong to the oil-producing countries or the Bank of China? Or perhaps it will be owned by the tax havens in which many of these actors will have sought refuge. It would be absurd not to raise the question of who will own what and

—-1
—0
—+1

simply to assume from the outset that growth is naturally “balanced” in the long run.

In a way, we are in the same position at the beginning of the twenty-first century as our forebears were in the early nineteenth century: we are witnessing impressive changes in economies around the world, and it is very difficult to know how extensive they will turn out to be or what the global distribution of wealth, both within and between countries, will look like several decades from now. The economists of the nineteenth century deserve immense credit for placing the distributional question at the heart of economic analysis and for seeking to study long-term trends. Their answers were not always satisfactory, but at least they were asking the right questions. There is no fundamental reason why we should believe that growth is automatically balanced. It is long since past the time when we should have put the question of inequality back at the center of economic analysis and begun asking questions first raised in the nineteenth century. For far too long, economists have neglected the distribution of wealth, partly because of Kuznets’s optimistic conclusions and partly because of the profession’s undue enthusiasm for simplistic mathematical models based on so-called representative agents.¹⁸ If the question of inequality is again to become central, we must begin by gathering as extensive as possible a set of historical data for the purpose of understanding past and present trends. For it is by patiently establishing facts and patterns and then comparing different countries that we can hope to identify the mechanisms at work and gain a clearer idea of the future.

The Sources Used in This Book

This book is based on sources of two main types, which together make it possible to study the historical dynamics of wealth distribution: sources dealing with the inequality and distribution of income, and sources dealing with the distribution of wealth and the relation of wealth to income.

To begin with income: in large part, my work has simply broadened the spatial and temporal limits of Kuznets’s innovative and pioneering work on the evolution of income inequality in the United States between 1913 and 1948. In this way I have been able to put Kuznets’s findings (which are quite accurate) into a wider perspective and thus radically challenge his optimistic view of the relation between economic development and the distribution of wealth.

INTRODUCTION

Oddly, no one has ever systematically pursued Kuznets's work, no doubt in part because the historical and statistical study of tax records falls into a sort of academic no-man's-land, too historical for economists and too economic for historians. That is a pity, because the dynamics of income inequality can only be studied in a long-run perspective, which is possible only if one makes use of tax records.¹⁹

I began by extending Kuznets's methods to France, and I published the results of that study in a book that appeared in 2001.²⁰ I then joined forces with several colleagues—Anthony Atkinson and Emmanuel Saez foremost among them—and with their help was able to expand the coverage to a much wider range of countries. Anthony Atkinson looked at Great Britain and a number of other countries, and together we edited two volumes that appeared in 2007 and 2010, in which we reported the results for some twenty countries throughout the world.²¹ Together with Emmanuel Saez, I extended Kuznets's series for the United States by half a century.²² Saez himself looked at a number of other key countries, such as Canada and Japan. Many other investigators contributed to this joint effort: in particular, Facundo Alvaredo studied Argentina, Spain, and Portugal; Fabien Dell looked at Germany and Switzerland; and Abhijit Banerjee and I investigated the Indian case. With the help of Nancy Qian I was able to work on China. And so on.²³

In each case, we tried to use the same types of sources, the same methods, and the same concepts. Deciles and centiles of high incomes were estimated from tax data based on stated incomes (corrected in various ways to ensure temporal and geographic homogeneity of data and concepts). National income and average income were derived from national accounts, which in some cases had to be fleshed out or extended. Broadly speaking, our data series begin in each country when an income tax was established (generally between 1910 and 1920 but in some countries, such as Japan and Germany, as early as the 1880s and in other countries somewhat later). These series are regularly updated and at this writing extend to the early 2010s.

Ultimately, the World Top Incomes Database (WTID), which is based on the joint work of some thirty researchers around the world, is the largest historical database available concerning the evolution of income inequality; it is the primary source of data for this book.²⁴

The book's second most important source of data, on which I will actually draw first, concerns wealth, including both the distribution of wealth and its

—-1
—0
—+1

relation to income. Wealth also generates income and is therefore important on the income study side of things as well. Indeed, income consists of two components: income from labor (wages, salaries, bonuses, earnings from nonwage labor, and other remuneration statutorily classified as labor related) and income from capital (rent, dividends, interest, profits, capital gains, royalties, and other income derived from the mere fact of owning capital in the form of land, real estate, financial instruments, industrial equipment, etc., again regardless of its precise legal classification). The WTID contains a great deal of information about the evolution of income from capital over the course of the twentieth century. It is nevertheless essential to complete this information by looking at sources directly concerned with wealth. Here I rely on three distinct types of historical data and methodology, each of which is complementary to the others.²⁵

In the first place, just as income tax returns allow us to study changes in income inequality, estate tax returns enable us to study changes in the inequality of wealth.²⁶ This approach was introduced by Robert Lampman in 1962 to study changes in the inequality of wealth in the United States from 1922 to 1956. Later, in 1978, Anthony Atkinson and Alan Harrison studied the British case from 1923 to 1972.²⁷ These results were recently updated and extended to other countries such as France and Sweden. Unfortunately, data are available for fewer countries than in the case of income inequality. In a few cases, however, estate tax data extend back much further in time, often to the beginning of the nineteenth century, because estate taxes predate income taxes. In particular, I have compiled data collected by the French government at various times and, together with Gilles Postel-Vinay and Jean-Laurent Rosenthal, have put together a huge collection of individual estate tax returns, with which it has been possible to establish homogeneous series of data on the concentration of wealth in France since the Revolution.²⁸ This will allow us to see the shocks due to World War I in a much broader context than the series dealing with income inequality (which unfortunately date back only as far as 1910 or so). The work of Jesper Roine and Daniel Waldenström on Swedish historical sources is also instructive.²⁹

The data on wealth and inheritance also enable us to study changes in the relative importance of inherited wealth and savings in the constitution of fortunes and the dynamics of wealth inequality. This work is fairly complete in the case of France, where the very rich historical sources offer a unique

INTRODUCTION

vantage point from which to observe changing inheritance patterns over the long run.³⁰ To one degree or another, my colleagues and I have extended this work to other countries, especially Great Britain, Germany, Sweden, and the United States. These materials play a crucial role in this study, because the significance of inequalities of wealth differs depending on whether those inequalities derive from inherited wealth or savings. In this book, I focus not only on the level of inequality as such but to an even greater extent on the structure of inequality, that is, on the origins of disparities in income and wealth between social groups and on the various systems of economic, social, moral, and political justification that have been invoked to defend or condemn those disparities. Inequality is not necessarily bad in itself: the key question is to decide whether it is justified, whether there are reasons for it.

Last but not least, we can also use data that allow us to measure the total stock of national wealth (including land, other real estate, and industrial and financial capital) over a very long period of time. We can measure this wealth for each country in terms of the number of years of national income required to amass it. This type of global study of the capital/income ratio has its limits. It is always preferable to analyze wealth inequality at the individual level as well, and to gauge the relative importance of inheritance and saving in capital formation. Nevertheless, the capital/income approach can give us an overview of the importance of capital to the society as a whole. Moreover, in some cases (especially Britain and France) it is possible to collect and compare estimates for different periods and thus push the analysis back to the early eighteenth century, which allows us to view the Industrial Revolution in relation to the history of capital. For this I will rely on historical data Gabriel Zucman and I recently collected.³¹ Broadly speaking, this research is merely an extension and generalization of Raymond Goldsmith's work on national balance sheets in the 1970s.³²

Compared with previous works, one reason why this book stands out is that I have made an effort to collect as complete and consistent a set of historical sources as possible in order to study the dynamics of income and wealth distribution over the long run. To that end, I had two advantages over previous authors. First, this work benefits, naturally enough, from a longer historical perspective than its predecessors had (and some long-term changes did not emerge clearly until data for the 2000s became available, largely owing to the fact that certain shocks due to the world wars persisted for a very long time). Second,

—-1
—0
—+1

advances in computer technology have made it much easier to collect and process large amounts of historical data.

Although I have no wish to exaggerate the role of technology in the history of ideas, the purely technical issues are worth a moment's reflection. Objectively speaking, it was far more difficult to deal with large volumes of historical data in Kuznets's time than it is today. This was true to a large extent as recently as the 1980s. In the 1970s, when Alice Hanson Jones collected US estate inventories from the colonial era and Adeline Daumard worked on French estate records from the nineteenth century,³³ they worked mainly by hand, using index cards. When we reread their remarkable work today, or look at François Siminad's work on the evolution of wages in the nineteenth century or Ernest Labrousse's work on the history of prices and incomes in the eighteenth century or Jean Bouvier and François Furet's work on the variability of profits in the nineteenth century, it is clear that these scholars had to overcome major material difficulties in order to compile and process their data.³⁴ In many cases, the technical difficulties absorbed much of their energy, taking precedence over analysis and interpretation, especially since the technical problems imposed strict limits on their ability to make international and temporal comparisons. It is much easier to study the history of the distribution of wealth today than in the past. This book is heavily indebted to recent improvements in the technology of research.³⁵

The Major Results of This Study

What are the major conclusions to which these novel historical sources have led me? The first is that one should be wary of any economic determinism in regard to inequalities of wealth and income. The history of the distribution of wealth has always been deeply political, and it cannot be reduced to purely economic mechanisms. In particular, the reduction of inequality that took place in most developed countries between 1910 and 1950 was above all a consequence of war and of policies adopted to cope with the shocks of war. Similarly, the resurgence of inequality after 1980 is due largely to the political shifts of the past several decades, especially in regard to taxation and finance. The history of inequality is shaped by the way economic, social, and political actors view what is just and what is not, as well as by the relative power of those actors and the collective choices that result. It is the joint product of all relevant actors combined.

INTRODUCTION

The second conclusion, which is the heart of the book, is that the dynamics of wealth distribution reveal powerful mechanisms pushing alternately toward convergence and divergence. Furthermore, there is no natural, spontaneous process to prevent destabilizing, inequalitarian forces from prevailing permanently.

Consider first the mechanisms pushing toward convergence, that is, toward reduction and compression of inequalities. The main forces for convergence are the diffusion of knowledge and investment in training and skills. The law of supply and demand, as well as the mobility of capital and labor, which is a variant of that law, may always tend toward convergence as well, but the influence of this economic law is less powerful than the diffusion of knowledge and skill and is frequently ambiguous or contradictory in its implications. Knowledge and skill diffusion is the key to overall productivity growth as well as the reduction of inequality both within and between countries. We see this at present in the advances made by a number of previously poor countries, led by China. These emergent economies are now in the process of catching up with the advanced ones. By adopting the modes of production of the rich countries and acquiring skills comparable to those found elsewhere, the less developed countries have leapt forward in productivity and increased their national incomes. The technological convergence process may be abetted by open borders for trade, but it is fundamentally a process of the diffusion and sharing of knowledge—the public good par excellence—rather than a market mechanism.

From a strictly theoretical standpoint, other forces pushing toward greater equality might exist. One might, for example, assume that production technologies tend over time to require greater skills on the part of workers, so that labor's share of income will rise as capital's share falls: one might call this the "rising human capital hypothesis." In other words, the progress of technological rationality is supposed to lead automatically to the triumph of human capital over financial capital and real estate, capable managers over fat cat stockholders, and skill over nepotism. Inequalities would thus become more meritocratic and less static (though not necessarily smaller): economic rationality would then in some sense automatically give rise to democratic rationality.

Another optimistic belief, which is current at the moment, is the idea that "class warfare" will automatically give way, owing to the recent increase in life

—-1
—0
—+1

expectancy, to “generational warfare” (which is less divisive because everyone is first young and then old). Put differently, this inescapable biological fact is supposed to imply that the accumulation and distribution of wealth no longer presage an inevitable clash between dynasties of rentiers and dynasties owning nothing but their labor power. The governing logic is rather one of saving over the life cycle: people accumulate wealth when young in order to provide for their old age. Progress in medicine together with improved living conditions has therefore, it is argued, totally transformed the very essence of capital.

Unfortunately, these two optimistic beliefs (the human capital hypothesis and the substitution of generational conflict for class warfare) are largely illusory. Transformations of this sort are both logically possible and to some extent real, but their influence is far less consequential than one might imagine. There is little evidence that labor’s share in national income has increased significantly in a very long time: “nonhuman” capital seems almost as indispensable in the twenty-first century as it was in the eighteenth or nineteenth, and there is no reason why it may not become even more so. Now as in the past, moreover, inequalities of wealth exist primarily within age cohorts, and inherited wealth comes close to being as decisive at the beginning of the twenty-first century as it was in the age of Balzac’s *Père Goriot*. Over a long period of time, the main force in favor of greater equality has been the diffusion of knowledge and skills.

Forces of Convergence, Forces of Divergence

The crucial fact is that no matter how potent a force the diffusion of knowledge and skills may be, especially in promoting convergence between countries, it can nevertheless be thwarted and overwhelmed by powerful forces pushing in the opposite direction, toward greater inequality. It is obvious that lack of adequate investment in training can exclude entire social groups from the benefits of economic growth. Growth can harm some groups while benefiting others (witness the recent displacement of workers in the more advanced economies by workers in China). In short, the principal force for convergence—the diffusion of knowledge—is only partly natural and spontaneous. It also depends in large part on educational policies, access to training and to the acquisition of appropriate skills, and associated institutions.

INTRODUCTION

I will pay particular attention in this study to certain worrisome forces of divergence—particularly worrisome in that they can exist even in a world where there is adequate investment in skills and where all the conditions of “market efficiency” (as economists understand that term) appear to be satisfied. What are these forces of divergence? First, top earners can quickly separate themselves from the rest by a wide margin (although the problem to date remains relatively localized). More important, there is a set of forces of divergence associated with the process of accumulation and concentration of wealth when growth is weak and the return on capital is high. This second process is potentially more destabilizing than the first, and it no doubt represents the principal threat to an equal distribution of wealth over the long run.

To cut straight to the heart of the matter: in Figures I.1 and I.2 I show two basic patterns that I will try to explain in what follows. Each graph represents the importance of one of these divergent processes. Both graphs depict “U-shaped curves,” that is, a period of decreasing inequality followed by one of increasing inequality. One might assume that the realities the two graphs represent are similar. In fact they are not. The phenomena underlying the various curves are quite different and involve distinct economic, social, and political processes. Furthermore, the curve in Figure I.1 represents income inequality in the United States, while the curves in Figure I.2 depict the capital/income ratio in several European countries (Japan, though not shown, is similar). It is not out of the question that the two forces of divergence will ultimately come together in the twenty-first century. This has already happened to some extent and may yet become a global phenomenon, which could lead to levels of inequality never before seen, as well as to a radically new structure of inequality. Thus far, however, these striking patterns reflect two distinct underlying phenomena.

The US curve, shown in Figure I.1, indicates the share of the upper decile of the income hierarchy in US national income from 1910 to 2010. It is nothing more than an extension of the historical series Kuznets established for the period 1913–1948. The top decile claimed as much as 45–50 percent of national income in the 1910s–1920s before dropping to 30–35 percent by the end of the 1940s. Inequality then stabilized at that level from 1950 to 1970. We subsequently see a rapid rise in inequality in the 1980s, until by 2000 we have returned to a level on the order of 45–50 percent of national income. The magnitude of the change is impressive. It is natural to ask how far such a trend might continue.

CAPITAL IN THE TWENTY-FIRST CENTURY

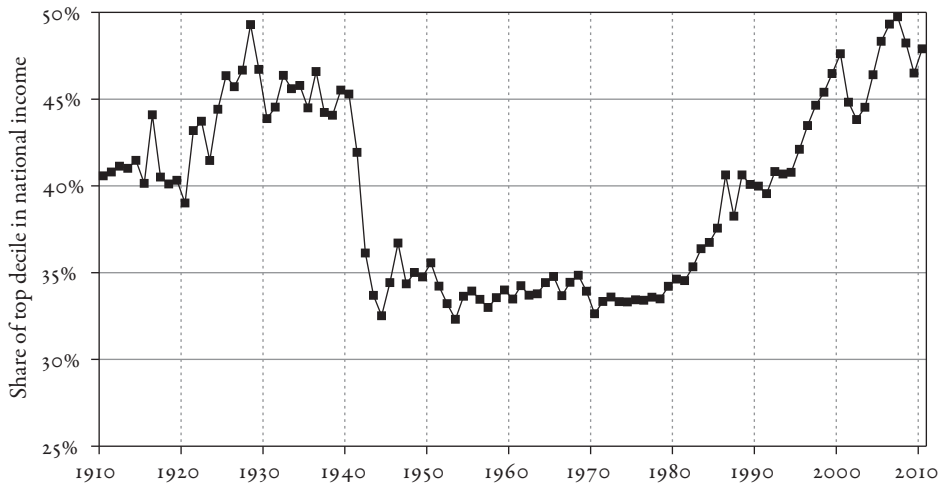


FIGURE I.1. Income inequality in the United States, 1910–2010

The top decile share in US national income dropped from 45–50 percent in the 1910s–1920s to less than 35 percent in the 1950s (this is the fall documented by Kuznets); it then rose from less than 35 percent in the 1970s to 45–50 percent in the 2000s–2010s. Sources and series: see piketty.pse.ens.fr/capital21c.

I will show that this spectacular increase in inequality largely reflects an unprecedented explosion of very elevated incomes from labor, a veritable separation of the top managers of large firms from the rest of the population. One possible explanation of this is that the skills and productivity of these top managers rose suddenly in relation to those of other workers. Another explanation, which to me seems more plausible and turns out to be much more consistent with the evidence, is that these top managers by and large have the power to set their own remuneration, in some cases without limit and in many cases without any clear relation to their individual productivity, which in any case is very difficult to estimate in a large organization. This phenomenon is seen mainly in the United States and to a lesser degree in Britain, and it may be possible to explain it in terms of the history of social and fiscal norms in those two countries over the past century. The tendency is less marked in other wealthy countries (such as Japan, Germany, France, and other continental European states), but the trend is in the same direction. To expect that the phenomenon will attain the same proportions elsewhere as it has done in the United States would be risky until we have subjected it to a

INTRODUCTION

full analysis—which unfortunately is not that simple, given the limits of the available data.

The Fundamental Force for Divergence: $r > g$

The second pattern, represented in Figure I.2, reflects a divergence mechanism that is in some ways simpler and more transparent and no doubt exerts greater influence on the long-run evolution of the wealth distribution. Figure I.2 shows the total value of private wealth (in real estate, financial assets, and professional capital, net of debt) in Britain, France and Germany, expressed in years of national income, for the period 1870–2010. Note, first of all, the very high level of private wealth in Europe in the late nineteenth century: the total amount of private wealth hovered around six or seven years of national income, which is a lot. It then fell sharply in response to the shocks of the period 1914–1945: the capital/income ratio decreased to just 2 or 3. We then observe a steady rise from 1950 on, a rise so sharp that private fortunes in the early twenty-first century seem to be on the verge of returning to five or six years of national income in both Britain and France. (Private wealth in Germany, which started at a lower level, remains lower, but the upward trend is just as clear.)

This “U-shaped curve” reflects an absolutely crucial transformation, which will figure largely in this study. In particular, I will show that the return of high capital/income ratios over the past few decades can be explained in large part by the return to a regime of relatively slow growth. In slowly growing economies, past wealth naturally takes on disproportionate importance, because it takes only a small flow of new savings to increase the stock of wealth steadily and substantially.

If, moreover, the rate of return on capital remains significantly above the growth rate for an extended period of time (which is more likely when the growth rate is low, though not automatic), then the risk of divergence in the distribution of wealth is very high.

This fundamental inequality, which I will write as $r > g$ (where r stands for the average annual rate of return on capital, including profits, dividends, interest, rents, and other income from capital, expressed as a percentage of its total value, and g stands for the rate of growth of the economy, that is, the annual increase in income or output), will play a crucial role in this book. In a sense, it sums up the overall logic of my conclusions.

CAPITAL IN THE TWENTY-FIRST CENTURY

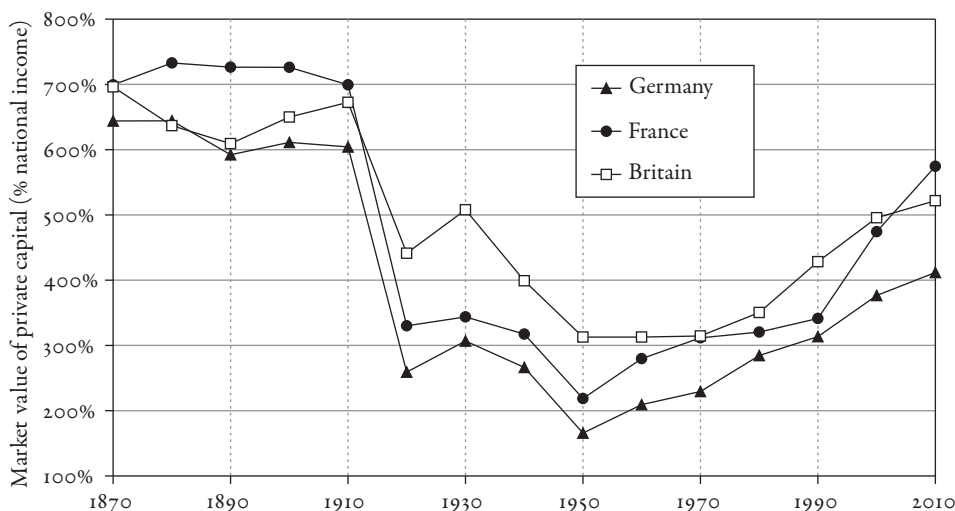


FIGURE 1.2. The capital/income ratio in Europe, 1870–2010

Aggregate private wealth was worth about six to seven years of national income in Europe in 1910, between two and three years in 1950, and between four and six years in 2010.

Sources and series: see piketty.pse.ens.fr/capital21c.

When the rate of return on capital significantly exceeds the growth rate of the economy (as it did through much of history until the nineteenth century and as is likely to be the case again in the twenty-first century), then it logically follows that inherited wealth grows faster than output and income. People with inherited wealth need save only a portion of their income from capital to see that capital grow more quickly than the economy as a whole. Under such conditions, it is almost inevitable that inherited wealth will dominate wealth amassed from a lifetime’s labor by a wide margin, and the concentration of capital will attain extremely high levels—levels potentially incompatible with the meritocratic values and principles of social justice fundamental to modern democratic societies.

What is more, this basic force for divergence can be reinforced by other mechanisms. For instance, the savings rate may increase sharply with wealth.³⁶ Or, even more important, the average effective rate of return on capital may be higher when the individual’s initial capital endowment is higher (as appears to be increasingly common). The fact that the return on capital is un-

-1—
0—
+1—

INTRODUCTION

predictable and arbitrary, so that wealth can be enhanced in a variety of ways, also poses a challenge to the meritocratic model. Finally, all of these factors can be aggravated by the Ricardian scarcity principle: the high price of real estate or petroleum may contribute to structural divergence.

To sum up what has been said thus far: the process by which wealth is accumulated and distributed contains powerful forces pushing toward divergence, or at any rate toward an extremely high level of inequality. Forces of convergence also exist, and in certain countries at certain times, these may prevail, but the forces of divergence can at any point regain the upper hand, as seems to be happening now, at the beginning of the twenty-first century. The likely decrease in the rate of growth of both the population and the economy in coming decades makes this trend all the more worrisome.

My conclusions are less apocalyptic than those implied by Marx's principle of infinite accumulation and perpetual divergence (since Marx's theory implicitly relies on a strict assumption of zero productivity growth over the long run). In the model I propose, divergence is not perpetual and is only one of several possible future directions for the distribution of wealth. But the possibilities are not heartening. Specifically, it is important to note that the fundamental $r > g$ inequality, the main force of divergence in my theory, has nothing to do with any market imperfection. Quite the contrary: the more perfect the capital market (in the economist's sense), the more likely r is to be greater than g . It is possible to imagine public institutions and policies that would counter the effects of this implacable logic: for instance, a progressive global tax on capital. But establishing such institutions and policies would require a considerable degree of international coordination. It is unfortunately likely that actual responses to the problem—including various nationalist responses—will in practice be far more modest and less effective.

The Geographical and Historical Boundaries of This Study

What will the geographical and historical boundaries of this study be? To the extent possible, I will explore the dynamics of the distribution of wealth between and within countries around the world since the eighteenth century. However, the limitations of the available data will often make it necessary to narrow the scope of inquiry rather severely. In regard to the between-country

—-1
—0
—+1

distribution of output and income, the subject of the first part of the book, a global approach is possible from 1700 on (thanks in particular to the national accounts data compiled by Angus Maddison). When it comes to studying the capital/income ratio and capital-labor split in Part Two, the absence of adequate historical data will force me to focus primarily on the wealthy countries and proceed by extrapolation to poor and emerging countries. The examination of the evolution of inequalities of income and wealth, the subject of Part Three, will also be narrowly constrained by the limitations of the available sources. I try to include as many poor and emergent countries as possible, using data from the WTID, which aims to cover five continents as thoroughly as possible. Nevertheless, the long-term trends are far better documented in the rich countries. To put it plainly, this book relies primarily on the historical experience of the leading developed countries: the United States, Japan, Germany, France, and Great Britain.

The British and French cases turn out to be particularly significant, because the most complete long-run historical sources pertain to these two countries. We have multiple estimates of both the magnitude and structure of national wealth for Britain and France as far back as the early eighteenth century. These two countries were also the leading colonial and financial powers in the nineteenth and early twentieth centuries. It is therefore clearly important to study them if we wish to understand the dynamics of the global distribution of wealth since the Industrial Revolution. In particular, their history is indispensable for studying what has been called the “first globalization” of finance and trade (1870–1914), a period that is in many ways similar to the “second globalization,” which has been under way since the 1970s. The period of the first globalization is as fascinating as it was prodigiously inequalitarian. It saw the invention of the electric light as well as the heyday of the ocean liner (the *Titanic* sailed in 1912), the advent of film and radio, and the rise of the automobile and international investment. Note, for example, that it was not until the coming of the twenty-first century that the wealthy countries regained the same level of stock-market capitalization relative to GDP that Paris and London achieved in the early 1900s. This comparison is quite instructive for understanding today’s world.

Some readers will no doubt be surprised that I accord special importance to the study of the French case and may suspect me of nationalism. I should

INTRODUCTION

therefore justify my decision. One reason for my choice has to do with sources. The French Revolution did not create a just or ideal society, but it did make it possible to observe the structure of wealth in unprecedented detail. The system established in the 1790s for recording wealth in land, buildings, and financial assets was astonishingly modern and comprehensive for its time. The Revolution is the reason why French estate records are probably the richest in the world over the long run.

My second reason is that because France was the first country to experience the demographic transition, it is in some respects a good place to observe what awaits the rest of the planet. Although the country's population has increased over the past two centuries, the rate of increase has been relatively low. The population of the country was roughly 30 million at the time of the Revolution, and it is slightly more than 60 million today. It is the same country, with a population whose order of magnitude has not changed. By contrast, the population of the United States at the time of the Declaration of Independence was barely 3 million. By 1900 it was 100 million, and today it is above 300 million. When a country goes from a population of 3 million to a population of 300 million (to say nothing of the radical increase in territory owing to westward expansion in the nineteenth century), it is clearly no longer the same country.

The dynamics and structure of inequality look very different in a country whose population increases by a factor of 100 compared with a country whose population merely doubles. In particular, the inheritance factor is much less important in the former than in the latter. It has been the demographic growth of the New World that has ensured that inherited wealth has always played a smaller role in the United States than in Europe. This factor also explains why the structure of inequality in the United States has always been so peculiar, and the same can be said of US representations of inequality and social class. But it also suggests that the US case is in some sense not generalizable (because it is unlikely that the population of the world will increase a hundredfold over the next two centuries) and that the French case is more typical and more pertinent for understanding the future. I am convinced that detailed analysis of the French case, and more generally of the various historical trajectories observed in other developed countries in Europe, Japan, North America, and Oceania, can tell us a great deal about the future dynamics of global wealth, including such emergent economies as China, Brazil, and

India, where demographic and economic growth will undoubtedly slow in the future (as they have done already).

Finally, the French case is interesting because the French Revolution—the “bourgeois” revolution par excellence—quickly established an ideal of legal equality in relation to the market. It is interesting to look at how this ideal affected the dynamics of wealth distribution. Although the English Revolution of 1688 established modern parliamentarism, it left standing a royal dynasty, primogeniture on landed estates (ended only in the 1920s), and political privileges for the hereditary nobility (reform of the House of Lords is still under discussion, a bit late in the day). Although the American Revolution established the republican principle, it allowed slavery to continue for nearly a century and legal racial discrimination for nearly two centuries. The race question still has a disproportionate influence on the social question in the United States today. In a way, the French Revolution of 1789 was more ambitious. It abolished all legal privileges and sought to create a political and social order based entirely on equality of rights and opportunities. The Civil Code guaranteed absolute equality before the laws of property as well as freedom of contract (for men, at any rate). In the late nineteenth century, conservative French economists such as Paul Leroy-Beaulieu often used this argument to explain why republican France, a nation of “small property owners” made egalitarian by the Revolution, had no need of a progressive or confiscatory income tax or estate tax, in contrast to aristocratic and monarchical Britain. The data show, however, that the concentration of wealth was as large at that time in France as in Britain, which clearly demonstrates that equality of rights in the marketplace cannot ensure equality of rights tout court. Here again, the French experience is quite relevant to today’s world, where many commentators continue to believe, as Leroy-Beaulieu did a little more than a century ago, that ever more fully guaranteed property rights, ever freer markets, and ever “purer and more perfect” competition are enough to ensure a just, prosperous, and harmonious society. Unfortunately, the task is more complex.

The Theoretical and Conceptual Framework

Before proceeding, it may be useful to say a little more about the theoretical and conceptual framework of this research as well as the intellectual itinerary that led me to write this book.

INTRODUCTION

I belong to a generation that turned eighteen in 1989, which was not only the bicentennial of the French Revolution but also the year when the Berlin Wall fell. I belong to a generation that came of age listening to news of the collapse of the Communist dictatorships and never felt the slightest affection or nostalgia for those regimes or for the Soviet Union. I was vaccinated for life against the conventional but lazy rhetoric of anticapitalism, some of which simply ignored the historic failure of Communism and much of which turned its back on the intellectual means necessary to push beyond it. I have no interest in denouncing inequality or capitalism per se—especially since social inequalities are not in themselves a problem as long as they are justified, that is, “founded only upon common utility,” as article 1 of the 1789 Declaration of the Rights of Man and the Citizen proclaims. (Although this definition of social justice is imprecise but seductive, it is rooted in history. Let us accept it for now. I will return to this point later on.) By contrast, I am interested in contributing, however modestly, to the debate about the best way to organize society and the most appropriate institutions and policies to achieve a just social order. Furthermore, I would like to see justice achieved effectively and efficiently under the rule of law, which should apply equally to all and derive from universally understood statutes subject to democratic debate.

I should perhaps add that I experienced the American dream at the age of twenty-two, when I was hired by a university near Boston just after finishing my doctorate. This experience proved to be decisive in more ways than one. It was the first time I had set foot in the United States, and it felt good to have my work recognized so quickly. Here was a country that knew how to attract immigrants when it wanted to! Yet I also realized quite soon that I wanted to return to France and Europe, which I did when I was twenty-five. Since then, I have not left Paris, except for a few brief trips. One important reason for my choice has a direct bearing on this book: I did not find the work of US economists entirely convincing. To be sure, they were all very intelligent, and I still have many friends from that period of my life. But something strange happened: I was only too aware of the fact that I knew nothing at all about the world’s economic problems. My thesis consisted of several relatively abstract mathematical theorems. Yet the profession liked my work. I quickly realized that there had been no significant effort to collect historical data on the dynamics of inequality since Kuznets, yet the profession continued to churn out purely theoretical results without even knowing what facts needed to be

—-1
—0
—+1

explained. And it expected me to do the same. When I returned to France, I set out to collect the missing data.

To put it bluntly, the discipline of economics has yet to get over its childish passion for mathematics and for purely theoretical and often highly ideological speculation, at the expense of historical research and collaboration with the other social sciences. Economists are all too often preoccupied with petty mathematical problems of interest only to themselves. This obsession with mathematics is an easy way of acquiring the appearance of scientificity without having to answer the far more complex questions posed by the world we live in. There is one great advantage to being an academic economist in France: here, economists are not highly respected in the academic and intellectual world or by political and financial elites. Hence they must set aside their contempt for other disciplines and their absurd claim to greater scientific legitimacy, despite the fact that they know almost nothing about anything. This, in any case, is the charm of the discipline and of the social sciences in general: one starts from square one, so that there is some hope of making major progress. In France, I believe, economists are slightly more interested in persuading historians and sociologists, as well as people outside the academic world, that what they are doing is interesting (although they are not always successful). My dream when I was teaching in Boston was to teach at the *École des Hautes Études en Sciences Sociales*, whose faculty has included such leading lights as Lucien Febvre, Fernand Braudel, Claude Lévi-Strauss, Pierre Bourdieu, Françoise Héritier, and Maurice Godelier, to name a few. Dare I admit this, at the risk of seeming chauvinistic in my view of the social sciences? I probably admire these scholars more than Robert Solow or even Simon Kuznets, even though I regret the fact that the social sciences have largely lost interest in the distribution of wealth and questions of social class since the 1970s. Before that, statistics about income, wages, prices, and wealth played an important part in historical and sociological research. In any case, I hope that both professional social scientists and amateurs of all fields will find something of interest in this book, starting with those who claim to “know nothing about economics” but who nevertheless have very strong opinions about inequality of income and wealth, as is only natural.

The truth is that economics should never have sought to divorce itself from the other social sciences and can advance only in conjunction with them. The social sciences collectively know too little to waste time on foolish disci-

INTRODUCTION

plinary squabbles. If we are to progress in our understanding of the historical dynamics of the wealth distribution and the structure of social classes, we must obviously take a pragmatic approach and avail ourselves of the methods of historians, sociologists, and political scientists as well as economists. We must start with fundamental questions and try to answer them. Disciplinary disputes and turf wars are of little or no importance. In my mind, this book is as much a work of history as of economics.

As I explained earlier, I began this work by collecting sources and establishing historical time series pertaining to the distribution of income and wealth. As the book proceeds, I sometimes appeal to theory and to abstract models and concepts, but I try to do so sparingly, and only to the extent that theory enhances our understanding of the changes we observe. For example, income, capital, the economic growth rate, and the rate of return on capital are abstract concepts—theoretical constructs rather than mathematical certainties. Yet I will show that these concepts allow us to analyze historical reality in interesting ways, provided that we remain clear-eyed and critical about the limited precision with which we can measure these things. I will also use a few equations, such as $\alpha = r \times \beta$ (which says that the share of capital in national income is equal to the product of the return on capital and the capital/income ratio), or $\beta = s/g$ (which says that the capital/income ratio is equal in the long run to the savings rate divided by the growth rate). I ask readers not well versed in mathematics to be patient and not immediately close the book: these are elementary equations, which can be explained in a simple, intuitive way and can be understood without any specialized technical knowledge. Above all, I try to show that this minimal theoretical framework is sufficient to give a clear account of what everyone will recognize as important historical developments.

Outline of the Book

The remainder of the book consists of sixteen chapters divided into four parts. Part One, titled “Income and Capital,” contains two chapters and introduces basic ideas that are used repeatedly in the remainder of the book. Specifically, Chapter 1 presents the concepts of national income, capital, and the capital/income ratio and then describes in broad brushstrokes how the global distribution of income and output has evolved. Chapter 2 gives a more

—-1
—0
—+1

detailed analysis of how the growth rates of population and output have evolved since the Industrial Revolution. This first part of the book contains nothing really new, and the reader familiar with these ideas and with the history of global growth since the eighteenth century may wish to skip directly to Part Two.

The purpose of Part Two, titled “The Dynamics of the Capital/Income Ratio,” which consists of four chapters, is to examine the prospects for the long-run evolution of the capital/income ratio and the global division of national income between labor and capital in the twenty-first century. Chapter 3 looks at the metamorphoses of capital since the eighteenth century, starting with the British and French cases, about which we possess the most data over the long run. Chapter 4 introduces the German and US cases. Chapters 5 and 6 extend the geographical range of the analysis to the entire planet, insofar as the sources allow, and seek to draw the lessons from all of these historical experiences that can enable us to anticipate the possible evolution of the capital/income ratio and the relative shares of capital and labor in the decades to come.

Part Three, titled “The Structure of Inequality,” consists of six chapters. Chapter 7 familiarizes the reader with the orders of magnitude of inequality attained in practice by the distribution of income from labor on the one hand and of capital ownership and income from capital on the other. Chapter 8 then analyzes the historical dynamics of these inequalities, starting with a comparison of France and the United States. Chapters 9 and 10 extend the analysis to all the countries for which we have historical data (in the WTID), looking separately at inequalities related to labor and capital, respectively. Chapter 11 studies the changing importance of inherited wealth over the long run. Finally, Chapter 12 looks at the prospects for the global distribution of wealth over the first few decades of the twenty-first century.

The purpose of Part Four, titled “Regulating Capital in the Twenty-First Century” and consisting of four chapters, is to draw normative and policy lessons from the previous three parts, whose purpose is primarily to establish the facts and understand the reasons for the observed changes. Chapter 13 examines what a “social state” suited to present conditions might look like. Chapter 14 proposes a rethinking of the progressive income tax based on past experience and recent trends. Chapter 15 describes what a progressive tax on capital adapted to twenty-first century conditions might look like and compares this idealized tool to other types of regulation that might emerge from

INTRODUCTION

the political process, ranging from a wealth tax in Europe to capital controls in China, immigration reform in the United States, and revival of protectionism in many countries. Chapter 16 deals with the pressing question of public debt and the related issue of the optimal accumulation of public capital at a time when natural capital may be deteriorating.

One final word. It would have been quite presumptuous in 1913 to publish a book called “Capital in the Twentieth Century.” I beg the reader’s indulgence for giving the title *Capital in the Twenty-First Century* to this book, which appeared in French in 2013 and in English in 2014. I am only too well aware of my total inability to predict what form capital will take in 2063 or 2113. As I already noted, and as I will frequently show in what follows, the history of income and wealth is always deeply political, chaotic, and unpredictable. How this history plays out depends on how societies view inequalities and what kinds of policies and institutions they adopt to measure and transform them. No one can foresee how these things will change in the decades to come. The lessons of history are nevertheless useful, because they help us to see a little more clearly what kinds of choices we will face in the coming century and what sorts of dynamics will be at work. The sole purpose of the book, which logically speaking should have been entitled “Capital at the Dawn of the Twenty-First Century,” is to draw from the past a few modest keys to the future. Since history always invents its own pathways, the actual usefulness of these lessons from the past remains to be seen. I offer them to readers without presuming to know their full import.

-1—
0—
+1—

PART ONE

INCOME AND CAPITAL

—-1
—0
—+1

-1—
0—
+1—

Income and Output

On August 16, 2012, the South African police intervened in a labor conflict between workers at the Marikana platinum mine near Johannesburg and the mine's owners: the stockholders of Lonmin, Inc., based in London. Police fired on the strikers with live ammunition. Thirty-four miners were killed.¹ As often in such strikes, the conflict primarily concerned wages: the miners had asked for a doubling of their wage from 500 to 1,000 euros a month. After the tragic loss of life, the company finally proposed a monthly raise of 75 euros.²

This episode reminds us, if we needed reminding, that the question of what share of output should go to wages and what share to profits—in other words, how should the income from production be divided between labor and capital?—has always been at the heart of distributional conflict. In traditional societies, the basis of social inequality and most common cause of rebellion was the conflict of interest between landlord and peasant, between those who owned land and those who cultivated it with their labor, those who received land rents and those who paid them. The Industrial Revolution exacerbated the conflict between capital and labor, perhaps because production became more capital intensive than in the past (making use of machinery and exploiting natural resources more than ever before) and perhaps, too, because hopes for a more equitable distribution of income and a more democratic social order were dashed. I will come back to this point.

The Marikana tragedy calls to mind earlier instances of violence. At Haymarket Square in Chicago on May 1, 1886, and then at Fourmies, in northern France, on May 1, 1891, police fired on workers striking for higher wages. Does this kind of violent clash between labor and capital belong to the past, or will it be an integral part of twenty-first-century history?

The first two parts of this book focus on the respective shares of global income going to labor and capital and on how those shares have changed since the eighteenth century. I will temporarily set aside the issue of income inequality between workers (for example, between an ordinary worker, an engineer,

—-1
—0
—+1

and a plant manager) and between capitalists (for example, between small, medium, and large stockholders or landlords) until Part Three. Clearly, each of these two dimensions of the distribution of wealth—the “factorial” distribution in which labor and capital are treated as “factors of production,” viewed in the abstract as homogeneous entities, and the “individual” distribution, which takes account of inequalities of income from labor and capital at the individual level—is in practice fundamentally important. It is impossible to achieve a satisfactory understanding of the distributional problem without analyzing both.³

In any case, the Marikana miners were striking not only against what they took to be Lonmin’s excessive profits but also against the apparently fabulous salary awarded to the mine’s manager and the difference between his compensation and theirs.⁴ Indeed, if capital ownership were equally distributed and each worker received an equal share of profits in addition to his or her wages, virtually no one would be interested in the division of earnings between profits and wages. If the capital-labor split gives rise to so many conflicts, it is due first and foremost to the extreme concentration of the ownership of capital. Inequality of wealth—and of the consequent income from capital—is in fact always much greater than inequality of income from labor. I will analyze this phenomenon and its causes in Part Three. For now, I will take the inequality of income from labor and capital as given and focus on the global division of national income between capital and labor.

To be clear, my purpose here is not to plead the case of workers against owners but rather to gain as clear as possible a view of reality. Symbolically, the inequality of capital and labor is an issue that arouses strong emotions. It clashes with widely held ideas of what is and is not just, and it is hardly surprising if this sometimes leads to physical violence. For those who own nothing but their labor power and who often live in humble conditions (not to say wretched conditions in the case of eighteenth-century peasants or the Marikana miners), it is difficult to accept that the owners of capital—some of whom have inherited at least part of their wealth—are able to appropriate so much of the wealth produced by their labor. Capital’s share can be quite large: often as much as one-quarter of total output and sometimes as high as one-half in capital-intensive sectors such as mining, or even more where local monopolies allow the owners of capital to demand an even larger share.

Of course, everyone can also understand that if all the company’s earnings from its output went to paying wages and nothing to profits, it would proba-

bly be difficult to attract the capital needed to finance new investments, at least as our economies are currently organized (to be sure, one can imagine other forms of organization). Furthermore, it is not necessarily just to deny any remuneration to those who choose to save more than others—assuming, of course, that differences in saving are an important reason for the inequality of wealth. Bear in mind, too, that a portion of what is called “the income of capital” may be remuneration for “entrepreneurial” labor, and this should no doubt be treated as we treat other forms of labor. This classic argument deserves closer scrutiny. Taking all these elements into account, what is the “right” split between capital and labor? Can we be sure that an economy based on the “free market” and private property always and everywhere leads to an optimal division, as if by magic? In an ideal society, how would one arrange the division between capital and labor? How should one think about the problem?

The Capital-Labor Split in the Long Run: Not So Stable

If this study is to make even modest progress on these questions and at least clarify the terms of a debate that appears to be endless, it will be useful to begin by establishing some facts as accurately and carefully as possible. What exactly do we know about the evolution of the capital-labor split since the eighteenth century? For a long time, the idea accepted by most economists and uncritically repeated in textbooks was that the relative shares of labor and capital in national income were quite stable over the long run, with the generally accepted figure being two-thirds for labor and one-third for capital.⁵ Today, with the advantage of greater historical perspective and newly available data, it is clear that the reality was quite a bit more complex.

For one thing, the capital-labor split varied widely over the course of the twentieth century. The changes observed in the nineteenth century, which I touched on in the Introduction (an increase in the capital share in the first half of the century, followed by a slight decrease and then a period of stability), seem mild by comparison. Briefly, the shocks that buffeted the economy in the period 1914–1945—World War I, the Bolshevik Revolution of 1917, the Great Depression, World War II, and the consequent advent of new regulatory and tax policies along with controls on capital—reduced capital’s share of income to historically low levels in the 1950s. Very soon, however, capital

—-1
—0
—+1

began to reconstitute itself. The growth of capital's share accelerated with the victories of Margaret Thatcher in England in 1979 and Ronald Reagan in the United States in 1980, marking the beginning of a conservative revolution. Then came the collapse of the Soviet bloc in 1989, followed by financial globalization and deregulation in the 1990s. All of these events marked a political turn in the opposite direction from that observed in the first half of the twentieth century. By 2010, and despite the crisis that began in 2007–2008, capital was prospering as it had not done since 1913. Not all of the consequences of capital's renewed prosperity were negative; to some extent it was a natural and desirable development. But it has changed the way we look at the capital-labor split since the beginning of the twenty-first century, as well as our view of changes likely to occur in the decades to come.

Furthermore, if we look beyond the twentieth century and adopt a very long-term view, the idea of a stable capital-labor split must somehow deal with the fact that the nature of capital itself has changed radically (from land and other real estate in the eighteenth century to industrial and financial capital in the twenty-first century). There is also the idea, widespread among economists, that modern economic growth depends largely on the rise of "human capital." At first glance, this would seem to imply that labor should claim a growing share of national income. And one does indeed find that there may be a tendency for labor's share to increase over the very long run, but the gains are relatively modest: capital's share (excluding human capital) in the early decades of the twenty-first century is only slightly smaller than it was at the beginning of the nineteenth century. The importance of capital in the wealthy countries today is primarily due to a slowing of both demographic growth and productivity growth, coupled with political regimes that objectively favor private capital.

The most fruitful way to understand these changes is to analyze the evolution of the capital/income ratio (that is, the ratio of the total stock of capital to the annual flow of income) rather than focus exclusively on the capital-labor split (that is, the share of income going to capital and labor, respectively). In the past, scholars have mainly studied the latter, largely owing to the lack of adequate data to do anything else.

Before presenting my results in detail, it is best to proceed by stages. The purpose of Part One of this book is to introduce certain basic notions. In the remainder of this chapter, I will begin by presenting the concepts of domestic product and national income, capital and labor, and the capital/income ratio.

Then I will look at how the global distribution of income has changed since the Industrial Revolution. In Chapter 2, I will analyze the general evolution of growth rates over time. This will play a central role in the subsequent analysis.

With these preliminaries out of the way, Part Two takes up the dynamics of the capital/income ratio and the capital-labor split, once again proceeding by stages. Chapter 3 will look at changes in the composition of capital and the capital/income ratio since the eighteenth century, beginning with Britain and France, about which we have the best long-run data. Chapter 4 introduces the German case and above all looks at the United States, which serves as a useful complement to the European prism. Finally, Chapters 5 and 6 attempt to extend the analysis to all the rich countries of the world and, insofar as possible, to the entire planet. I also attempt to draw conclusions relevant to the global dynamics of the capital/income ratio and capital-labor split in the twenty-first century.

The Idea of National Income

It will be useful to begin with the concept of “national income,” to which I will frequently refer in what follows. National income is defined as the sum of all income available to the residents of a given country in a given year, regardless of the legal classification of that income.

National income is closely related to the idea of GDP, which comes up often in public debate. There are, however, two important differences between GDP and national income. GDP measures the total of goods and services produced in a given year within the borders of a given country. In order to calculate national income, one must first subtract from GDP the depreciation of the capital that made this production possible: in other words, one must deduct wear and tear on buildings, infrastructure, machinery, vehicles, computers, and other items during the year in question. This depreciation is substantial, today on the order of 10 percent of GDP in most countries, and it does not correspond to anyone’s income: before wages are distributed to workers or dividends to stockholders, and before genuinely new investments are made, worn-out capital must be replaced or repaired. If this is not done, wealth is lost, resulting in negative income for the owners. When depreciation is subtracted from GDP, one obtains the “net domestic product,” which I will refer

—-1
—0
—+1

to more simply as “domestic output” or “domestic production,” which is typically 90 percent of GDP.

Then one must add net income received from abroad (or subtract net income paid to foreigners, depending on each country’s situation). For example, a country whose firms and other capital assets are owned by foreigners may well have a high domestic product but a much lower national income, once profits and rents flowing abroad are deducted from the total. Conversely, a country that owns a large portion of the capital of other countries may enjoy a national income much higher than its domestic product.

Later I will give examples of both of these situations, drawn from the history of capitalism as well as from today’s world. I should say at once that this type of international inequality can give rise to great political tension. It is not an insignificant thing when one country works for another and pays out a substantial share of its output as dividends and rent to foreigners over a long period of time. In many cases, such a system can survive (to a point) only if sustained by relations of political domination, as was the case in the colonial era, when Europe effectively owned much of the rest of the world. A key question of this research is the following: Under what conditions is this type of situation likely to recur in the twenty-first century, possibly in some novel geographic configuration? For example, Europe, rather than being the owner, may find itself owned. Such fears are currently widespread in the Old World—perhaps too widespread. We shall see.

At this stage, suffice it to say that most countries, whether wealthy or emergent, are currently in much more balanced situations than one sometimes imagines. In France as in the United States, Germany as well as Great Britain, China as well as Brazil, and Japan as well as Italy, national income is within 1 or 2 percent of domestic product. In all these countries, in other words, the inflow of profits, interest, dividends, rent, and so on is more or less balanced by a comparable outflow. In wealthy countries, net income from abroad is generally slightly positive. To a first approximation, the residents of these countries own as much in foreign real estate and financial instruments as foreigners own of theirs. Contrary to a tenacious myth, France is not owned by California pension funds or the Bank of China, any more than the United States belongs to Japanese and German investors. The fear of getting into such a predicament is so strong today that fantasy often outstrips reality. The reality is that inequality with respect to capital is a far greater domestic issue than it is an international one. Inequality in the ownership of capital brings the rich and poor within each country into

INCOME AND OUTPUT

conflict with one another far more than it pits one country against another. This has not always been the case, however, and it is perfectly legitimate to ask whether our future may not look more like our past, particularly since certain countries—Japan, Germany, the oil-exporting countries, and to a lesser degree China—have in recent years accumulated substantial claims on the rest of the world (though by no means as large as the record claims of the colonial era). Furthermore, the very substantial increase in cross-ownership, in which various countries own substantial shares of one another, can give rise to a legitimate sense of dispossession, even when net asset positions are close to zero.

To sum up, a country's national income may be greater or smaller than its domestic product, depending on whether net income from abroad is positive or negative.

National income = domestic output + net income from abroad⁶

At the global level, income received from abroad and paid abroad must balance, so that income is by definition equal to output:

Global income = global output⁷

This equality between two annual flows, income and output, is an accounting identity, yet it reflects an important reality. In any given year, it is impossible for total income to exceed the amount of new wealth that is produced (globally speaking; a single country may of course borrow from abroad). Conversely, all production must be distributed as income in one form or another, to either labor or capital: whether as wages, salaries, honoraria, bonuses, and so on (that is, as payments to workers and others who contributed labor to the process of production) or else as profits, dividends, interest, rents, royalties, and so on (that is, as payments to the owners of capital used in the process of production).

What Is Capital?

To recapitulate: regardless of whether we are looking at the accounts of a company, a nation, or the global economy, the associated output and income can be decomposed as the sum of income to capital and income to labor:

National income = capital income + labor income

—-1
—0
—+1

But what is capital? What are its limits? What forms does it take? How has its composition changed over time? This question, central to this investigation, will be examined in greater detail in subsequent chapters. For now it will suffice to make the following points:

First, throughout this book, when I speak of “capital” without further qualification, I always exclude what economists often call (unfortunately, to my mind) “human capital,” which consists of an individual’s labor power, skills, training, and abilities. In this book, capital is defined as the sum total of nonhuman assets that can be owned and exchanged on some market. Capital includes all forms of real property (including residential real estate) as well as financial and professional capital (plants, infrastructure, machinery, patents, and so on) used by firms and government agencies.

There are many reasons for excluding human capital from our definition of capital. The most obvious is that human capital cannot be owned by another person or traded on a market (not permanently, at any rate). This is a key difference from other forms of capital. One can of course put one’s labor services up for hire under a labor contract of some sort. In all modern legal systems, however, such an arrangement has to be limited in both time and scope. In slave societies, of course, this is obviously not true: there, a slaveholder can fully and completely own the human capital of another person and even of that person’s offspring. In such societies, slaves can be bought and sold on the market and conveyed by inheritance, and it is common to include slaves in calculating a slaveholder’s wealth. I will show how this worked when I examine the composition of private capital in the southern United States before 1865. Leaving such special (and for now historical) cases aside, it makes little sense to attempt to add human and nonhuman capital. Throughout history, both forms of wealth have played fundamental and complementary roles in economic growth and development and will continue to do so in the twenty-first century. But in order to understand the growth process and the inequalities it engenders, we must distinguish carefully between human and nonhuman capital and treat each one separately.

Nonhuman capital, which in this book I will call simply “capital,” includes all forms of wealth that individuals (or groups of individuals) can own and that can be transferred or traded through the market on a permanent basis. In practice, capital can be owned by private individuals (in which case we speak of “private capital”) or by the government or government agencies

(in which case we speak of “public capital”). There are also intermediate forms of collective property owned by “moral persons” (that is, entities such as foundations and churches) pursuing specific aims. I will come back to this. The boundary between what private individuals can and cannot own has evolved considerably over time and around the world, as the extreme case of slavery indicates. The same is true of property in the atmosphere, the sea, mountains, historical monuments, and knowledge. Certain private interests would like to own these things, and sometimes they justify this desire on grounds of efficiency rather than mere self-interest. But there is no guarantee that this desire coincides with the general interest. Capital is not an immutable concept: it reflects the state of development and prevailing social relations of each society.

Capital and Wealth

To simplify the text, I use the words “capital” and “wealth” interchangeably, as if they were perfectly synonymous. By some definitions, it would be better to reserve the word “capital” to describe forms of wealth accumulated by human beings (buildings, machinery, infrastructure, etc.) and therefore to exclude land and natural resources, with which humans have been endowed without having to accumulate them. Land would then be a component of wealth but not of capital. The problem is that it is not always easy to distinguish the value of buildings from the value of the land on which they are built. An even greater difficulty is that it is very hard to gauge the value of “virgin” land (as humans found it centuries or millennia ago) apart from improvements due to human intervention, such as drainage, irrigation, fertilization, and so on. The same problem arises in connection with natural resources such as petroleum, gas, rare earth elements, and the like, whose pure value is hard to distinguish from the value added by the investments needed to discover new deposits and prepare them for exploitation. I therefore include all these forms of wealth in capital. Of course, this choice does not eliminate the need to look closely at the origins of wealth, especially the boundary line between accumulation and appropriation.

Some definitions of “capital” hold that the term should apply only to those components of wealth directly employed in the production process. For instance, gold might be counted as part of wealth but not of capital, because

—-1
—0
—+1

gold is said to be useful only as a store of value. Once again, this limitation strikes me as neither desirable nor practical (because gold can be a factor of production, not only in the manufacture of jewelry but also in electronics and nanotechnology). Capital in all its forms has always played a dual role, as both a store of value and a factor of production. I therefore decided that it was simpler not to impose a rigid distinction between wealth and capital.

Similarly, I ruled out the idea of excluding residential real estate from capital on the grounds that it is “unproductive,” unlike the “productive capital” used by firms and government: industrial plants, office buildings, machinery, infrastructure, and so on. The truth is that all these forms of wealth are useful and productive and reflect capital’s two major economic functions. Residential real estate can be seen as a capital asset that yields “housing services,” whose value is measured by their rental equivalent. Other capital assets can serve as factors of production for firms and government agencies that produce goods and services (and need plants, offices, machinery, infrastructure, etc. to do so). Each of these two types of capital currently accounts for roughly half the capital stock in the developed countries.

To summarize, I define “national wealth” or “national capital” as the total market value of everything owned by the residents and government of a given country at a given point in time, provided that it can be traded on some market.⁸ It consists of the sum total of nonfinancial assets (land, dwellings, commercial inventory, other buildings, machinery, infrastructure, patents, and other directly owned professional assets) and financial assets (bank accounts, mutual funds, bonds, stocks, financial investments of all kinds, insurance policies, pension funds, etc.), less the total amount of financial liabilities (debt).⁹ If we look only at the assets and liabilities of private individuals, the result is private wealth or private capital. If we consider assets and liabilities held by the government and other governmental entities (such as towns, social insurance agencies, etc.), the result is public wealth or public capital. By definition, national wealth is the sum of these two terms:

National wealth = private wealth + public wealth

Public wealth in most developed countries is currently insignificant (or even negative, where the public debt exceeds public assets). As I will show, private wealth accounts for nearly all of national wealth almost everywhere.

INCOME AND OUTPUT

This has not always been the case, however, so it is important to distinguish clearly between the two notions.

To be clear, although my concept of capital excludes human capital (which cannot be exchanged on any market in nonslave societies), it is not limited to “physical” capital (land, buildings, infrastructure, and other material goods). I include “immaterial” capital such as patents and other intellectual property, which are counted either as nonfinancial assets (if individuals hold patents directly) or as financial assets (when an individual owns shares of a corporation that holds patents, as is more commonly the case). More broadly, many forms of immaterial capital are taken into account by way of the stock market capitalization of corporations. For instance, the stock market value of a company often depends on its reputation and trademarks, its information systems and modes of organization, its investments, whether material or immaterial, for the purpose of making its products and services more visible and attractive, and so on. All of this is reflected in the price of common stock and other corporate financial assets and therefore in national wealth.

To be sure, the price that the financial markets sets on a company’s or even a sector’s immaterial capital at any given moment is largely arbitrary and uncertain. We see this in the collapse of the Internet bubble in 2000, in the financial crisis that began in 2007–2008, and more generally in the enormous volatility of the stock market. The important fact to note for now is that this is a characteristic of all forms of capital, not just immaterial capital. Whether we are speaking of a building or a company, a manufacturing firm or a service firm, it is always very difficult to set a price on capital. Yet as I will show, total national wealth, that is, the wealth of a country as a whole and not of any particular type of asset, obeys certain laws and conforms to certain regular patterns.

One further point: total national wealth can always be broken down into domestic capital and foreign capital:

National wealth = national capital = domestic capital + net foreign capital

Domestic capital is the value of the capital stock (buildings, firms, etc.) located within the borders of the country in question. Net foreign capital—or net foreign assets—measures the country’s position vis-à-vis the rest of the world: more specifically, it is the difference between assets owned by the

—-1
—0
—+1

country's citizens in the rest of the world and assets of the country owned by citizens of other countries. On the eve of World War I, Britain and France both enjoyed significant net positive asset positions vis-à-vis the rest of the world. One characteristic of the financial globalization that has taken place since the 1980s is that many countries have more or less balanced net asset positions, but those positions are quite large in absolute terms. In other words, many countries have large capital stakes in other countries, but those other countries also have stakes in the country in question, and the two positions are more or less equal, so that net foreign capital is close to zero. Globally, of course, all the net positions must add up to zero, so that total global wealth equals the "domestic" capital of the planet as a whole.

The Capital/Income Ratio

Now that income and capital have been defined, I can move on to the first basic law tying these two ideas together. I begin by defining the capital/income ratio.

Income is a flow. It corresponds to the quantity of goods produced and distributed in a given period (which we generally take to be a year).

Capital is a stock. It corresponds to the total wealth owned at a given point in time. This stock comes from the wealth appropriated or accumulated in all prior years combined.

The most natural and useful way to measure the capital stock in a particular country is to divide that stock by the annual flow of income. This gives us the capital/income ratio, which I denote by the Greek letter β .

For example, if a country's total capital stock is the equivalent of six years of national income, we write $\beta = 6$ (or $\beta = 600\%$).

In the developed countries today, the capital/income ratio generally varies between 5 and 6, and the capital stock consists almost entirely of private capital. In France and Britain, Germany and Italy, the United States and Japan, national income was roughly 30,000–35,000 euros per capita in 2010, whereas total private wealth (net of debt) was typically on the order of 150,000–200,000 euros per capita, or five to six times annual national income. There are interesting variations both within Europe and around the world. For instance, β is greater than 6 in Japan and Italy and less than 5 in the United States and Germany. Public wealth is just barely positive in some countries

and slightly negative in others. And so on. I examine all this in detail in the next few chapters. At this point, it is enough to keep these orders of magnitude in mind, in order to make the ideas as concrete as possible.¹⁰

The fact that national income in the wealthy countries of the world in 2010 was on the order of 30,000 euros per capita per annum (or 2,500 euros per month) obviously does not mean that everyone earns that amount. Like all averages, this average income figure hides enormous disparities. In practice, many people earn much less than 2,500 euros a month, while others earn dozens of times that much. Income disparities are partly the result of unequal pay for work and partly of much larger inequalities in income from capital, which are themselves a consequence of the extreme concentration of wealth. The average national income per capita is simply the amount that one could distribute to each individual if it were possible to equalize the income distribution without altering total output or national income.¹¹

Similarly, private per capita wealth on the order of 180,000 euros, or six years of national income, does not mean that everyone owns that much capital. Many people have much less, while some own millions or tens of millions of euros' worth of capital assets. Much of the population has very little accumulated wealth—significantly less than one year's income: a few thousand euros in a bank account, the equivalent of a few weeks' or months' worth of wages. Some people even have negative wealth: in other words, the goods they own are worth less than the debts they owe. By contrast, others have considerable fortunes, ranging from ten to twenty times their annual income or even more. The capital/income ratio for the country as a whole tells us nothing about inequalities within the country. But β does measure the overall importance of capital in a society, so analyzing this ratio is a necessary first step in the study of inequality. The main purpose of Part Two is to understand how and why the capital/income ratio varies from country to country, and how it has evolved over time.

To appreciate the concrete form that wealth takes in today's world, it is useful to note that the capital stock in the developed countries currently consists of two roughly equal shares: residential capital and professional capital used by firms and government. To sum up, each citizen of one of the wealthy countries earned an average of 30,000 euros per year in 2010, owned approximately 180,000 euros of capital, 90,000 in the form of a dwelling and another 90,000 in stocks, bonds, savings, or other investments.¹² There are interesting

—-1
—0
—+1

variations across countries, which I will analyze in Chapter 2. For now, the fact that capital can be divided into two roughly equal shares will be useful to keep in mind.

The First Fundamental Law of Capitalism: $\alpha = r \times \beta$

I can now present the first fundamental law of capitalism, which links the capital stock to the flow of income from capital. The capital/income ratio β is related in a simple way to the share of income from capital in national income, denoted α . The formula is

$$\alpha = r \times \beta$$

where r is *the rate of return on capital*.

For example, if $\beta = 600\%$ and $r = 5\%$, then $\alpha = r \times \beta = 30\%$.¹³

In other words, if national wealth represents the equivalent of six years of national income, and if the rate of return on capital is 5 percent per year, then capital's share in national income is 30 percent.

The formula $\alpha = r \times \beta$ is a pure accounting identity. It can be applied to all societies in all periods of history, by definition. Though tautological, it should nevertheless be regarded as the first fundamental law of capitalism, because it expresses a simple, transparent relationship among the three most important concepts for analyzing the capitalist system: the capital/income ratio, the share of capital in income, and the rate of return on capital.

The rate of return on capital is a central concept in many economic theories. In particular, Marxist analysis emphasizes the falling rate of profit—a historical prediction that turned out to be quite wrong, although it does contain an interesting intuition. The concept of the rate of return on capital also plays a central role in many other theories. In any case, the rate of return on capital measures the yield on capital over the course of a year regardless of its legal form (profits, rents, dividends, interest, royalties, capital gains, etc.), expressed as a percentage of the value of capital invested. It is therefore a broader notion than the “rate of profit,”¹⁴ and much broader than the “rate of interest,”¹⁵ while incorporating both.

Obviously, the rate of return can vary widely, depending on the type of investment. Some firms generate rates of return greater than 10 percent per

INCOME AND OUTPUT

year; others make losses (negative rate of return). The average long-run rate of return on stocks is 7–8 percent in many countries. Investments in real estate and bonds frequently return 3–4 percent, while the real rate of interest on public debt is sometimes much lower. The formula $\alpha = r \times \beta$ tells us nothing about these subtleties, but it does tell us how to relate these three quantities, which can be useful for framing discussion.

For example, in the wealthy countries around 2010, income from capital (profits, interests, dividends, rents, etc.) generally hovered around 30 percent of national income. With a capital/income ratio on the order of 600 percent, this meant that the rate of return on capital was around 5 percent.

Concretely, this means that the current per capita national income of 30,000 euros per year in rich countries breaks down as 21,000 euros per year income from labor (70 percent) and 9,000 euros income from capital (30 percent). Each citizen owns an average of 180,000 euros of capital, and the 9,000 euros of income from capital thus corresponds to an average annual return on capital of 5 percent.

Once again, I am speaking here only of averages: some individuals receive far more than 9,000 euros per year in income from capital, while others receive nothing while paying rent to their landlords and interest to their creditors. Considerable country-to-country variation also exists. In addition, measuring the share of income from capital is often difficult in both a conceptual and a practical sense, because there are some categories of income (such as nonwage self-employment income and entrepreneurial income) that are hard to break down into income from capital and income from labor. In some cases this can make comparison misleading. When such problems arise, the least imperfect method of measuring the capital share of income may be to apply a plausible average rate of return to the capital/income ratio. At this stage, the orders of magnitude given above ($\beta = 600\%$, $\alpha = 30\%$, $r = 5\%$) may be taken as typical.

For the sake of concreteness, let us note, too, that the average rate of return on land in rural societies is typically on the order of 4–5 percent. In the novels of Jane Austen and Honoré de Balzac, the fact that land (like government bonds) yields roughly 5 percent of the amount of capital invested (or, equivalently, that the value of capital corresponds to roughly twenty years of annual rent) is so taken for granted that it often goes unmentioned. Contemporary readers were well aware that it took capital on the order of 1 million francs to produce an annual rent of 50,000 francs. For nineteenth-century novelists

and their readers, the relation between capital and annual rent was self-evident, and the two measuring scales were used interchangeably, as if rent and capital were synonymous, or perfect equivalents in two different languages.

Now, at the beginning of the twenty-first century, we find roughly the same return on real estate, 4–5 percent, sometimes a little less, especially where prices have risen rapidly without dragging rents upward at the same rate. For example, in 2010, a large apartment in Paris, valued at 1 million euros, typically rents for slightly more than 2,500 euros per month, or annual rent of 30,000 euros, which corresponds to a return on capital of only 3 percent per year from the landlord's point of view. Such a rent is nevertheless quite high for a tenant living solely on income from labor (one hopes he or she is paid well) while it represents a significant income for the landlord. The bad news (or good news, depending on your point of view) is that things have always been like this. This type of rent tends to rise until the return on capital is around 4 percent (which in this example would correspond to a rent of 3,000–3,500 euros per month, or 40,000 per year). Hence this tenant's rent is likely to rise in the future. The landlord's annual return on investment may eventually be enhanced by a long-term capital gain on the value of the apartment. Smaller apartments yield a similar or perhaps slightly higher return. An apartment valued at 100,000 euros may yield 400 euros a month in rent, or nearly 5,000 per year (5 percent). A person who owns such an apartment and chooses to live in it can save the rental equivalent and devote that money to other uses, which yields a similar return on investment.

Capital invested in businesses is of course at greater risk, so the average return is often higher. The stock-market capitalization of listed companies in various countries generally represents 12 to 15 years of annual profits, which corresponds to an annual return on investment of 6–8 percent (before taxes).

The formula $\alpha = r \times \beta$ allows us to analyze the importance of capital for an entire country or even for the planet as a whole. It can also be used to study the accounts of a specific company. For example, take a firm that uses capital valued at 5 million euros (including offices, infrastructure, machinery, etc.) to produce 1 million euros worth of goods annually, with 600,000 euros going to pay workers and 400,000 euros in profits.¹⁶ The capital/income ratio of this company is $\beta = 5$ (its capital is equivalent to five years of output), the capital share α is 40 percent, and the rate of return on capital is $r = 8$ percent.

Imagine another company that uses less capital (3 million euros) to produce the same output (1 million euros), but using more labor (700,000 euros in wages, 300,000 in profits). For this company, $\beta = 3$, $\alpha = 30$ percent, and $r = 10$ percent. The second firm is less capital intensive than the first, but it is more profitable (the rate of return on its capital is significantly higher).

In all countries, the magnitudes of β , α , and r vary a great deal from company to company. Some sectors are more capital intensive than others: for example, the metal and energy sectors are more capital intensive than the textile and food processing sectors, and the manufacturing sector is more capital intensive than the service sector. There are also significant variations between firms in the same sector, depending on their choice of production technology and market position. The levels of β , α , and r in a given country also depend on the relative shares of residential real estate and natural resources in total capital.

It bears emphasizing that the law $\alpha = r \times \beta$ does not tell us how each of these three variables is determined, or, in particular, how the national capital/income ratio (β) is determined, the latter being in some sense a measure of how intensely capitalistic the society in question is. To answer that question, we must introduce additional ideas and relationships, in particular the savings and investment rates and the rate of growth. This will lead us to the second fundamental law of capitalism: the higher the savings rate and the lower the growth rate, the higher the capital/income ratio (β). This will be shown in the next few chapters; at this stage, the law $\alpha = r \times \beta$ simply means that regardless of what economic, social, and political forces determine the level of the capital/income ratio (β), capital's share in income (α), and the rate of return on capital (r), these three variables are not independent of one another. Conceptually, there are two degrees of freedom, not three.

National Accounts: An Evolving Social Construct

Now that the key concepts of output and income, capital and wealth, capital/income ratio, and rate of return on capital have been explained, I will examine in greater detail how these abstract quantities can be measured and what such measurements can tell us about the historical evolution of the distribution of wealth in various countries. I will briefly review the main stages in the history of national accounts and then present a portrait in broad brushstrokes of how the global distribution of output and income has changed since the

—-1
—0
—+1

eighteenth century, along with a discussion of how demographic and economic growth rates have changed over the same period. These growth rates will play an important part in the analysis.

As noted, the first attempts to measure national income and capital date back to the late seventeenth and early eighteenth century. Around 1700, several isolated estimates appeared in Britain and France (apparently independently of one another). I am speaking primarily of the work of William Petty (1664) and Gregory King (1696) for England and Pierre le Pesant, sieur de Boisguillebert (1695), and Sébastien Le Prestre de Vauban (1707) for France. Their work focused on both the national stock of capital and the annual flow of national income. One of their primary objectives was to calculate the total value of land, by far the most important source of wealth in the agrarian societies of the day, and then to relate the quantity of landed wealth to the level of agricultural output and land rents.

It is worth noting that these authors often had a political objective in mind, generally having to do with modernization of the tax system. By calculating the nation's income and wealth, they hoped to show the sovereign that it would be possible to raise tax receipts considerably while keeping tax rates relatively low, provided that all property and goods produced were subject to taxation and everyone was required to pay, including landlords of both aristocratic and common descent. This objective is obvious in Vauban's *Projet de dîme royale* (Plan for a Royal Tithe), but it is just as clear in the works of Boisguillebert and King (though less so in Petty's writing).

The late eighteenth century saw further attempts to measure income and wealth, especially around the time of the French Revolution. Antoine Lavoisier published his estimates for the year 1789 in his book *La Richesse territoriale du Royaume de France* (The Territorial Wealth of the Kingdom of France), published in 1791. The new tax system established after the Revolution, which ended the privileges of the nobility and imposed a tax on all property in land, was largely inspired by this work, which was widely used to estimate expected receipts from new taxes.

It was above all in the nineteenth century, however, that estimates of national wealth proliferated. From 1870 to 1900, Robert Giffen regularly updated his estimates of Britain's stock of national capital, which he compared to estimates by other authors (especially Patrick Colquhoun) from the early 1800s. Giffen marveled at the size of Britain's stock of industrial capital as

well as the stock of foreign assets acquired since the Napoleonic wars, which was many times larger than the entire public debt due to those wars.¹⁷ In France at about the same time, Alfred de Foville and Clément Colson published estimates of “national wealth” and “private wealth,” and, like Giffen, both writers also marveled at the considerable accumulation of private capital over the course of the nineteenth century. It was glaringly obvious to everyone that private fortunes were prospering in the period 1870–1914. For the economists of the day, the problem was to measure that wealth and compare different countries (the Franco-British rivalry was never far from their minds). Until World War I, estimates of wealth received much more attention than estimates of income and output, and there were in any case more of them, not only in Britain and France but also in Germany, the United States, and other industrial powers. In those days, being an economist meant first and foremost being able to estimate the national capital of one’s country: this was almost a rite of initiation.

It was not until the period between the two world wars that national accounts began to be established on an annual basis. Previous estimates had always focused on isolated years, with successive estimates separated by ten or more years, as in the case of Giffen’s calculations of British national capital in the nineteenth century. In the 1930s, improvements in the primary statistical sources made the first annual series of national income data possible. These generally went back as far as the beginning of the twentieth century or the last decades of the nineteenth. They were established for the United States by Kuznets and Kendrick, for Britain by Bowley and Clark, and for France by Dugé de Bernonville. After World War II, government statistical offices supplanted economists and began to compile and publish official annual data on GDP and national income. These official series continue to this day.

Compared with the pre-World War I period, however, the focal point of the data had changed entirely. From the 1940s on, the primary motivation was to respond to the trauma of the Great Depression, during which governments had no reliable annual estimates of economic output. There was therefore a need for statistical and political tools in order to steer the economy properly and avoid a repeat of the catastrophe. Governments thus insisted on annual or even quarterly data on output and income. Estimates of national wealth, which had been so prized before 1914, now took a backseat, especially after the economic and political chaos of 1914–1945 made it difficult to interpret their meaning. Specifically, the prices of real estate and financial assets

—-1
—0
—+1

fell to extremely low levels, so low that private capital seemed to have evaporated. In the 1950s and 1960s, a period of reconstruction, the main goal was to measure the remarkable growth of output in various branches of industry.

In the 1990s–2000s, wealth accounting again came to the fore. Economists and political leaders were well aware that the financial capitalism of the twenty-first century could not be properly analyzed with the tools of the 1950s and 1960s. In collaboration with central banks, government statistical agencies in various developed countries compiled and published annual series of data on the assets and liabilities of different groups, in addition to the usual income and output data. These wealth accounts are still far from perfect: for example, natural capital and damages to the environment are not well accounted for. Nevertheless, they represent real progress in comparison with national accounts from the early postwar years, which were concerned solely with endless growth in output.¹⁸ These are the official series that I use in this book to analyze aggregate wealth and the current capital/income ratio in the wealthy countries.

One conclusion stands out in this brief history of national accounting: national accounts are a social construct in perpetual evolution. They always reflect the preoccupations of the era when they were conceived.¹⁹ We should be careful not to make a fetish of the published figures. When a country's national income per capita is said to be 30,000 euros, it is obvious that this number, like all economic and social statistics, should be regarded as an estimate, a construct, and not a mathematical certainty. It is simply the best estimate we have. National accounts represent the only consistent, systematic attempt to analyze a country's economic activity. They should be regarded as a limited and imperfect research tool, a compilation and arrangement of data from highly disparate sources. In all developed countries, national accounts are currently compiled by government statistical offices and central banks from the balance sheets and account books of financial and nonfinancial corporations together with many other statistical sources and surveys. We have no reason to think a priori that the officials involved in these efforts do not do their best to spot inconsistencies in the data in order to achieve the best possible estimates. Provided we use these data with caution and in a critical spirit and complement them with other data where there are errors or gaps (say, in

-1—
0—
+1—

dealing with tax havens), these national accounts are an indispensable tool for estimating aggregate income and wealth.

In particular, as I will show in Part Two, we can put together a consistent analysis of the historical evolution of the capital/income ratio by meticulously compiling and comparing national wealth estimates by many authors from the eighteenth to the early twentieth century and connecting them up with official capital accounts from the late twentieth and early twenty-first century. The other major limitation of official national accounts, apart from their lack of historical perspective, is that they are deliberately concerned only with aggregates and averages and not with distributions and inequalities. We must therefore draw on other sources to measure the distribution of income and wealth and to study inequalities. National accounts thus constitute a crucial element of our analyses, but only when completed with additional historical and distributional data.

The Global Distribution of Production

I begin by examining the evolution of the global distribution of production, which is relatively well known from the early nineteenth century on. For earlier periods, estimates are more approximate, but we know the broad outlines, thanks most notably to the historical work of Angus Maddison, especially since the overall pattern is relatively simple.²⁰

From 1900 to 1980, 70–80 percent of the global production of goods and services was concentrated in Europe and America, which incontestably dominated the rest of the world. By 2010, the European–American share had declined to roughly 50 percent, or approximately the same level as in 1860. In all probability, it will continue to fall and may go as low as 20–30 percent at some point in the twenty-first century. This was the level maintained up to the turn of the nineteenth century and would be consistent with the European–American share of the world’s population (see Figures 1.1 and 1.2).

In other words, the lead that Europe and America achieved during the Industrial Revolution allowed these two regions to claim a share of global output that was two to three times greater than their share of the world’s population simply because their output per capita was two to three times

—-1
—0
—+1

INCOME AND CAPITAL

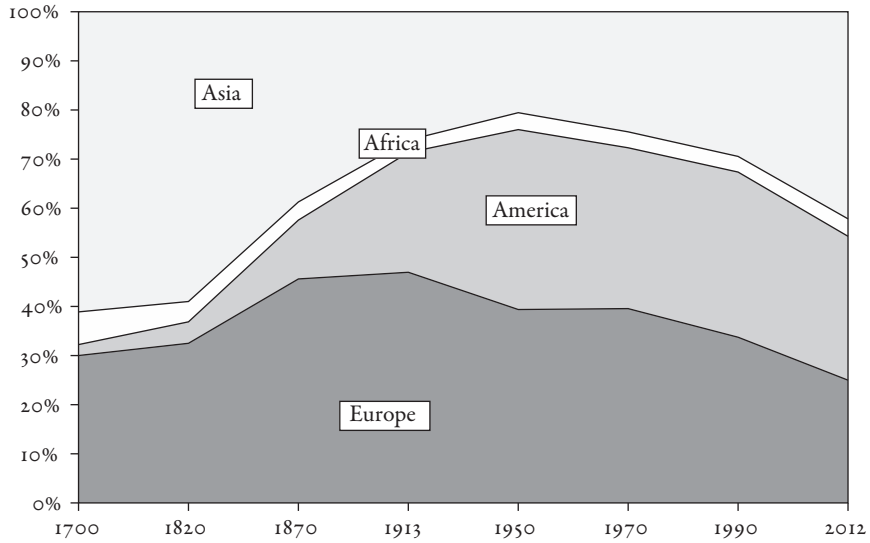


FIGURE 1.1. The distribution of world output, 1700–2012

Europe’s GDP made 47 percent of world GDP in 1913, down to 25 percent in 2012. Sources and series: see piketty.pse.ens.fr/capital21c.

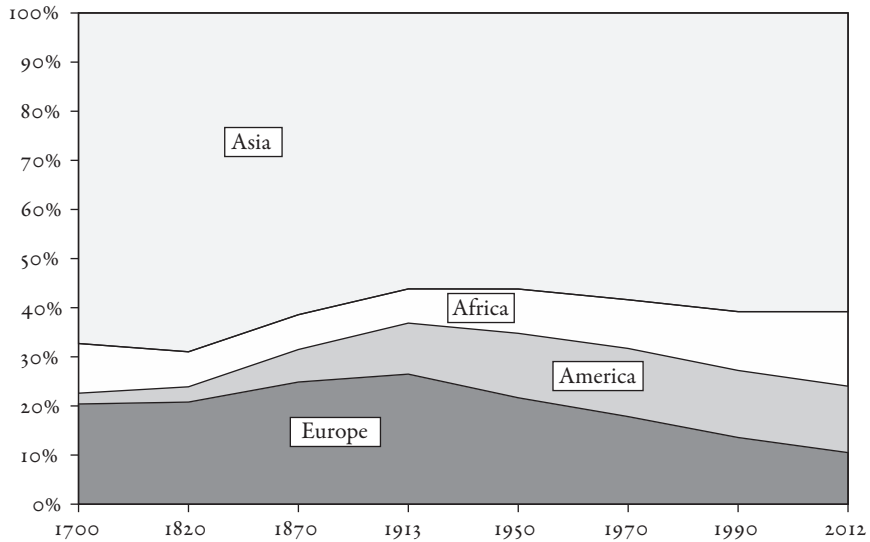


FIGURE 1.2. The distribution of world population, 1700–2012

Europe’s population made 26 percent of world population in 1913, down to 10 percent in 2012. Sources and series: see piketty.pse.ens.fr/capital21c.

-1—
0—
+1—

INCOME AND OUTPUT

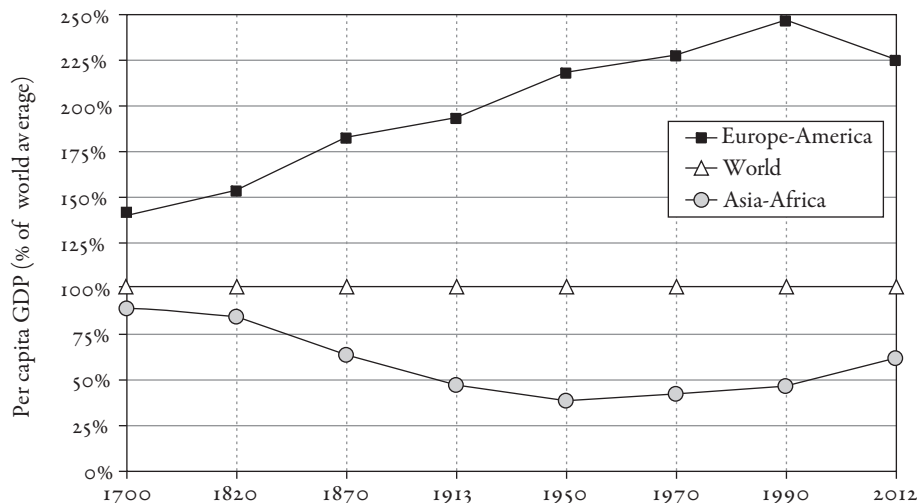


FIGURE 1.3. Global inequality 1700–2012: divergence then convergence?

Per capita GDP in Asia-Africa went from 37 percent of world average in 1950 to 61 percent in 2012.

Sources and series: see piketty.pse.ens.fr/capital21c.

greater than the global average.²¹ All signs are that this phase of divergence in per capita output is over and that we have embarked on a period of convergence. The resulting “catch-up” phenomenon is far from over, however (see Figure 1.3). It is far too early to predict when it might end, especially since the possibility of economic and/or political reversals in China and elsewhere obviously cannot be ruled out.

From Continental Blocs to Regional Blocs

The general pattern just described is well known, but a number of points need to be clarified and refined. First, putting Europe and the Americas together as a single “Western bloc” simplifies the presentation but is largely artificial. Europe attained its maximal economic weight on the eve of World War I, when it accounted for nearly 50 percent of global output, and it has declined steadily since then, whereas America attained its peak in the 1950s, when it accounted for nearly 40 percent of global output.

Furthermore, both Europe and the Americas can be broken down into two highly unequal subregions: a hyperdeveloped core and a less developed

—-1
—0
—+1

periphery. Broadly speaking, global inequality is best analyzed in terms of regional blocs rather than continental blocs. This can be seen clearly in Table 1.1, which shows the distribution of global output in 2012. All these numbers are of no interest in themselves, but it is useful to familiarize oneself with the principal orders of magnitude.

The population of the planet is close to 7 billion in 2012, and global output is slightly greater than 70 trillion euros, so that global output per capita is almost exactly 10,000 euros. If we subtract 10 percent for capital depreciation and divide by 12, we find that this yields an average per capita monthly income of 760 euros, which may be a clearer way of making the point. In other words, if global output and the income to which it gives rise were equally divided, each individual in the world would have an income of about 760 euros per month.

The population of Europe is about 740 million, about 540 million of whom live in member countries of the European Union, whose per capita output exceeds 27,000 euros per year. The remaining 200 million people live in Russia and Ukraine, where the per capita output is about 15,000 euros per year, barely 50 percent above the global average.²² The European Union itself is relatively heterogeneous: 410 million of its citizens live in what used to be called Western Europe, three-quarters of them in the five most populous countries of the Union, namely Germany, France, Great Britain, Italy, and Spain, with an average per capita GDP of 31,000 euros per year, while the remaining 130 million live in what used to be Eastern Europe, with an average per capita output on the order of 16,000 euros per year, not very different from the Russia-Ukraine bloc.²³

The Americas can also be divided into distinct regions that are even more unequal than the European center and periphery: the US-Canada bloc has 350 million people with a per capita output of 40,000 euros, while Latin America has 600 million people with a per capita output of 10,000 euros, exactly equal to the world average.

Sub-Saharan Africa, with a population of 900 million and an annual output of only 1.8 trillion euros (less than the French GDP of 2 trillion), is economically the poorest region of the world, with a per capita output of only 2,000 euros per year. India is slightly higher, while North Africa does markedly better, and China even better than that: with a per capita output of

-1—
0—
+1—

TABLE I.1.
Distribution of world GDP, 2012.

Region	Population (million inhabitants)	GDP (billion euros 2012)	Per capita GDP (euros 2012)	Equivalent per capita monthly income (euros 2012)
World	7,050	71,200	10,100	760
Europe	740	17,800	24,000	1,800
incl. European Union	540	14,700	27,300	2,040
incl. Russia/Ukraine	200	3,100	15,400	1,150
America	950	20,600	21,500	1,620
incl. United States/Canada	350	14,300	40,700	3,050
incl. Latin America	600	6,300	10,400	780
Africa	1,070	2,800	2,600	200
incl. North Africa	170	1,000	5,700	430
incl. Sub-Saharan Africa	900	1,800	2,000	150
Asia	4,290	30,000	7,000	520
incl. China	1,350	10,400	7,700	580
incl. India	1,260	4,000	3,200	240
incl. Japan	130	3,800	30,000	2,250
incl. other	1,550	11,800	7,600	570

Note: World GDP, estimated in purchasing power parity, was about 71,200 billion euros in 2012. World population was about 7,050 billion inhabitants, hence a per capita GDP of €10,100 (equivalent to a monthly income of about €760 per month). All numbers were rounded to the closed dozen or hundred.

Sources: See piketty.pse.ens.fr/capital21c.

8,000 euros per year, China in 2012 is not far below the world average. Japan's annual per capita output is equal to that of the wealthiest European countries (approximately 30,000 euros), but its population is such a small minority in the greater Asian population that it has little influence on the continental average, which is close to that of China.²⁴

*Global Inequality: From 150 Euros per Month to
3,000 Euros per Month*

To sum up, global inequality ranges from regions in which the per capita income is on the order of 150–250 euros per month (sub-Saharan Africa, India) to regions where it is as high as 2,500–3,000 euros per month (Western Europe, North America, Japan), that is, ten to twenty times higher. The global average, which is roughly equal to the Chinese average, is around 600–800 euros per month.

These orders of magnitude are significant and worth remembering. Bear in mind, however, that the margin of error in these figures is considerable: it is always much more difficult to measure inequalities between countries (or between different periods) than within them.

For example, global inequality would be markedly higher if we used current exchange rates rather than purchasing power parities, as I have done thus far. To understand what these terms mean, first consider the euro/dollar exchange rate. In 2012, a euro was worth about \$1.30 on the foreign exchange market. A European with an income of 1,000 euros per month could go to his or her bank and exchange that amount for \$1,300. If that person then took that money to the United States to spend, his or her purchasing power would be \$1,300. But according to the official International Comparison Program (ICP), European prices are about 10 percent higher than American prices, so that if this same European spent the same money in Europe, his or her purchasing power would be closer to an American income of \$1,200. Thus we say that \$1.20 has “purchasing power parity” with 1 euro. I used this parity rather than the exchange rate to convert American GDP to euros in Table 1.1, and I did the same for the other countries listed. In other words, we compare the GDP of different countries on the basis of the actual purchasing power of their citizens, who generally spend their income at home rather than abroad.²⁵

INCOME AND OUTPUT

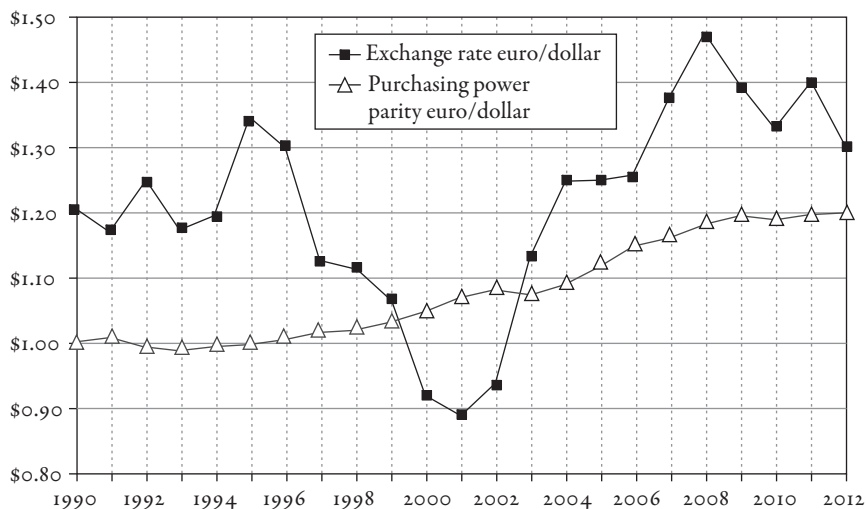


FIGURE 1.4. Exchange rate and purchasing power parity: euro/dollar

In 2012, 1 euro was worth \$1.30 according to current exchange rate, but \$1.20 in purchasing power parity.

Sources and series: see piketty.pse.ens.fr/capital21c.

The other advantage of using purchasing power parities is that they are more stable than exchange rates. Indeed, exchange rates reflect not only the supply and demand for the goods and services of different countries but also sudden changes in the investment strategies of international investors and volatile estimates of the political and/or financial stability of this or that country, to say nothing of unpredictable changes in monetary policy. Exchange rates are therefore extremely volatile, as a glance at the large fluctuations of the dollar over the past few decades will show. The dollar/euro rate went from \$1.30 per euro in the 1990s to less than \$0.90 in 2001 before rising to around \$1.50 in 2008 and then falling back to \$1.30 in 2012. During that time, the purchasing power parity of the euro rose gently from roughly \$1 per euro in the early 1990s to roughly \$1.20 in 2010 (see Figure 1.4).²⁶

Despite the best efforts of the international organizations involved in the ICP, there is no escaping the fact that these purchasing power parity estimates are rather uncertain, with margins of error on the order of 10 percent if not higher, even between countries at comparable levels of development. For example, the most recent available survey shows that while some European prices (for energy, housing, hotels, and restaurants) are indeed higher than

—-1
—0
—+1

comparable American prices, others are sharply lower (for health and education, for instance).²⁷ In theory, the official estimates weight all prices according to the weight of various goods and services in a typical budget for each country, but such calculations clearly leave a good deal of room for error, particularly since it is very hard to measure qualitative differences for many services. In any case, it is important to emphasize that each of these price indices measures a different aspect of social reality. The price of energy measures purchasing power for energy (which is greater in the United States), while the price of health care measures purchasing power in that area (which is greater in Europe). The reality of inequality between countries is multidimensional, and it is misleading to say that it can all be summed up with a single index leading to an unambiguous classification, especially between countries with fairly similar average incomes.

In the poorer countries, the corrections introduced by purchasing power parity are even larger: in Africa and Asia, prices are roughly half what they are in the rich countries, so that GDP roughly doubles when purchasing power parity is used for comparisons rather than the market exchange rate. This is chiefly a result of the fact that the prices of goods and services that cannot be traded internationally are lower, because these are usually relatively labor intensive and involve relatively unskilled labor (a relatively abundant factor of production in less developed countries), as opposed to skilled labor and capital (which are relatively scarce in less developed countries).²⁸ Broadly speaking, the poorer a country is, the greater the correction: in 2012, the correction coefficient was 1.6 in China and 2.5 in India.²⁹ At this moment, the euro is worth 8 Chinese yuan on the foreign exchange market but only 5 yuan in purchasing power parity. The gap is shrinking as China develops and revalues the yuan (see Figure 1.5). Some writers, including Angus Maddison, argue that the gap is not as small as it might appear and that official international statistics underestimate Chinese GDP.³⁰

Because of the uncertainties surrounding exchange rates and purchasing power parities, the average per capita monthly incomes discussed earlier (150–250 euros for the poorest countries, 600–800 euros for middling countries, and 2,500–3,000 euros for the richest countries) should be treated as approximations rather than mathematical certainties. For example, the share of the rich countries (European Union, United States, Canada, and Japan) in global income was 46 percent in 2012 if we use purchasing power parity but 57 per-

INCOME AND OUTPUT

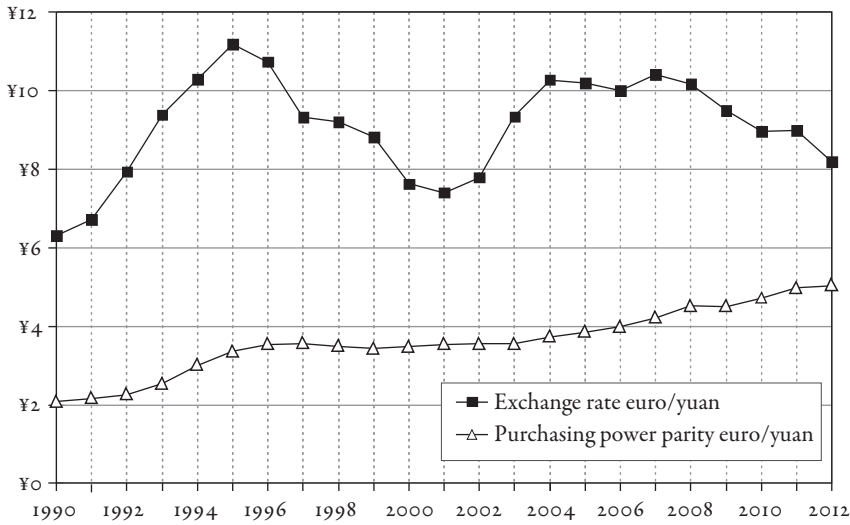


FIGURE 1.5. Exchange rate and purchasing power parity: euro/yuan

In 2012, one euro was worth eight yuan according to current exchange rate, but five yuan in purchasing power parity.

Sources and series: see piketty.pse.ens.fr/capital21c.

cent if we use current exchange rates.³¹ The “truth” probably lies somewhere between these two figures and is probably closer to the first. Still, the orders of magnitude remain the same, as does the fact that the share of income going to the wealthy countries has been declining steadily since the 1970s. Regardless of what measure is used, the world clearly seems to have entered a phase in which rich and poor countries are converging in income.

The Global Distribution of Income Is More Unequal Than the Distribution of Output

To simplify the exposition, the discussion thus far has assumed that the national income of each continental or regional grouping coincided with its domestic product: the monthly incomes indicated in Table 1.1 were obtained simply by deducting 10 percent from GDP (to account for depreciation of capital) and dividing by twelve.

In fact, it is valid to equate income and output only at the global level and not at the national or continental level. Generally speaking, the global income

—-1
—0
—+1

distribution is more unequal than the output distribution, because the countries with the highest per capita output are also more likely to own part of the capital of other countries and therefore to receive a positive flow of income from capital originating in countries with a lower level of per capita output. In other words, the rich countries are doubly wealthy: they both produce more at home and invest more abroad, so that their national income per head is greater than their output per head. The opposite is true for poor countries.

More specifically, all of the major developed countries (the United States, Japan, Germany, France, and Britain) currently enjoy a level of national income that is slightly greater than their domestic product. As noted, however, net income from abroad is just slightly positive and does not radically alter the standard of living in these countries. It amounts to about 1 or 2 percent of GDP in the United States, France, and Britain and 2–3 percent of GDP in Japan and Germany. This is nevertheless a significant boost to national income, especially for Japan and Germany, whose trade surpluses have enabled them to accumulate over the past several decades substantial reserves of foreign capital, the return on which is today considerable.

I turn now from the wealthiest countries taken individually to continental blocs taken as a whole. What we find in Europe, America, and Asia is something close to equilibrium: the wealthier countries in each bloc (generally in the north) receive a positive flow of income from capital, which is partly canceled by the flow out of other countries (generally in the south and east), so that at the continental level, total income is almost exactly equal to total output, generally within 0.5 percent.³²

The only continent not in equilibrium is Africa, where a substantial share of capital is owned by foreigners. According to the balance of payments data compiled since 1970 by the United Nations and other international organizations such as the World Bank and International Monetary Fund, the income of Africans is roughly 5 percent less than the continent's output (and as high as 10 percent lower in some individual countries).³³ With capital's share of income at about 30 percent, this means that nearly 20 percent of African capital is owned by foreigners: think of the London stockholders of the Marikana platinum mine discussed at the beginning of this chapter.

It is important to realize what such a figure means in practice. Since some kinds of wealth (such as residential real estate and agricultural capital) are rarely owned by foreign investors, it follows that the foreign-owned share of

Africa's manufacturing capital may exceed 40–50 percent and may be higher still in other sectors. Despite the fact that there are many imperfections in the balance of payments data, foreign ownership is clearly an important reality in Africa today.

If we look back farther in time, we find even more marked international imbalances. On the eve of World War I, the national income of Great Britain, the world's leading investor, was roughly 10 percent above its domestic product. The gap was more than 5 percent in France, the number two colonial power and global investor, and Germany was a close third, even though its colonial empire was insignificant, because its highly developed industrial sector accumulated large claims on the rest of the world. British, French, and German investment went partly to other European countries and the United States and partly to Asia and Africa. Overall, the European powers in 1913 owned an estimated one-third to one-half of the domestic capital of Asia and Africa and more than three-quarters of their industrial capital.³⁴

What Forces Favor Convergence?

In theory, the fact that the rich countries own part of the capital of poor countries can have virtuous effects by promoting convergence. If the rich countries are so flush with savings and capital that there is little reason to build new housing or add new machinery (in which case economists say that the “marginal productivity of capital,” that is, the additional output due to adding one new unit of capital “at the margin,” is very low), it can be collectively efficient to invest some part of domestic savings in poorer countries abroad. Thus the wealthy countries—or at any rate the residents of wealthy countries with capital to spare—will obtain a better return on their investment by investing abroad, and the poor countries will increase their productivity and thus close the gap between them and the rich countries. According to classical economic theory, this mechanism, based on the free flow of capital and equalization of the marginal productivity of capital at the global level, should lead to convergence of rich and poor countries and an eventual reduction of inequalities through market forces and competition.

This optimistic theory has two major defects, however. First, from a strictly logical point of view, the equalization mechanism does not guarantee global convergence of per capita income. At best it can give rise to convergence

—-1
—0
—+1

of per capita output, provided we assume perfect capital mobility and, even more important, total equality of skill levels and human capital across countries—no small assumption. In any case, the possible convergence of output per head does *not* imply convergence of income per head. After the wealthy countries have invested in their poorer neighbors, they may continue to own them indefinitely, and indeed their share of ownership may grow to massive proportions, so that the per capita national income of the wealthy countries remains permanently greater than that of the poorer countries, which must continue to pay to foreigners a substantial share of what their citizens produce (as African countries have done for decades). In order to determine how likely such a situation is to arise, we must compare the rate of return on capital that the poor countries must pay to the rich to the growth rates of rich and poor economies. Before proceeding down this road, we must first gain a better understanding of the dynamics of the capital/income ratio within a given country.

Furthermore, if we look at the historical record, it does not appear that capital mobility has been the primary factor promoting convergence of rich and poor nations. None of the Asian countries that have moved closer to the developed countries of the West in recent years has benefited from large foreign investments, whether it be Japan, South Korea, or Taiwan and more recently China. In essence, all of these countries themselves financed the necessary investments in physical capital and, even more, in human capital, which the latest research holds to be the key to long-term growth.³⁵ Conversely, countries owned by other countries, whether in the colonial period or in Africa today, have been less successful, most notably because they have tended to specialize in areas without much prospect of future development and because they have been subject to chronic political instability.

Part of the reason for that instability may be the following. When a country is largely owned by foreigners, there is a recurrent and almost irrepressible social demand for expropriation. Other political actors respond that investment and development are possible only if existing property rights are unconditionally protected. The country is thus caught in an endless alternation between revolutionary governments (whose success in improving actual living conditions for their citizens is often limited) and governments dedicated to the protection of existing property owners, thereby laying the groundwork for the next revolution or coup. Inequality of capital ownership is already dif-

difficult to accept and peacefully maintain within a single national community. Internationally, it is almost impossible to sustain without a colonial type of political domination.

Make no mistake: participation in the global economy is not negative in itself. Autarky has never promoted prosperity. The Asian countries that have lately been catching up with the rest of the world have clearly benefited from openness to foreign influences. But they have benefited far more from open markets for goods and services and advantageous terms of trade than from free capital flows. China, for example, still imposes controls on capital: foreigners cannot invest in the country freely, but that has not hindered capital accumulation, for which domestic savings largely suffice. Japan, South Korea, and Taiwan all financed investment out of savings. Many studies also show that gains from free trade come mainly from the diffusion of knowledge and from the productivity gains made necessary by open borders, not from static gains associated with specialization, which appear to be fairly modest.³⁶

To sum up, historical experience suggests that the principal mechanism for convergence at the international as well as the domestic level is the diffusion of knowledge. In other words, the poor catch up with the rich to the extent that they achieve the same level of technological know-how, skill, and education, not by becoming the property of the wealthy. The diffusion of knowledge is not like manna from heaven: it is often hastened by international openness and trade (autarky does not encourage technological transfer). Above all, knowledge diffusion depends on a country's ability to mobilize financing as well as institutions that encourage large-scale investment in education and training of the population while guaranteeing a stable legal framework that various economic actors can reliably count on. It is therefore closely associated with the achievement of legitimate and efficient government. Concisely stated, these are the main lessons that history has to teach about global growth and international inequalities.