

## CHAPTER 1

---

## Introduction

*A complete, consistent, unified theory . . . would be the ultimate triumph of human reason.*

—Stephen W. Hawking

The transition from an epoch of stagnation to an era of sustained economic growth has marked the onset of one of the most remarkable transformations in the course of human history. While living standards in the world economy stagnated during the millennia preceding the Industrial Revolution, income per capita has undergone an unprecedented tenfold increase over the past two centuries, profoundly altering the level and distribution of education, health, and wealth across the globe.

The rise in the standard of living has not been universally shared among societies. Variation in the timing of the take-off from stagnation to growth has led to a vast worldwide divergence in income per capita. Inequality, which had been modest until the nineteenth century, has widened considerably, and the ratio of income per capita between the richest and the poorest regions of the world has been magnified from a moderate 3:1 ratio in 1820 to a staggering 18:1 ratio in 2000 (Figure 1.1).

An equally striking development has emerged in the world distribution of population. The decline in population growth in Europe and North America toward the end of the nineteenth century and the long delay in the onset of a corresponding demographic transition in less developed regions, well into the second half of the twentieth century, have generated significant bifurcation in the global distribution of population. The share of world population that resides in the prosperous region of Europe has declined by nearly one-half over the past century, whereas the fraction of the human population that lives in the impoverished regions of Africa and Latin America has doubled.

Throughout most of human existence, the process of development was marked by Malthusian stagnation: resources generated by technological progress and land expansion were channeled primarily toward an increase in the size of the population, providing only a glacial contribution to the level of income per capita in the long run. While cross-country variations in technology and land productivity were reflected in differing population densities, their effect on variation in living standards was merely transitory.

## 2 • Chapter 1

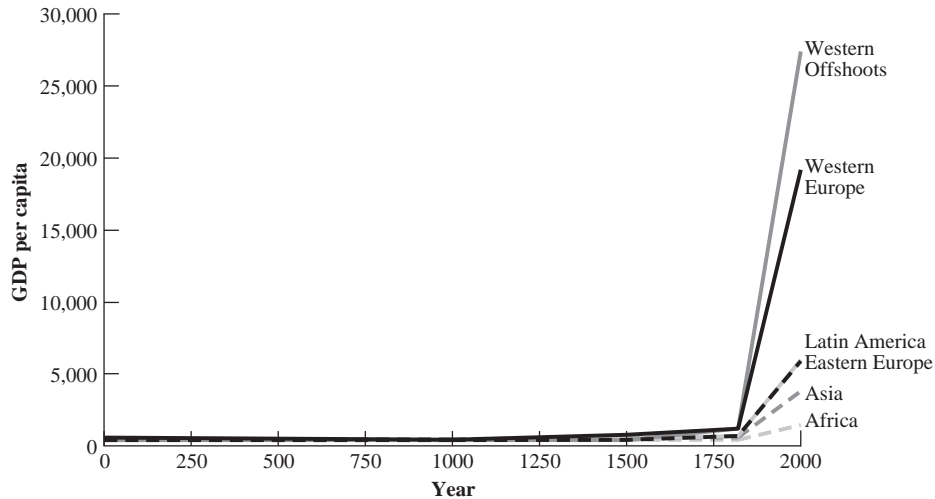


FIGURE 1.1. Evolution of regional income per capita over the past two thousand years.

*Data source:* Maddison (2001).

*Note:* The Western Offshoots are Australia, Canada, New Zealand, and the United States.

In contrast, over the past two centuries, various regions of the world have departed from the Malthusian trap and have witnessed a considerable increase in growth rates of income per capita. The decline in population growth over the course of the demographic transition has liberated productivity gains from the counterbalancing effect of population growth and enabled technological progress and human capital formation to pave the way for the emergence of an era of sustained economic growth.

The transition from an epoch of Malthusian stagnation to an era of sustained economic growth and the corresponding divergence in income per capita across the globe have been the center of intensive research during the past decade. The inconsistency of the predominant theories of economic growth with some of the most fundamental characteristics of the growth process and their limited ability to shed light on the origins of the vast global disparity in living standards have led to the development of a unified theory of economic growth that captures the growth process in its entirety.

Unified Growth Theory explores the fundamental factors that have contributed to the remarkable transition from stagnation to growth and examines their significance for the understanding of the contemporary growth process of developed and less developed economies. First, it unveils the factors that have generated the Malthusian trap. What accounts for the epoch of stagnation that

has characterized most of human history? Why did episodes of technological progress in the pre-industrial era fail to generate sustained economic growth? Why has population growth counterbalanced the expansion of resources per capita that could have been generated by technological progress?

Moreover, the theory uncovers the forces that triggered the take-off from stagnation to growth. What is the origin of the sudden spurt in the growth rates of income per capita and population during the course of industrialization? What was the source of the striking reversal in the positive relationship between income per capita and population growth that existed throughout most of human history? Would the transition to the modern state of sustained economic growth have been feasible without the decline in population growth? What are the hurdles faced by less developed economies in their attempts to transition to a sustained-growth regime?

Further, Unified Growth Theory sheds new light on the origins of the perplexing divergence in income per capita across developed and less developed regions in the past two centuries. What accounts for the sudden take-off from stagnation to growth among some countries in the world and the persistent stagnation in others? Why has the positive link between income per capita and population growth reversed its course in some economies but not in others? Has the transition to a state of sustained economic growth in advanced economies adversely affected the process of development in less developed ones? Have variations in prehistoric biogeographical factors had a persistent effect on the composition of human capital and economic development across the world?

## **1.1 Toward a Unified Theory of Economic Growth**

Non-unified theories of economic growth have been instrumental in advancing the understanding of the role that technological progress and the accumulation of factors of production have played in the modern era of economic growth. Nevertheless, they are inconsistent with the qualitative aspects of the growth process over most of human existence, and they fail to identify the forces that triggered the take-off from stagnation to sustained economic growth—insights that are instrumental for understanding the contemporary growth process and the origins of the great divergence in income per capita over the past two centuries.

The preoccupation of non-unified theories of economic growth with the growth process of developed economies in the past century and of less developed economies in the past few decades has become harder to justify in light of the disparity between the main features of the modern growth era and those that have characterized the growth process over most of human existence. It has

## 4 • Chapter 1

become evident that as long as growth theory rests on distinct and disjoint theories to characterize the process of development during the Malthusian Epoch and the Modern Growth Regime, the understanding of the contemporary growth process will be limited and distorted.<sup>1</sup> “It is as though an artist were to gather the hands, feet, head and other members for his images from diverse models, each part perfectly drawn, but not related to a single body, and since they in no way match each other, the result would be monster rather than man” (Copernicus quoted in Kuhn [1957, p. 137]).

The advancement of Unified Growth Theory has been fueled by the conviction that the understanding of global variation in economic development would be fragile and incomplete unless the prevailing theory of economic growth reflects the principal driving forces behind the entire process of development and captures the central role that historical factors have played in bringing about the current disparities in living standards.<sup>2</sup> Moreover, it has been fostered by the realization that a comprehensive understanding of the hurdles faced by less developed economies would remain obscure unless the factors that facilitated the transition of the currently developed economies from stagnation to growth could be identified and modified to account for the differences in the growth structure of less developed economies in an increasingly interdependent world.

Unified Growth Theory provides a fundamental framework of analysis for the evolution of individuals, societies, and economies over the entire course of human history. The theory captures in a single analytical framework the main characteristics of the process of development: (i) the epoch of Malthusian stagnation that has characterized most of human history; (ii) the escape from the Malthusian trap and the associated spike in the growth rates of income per capita and population; (iii) the emergence of human capital formation in the process of development; (iv) the onset of the demographic transition; (v) the

<sup>1</sup>The evolution of theories in older scientific disciplines suggests that theories founded on the basis of a subset of the existing observations may be attractive in the short run but are nonrobust and nondurable in the long run. For instance, classical thermodynamics, which lacked microfoundations, was ultimately superseded by the micro-based statistical mechanics. Moreover, attempts to develop unified theories in physics have been based on the conviction that all physical phenomena should eventually be explainable by some underlying unity. In particular, Unified Field Theory proposes to unify by a set of general laws the four distinct forces that are known to control all observed interactions in matter: electromagnetism, gravitation, the weak force, and the strong force.

<sup>2</sup>Clearly, the understanding of the contemporary world would be limited and incomplete in the absence of a historical perspective. However, the intensity of recent explorations of the interaction between economic development and economic history could be attributed to increasing frustration with the failure of the ahistorical branch of growth theory to capture some of the most fundamental aspects of the growth process.

contemporary era of sustained economic growth; and (vi) the divergence in income per capita across countries.<sup>3</sup>

The theory unveils the principal economic forces that have generated the remarkable transition from stagnation to growth and underlines their significance for understanding the contemporary growth process of both developed and less developed economies. Moreover, it sheds light on the role of historical and pre-historical characteristics in the divergence of income per capita across regions of the world in the past two centuries.

Unified Growth Theory suggests that the transition from stagnation to growth has been an inevitable by-product of the process of development. It argues that the inherent Malthusian interaction between the rate of technological progress and the size and composition of the population accelerated the pace of technological progress and ultimately raised the importance of education in coping with the rapidly changing technological environment.<sup>4</sup> The rise in industrial demand for education brought about significant reductions in fertility rates. It enabled economies to divert a larger share of the fruits of factor accumulation and technological progress to the enhancement of human capital formation and income per capita, paving the way for the emergence of sustained economic growth.

The theory further explores the dynamic interaction between human evolution and the process of economic development and advances the hypothesis that the forces of natural selection played a significant role in the evolution of the world economy from stagnation to growth. The Malthusian pressures have acted as the key determinant of population size and conceivably, via natural selection, have shaped the composition of the population as well. Lineages of individuals whose traits were complementary to the economic environment generated higher levels of income, and thus a larger number of surviving offspring, and the gradual increase in the representation of their traits in the population contributed to the process of development and the take-off from stagnation to growth.

<sup>3</sup>The term “Unified Growth Theory” was coined by Galor (2005) to categorize theories of economic growth that capture the entire growth process in a single framework of analysis. The only unified theory of economic growth that captures the endogenous evolution of population, technology, human capital, and income per capita over the entire course of economic development, while generating both a spontaneous transition from Malthusian stagnation to sustained growth and a great divergence has been developed by Galor (2005, 2010), based on Galor and Weil (1999, 2000), Galor and Moav (2002), and Galor and Mountford (2008). This theory therefore is the central pillar of this book.

<sup>4</sup>The increased demand for human capital has not necessarily resulted in an increase in the rate of return on human capital due to institutional changes (e.g., the provision of public education) that lowered the cost of investment in human capital and facilitated a massive increase in the supply of education.

## 1.2 Origins of Global Disparity in Living Standards

Unified Growth Theory sheds light on the notable divergence in income per capita across the globe during the past two centuries. The theory advances the understanding of three fundamental aspects of comparative economic development. First, it identifies the factors that have governed the transition from stagnation to growth and have thus contributed to the observed worldwide differences in economic development. Second, it highlights the persistent effects that variations in historical and prehistorical conditions have had on the composition of human capital and economic development across countries. Finally, it uncovers the forces that have sparked the emergence of convergence clubs, and it explores the characteristics that have determined the association of different economies with each club.

### 1.2.1 Catalysts for the Engine of Transition from Stagnation to Growth

The first layer of Unified Growth Theory explores the underlying forces that have determined the timing and pace of the transition from an epoch of Malthusian stagnation to an era of sustained economic growth and have thus contributed to the disparity in economic development across countries. Country-specific characteristics that have affected the intensity of the pivotal interaction between the rate of technological progress and the size and composition of the population have generated variations in the transition from stagnation to growth and contributed to the gap in income per capita across countries.

Variation in rates of technological progress has reinforced the differential pace of the emergence of demand for human capital, the onset of the demographic transition, and the shift from stagnation to growth, and has thus contributed to the divergence in income per capita in the past two centuries. In particular, worldwide variation in the pace of technological progress has been triggered by cross-country differences in (i) the stock of knowledge and its rate of creation and diffusion among members of society; (ii) the level of protection of intellectual property rights, its positive effect on the incentive to innovate, and its adverse effect on the proliferation of existing knowledge; (iii) financial constraints and the level of competitiveness of the innovation sector; (iv) the composition of cultural and religious attributes and their effects on knowledge creation and diffusion; (v) the composition of interest groups in society and their incentives to block or promote technological innovations; (vi) the level of human diversity and the degree to which it complements the implementation and advancement of new technological paradigms; (vii) the propensity to trade and its effect on technological diffusion; and (viii) the abundance of natural resources essential for an imminent technological paradigm.

Once the technologically driven demand for human capital emerged in the second phase of industrialization, the prevalence of characteristics conducive to human capital formation has determined the swiftness of its accumulation, the timing of the demographic transition, the pace of the transition from stagnation to growth, and the observed distribution of income in the world economy. Thus, variations in country-specific characteristics that have contributed to human capital formation have differentially affected the timing and pace of the transition from agriculture to industry and comparative economic development as a whole.

In particular, global variation in human capital formation has been influenced by cross-country differences in (i) the prevalence of human capital-promoting institutions or policies (e.g., the availability, accessibility, and quality of public education); (ii) the ability of individuals to finance the cost of education as well as the foregone earnings associated with schooling; (iii) the impact of the level of inequality and of the degree of credit market imperfections on the extent of underinvestment in education; (iv) the stock of knowledge in society and its effect on the productivity of investment of human capital; (v) the composition of cultural and religious groups in a society and their effects on the incentives of individuals to invest in human capital; (vi) the impact of geographical attributes on health and thus human capital formation; (vii) the propensity to trade and the patterns of comparative advantage with respect to the production of skill-intensive goods; and (viii) preferences for educated offspring that may reflect cultural attributes, the composition of religious groups and social status associated with education.

### **1.2.2 Persistence of Prehistorical Biogeographical Conditions**

In its second layer, Unified Growth Theory highlights the direct persistent effect that deep-rooted factors, determined as early as tens of thousands years ago, have had on the course of comparative economic development from the dawn of human civilization to the modern era.

The theory captures the thesis that part of the differences in the process of development across the globe can be traced to biogeographical factors that led to regional variation in the timing of the Neolithic Revolution (Diamond, 1997). According to this thesis, favorable biogeographical endowments that contributed to the emergence of agriculture gave some societies the early advantage of operating a superior production technology and generating resource surpluses. They permitted the establishment of a non-food-producing class, whose members were crucial for the development of written language and science and for the formation of cities, technology-based military powers, and nation states. The early dominance of these societies persisted throughout history, being further sustained by geopolitical and historical processes, such as colonization. The significance of the timing of agricultural transitions for precolonial economic

## 8 • Chapter 1

development has been confirmed empirically, although evidence appears to suggest that over the past five hundred years the initial dominance brought about by an earlier transition to agriculture has dissipated.

Moreover, the theory is consistent with the thesis that the exodus of modern humans from Africa, nearly a hundred thousand years ago, appears central to understanding comparative economic development across the globe (Ashraf and Galor, 2009). In the course of the exodus of *Homo sapiens* out of Africa, variation in migratory distance from the cradle of humankind to settlements around the globe affected the level of genetic diversity and has had a long-lasting, hump-shaped effect on the pattern of comparative economic development that cannot be captured by contemporary geographical, institutional, and cultural factors. While the intermediate level of genetic diversity prevalent among Asian and European populations has been conducive to development, the high degree of diversity among African populations and the low degree among Native American populations have acted as detrimental forces in the development of these regions.

### 1.2.3 Convergence Clubs

In its third layer, Unified Growth Theory advances the understanding of the forces that have contributed to the existence of multiple growth regimes and the emergence of convergence clubs (i.e., groups of countries among which the disparity in income per capita tends to narrow over time). The theory attributes these phenomena to variation in the position of economies across the distinct phases of development. It suggests that the differential timing of take-offs from stagnation to growth has segmented economies into three fundamental growth regimes: slowly growing economies in the vicinity of a Malthusian steady state, fast growing countries in a sustained-growth regime, and a third group of economies in transition from one regime to the other. Moreover, it suggests that the presence of multiple convergence clubs may reflect a temporary state, as endogenous forces may ultimately permit members of the Malthusian club to shift their positions and join the members of the sustained-growth club.