1

Explaining Levels of Colonialism and Postcolonial Development

Much of the developing world was dragged into the modern era by colonialism. However one judges it, this is a historical legacy with which all scholars interested in the political economy of development, especially political economy over the long duration, must come to terms.

– Atul Kohli

Comparative historical analysis serves as an ideal strategy for mediating between theory and history. Provided that it is not mechanically applied, it can prompt both theoretical extensions and reformulations, on the one hand, and new ways of looking at concrete historical cases, on the other.

– Theda Skocpol

Colonialism was a great force of change in the modern era. From the Americas to the Asian and African continents, colonial expansion brought Europeans and their institutions around the world. It stirred nationalist sentiments and intensified competition within the European core; and the colonies provided an outlet for citizens who sought or were compelled to pursue a new life overseas. By disseminating people and institutions, moreover, colonialism forever changed the structure of trade and production within what had been an almost exclusively European commercial system. Nothing less than a genuinely worldwide system of states and trade was born out of colonialism. In the judgment of Karl Marx and Friedrich Engels, “the colonization of America, trade with the colonies, the increase in the means of exchange and in commodities generally, gave to commerce, to navigation, to industry, an impulse never before known.”[1]

But the consequences of colonialism were, of course, felt most deeply in those territories and by those people subjected to this intervention. Preexisting societies were traumatically rearranged and sometimes destroyed. This was as true for precolonial societies renowned as great civilizations – such as the Aztec and Inca empires in the Americas – as it was for less well remembered precolonial chiefdoms and hunter-gatherer groups. The institutions established during colonialism, furthermore, exhibited over-time effects, whether directly through their own persistence or indirectly through the actors and processes that they brought into being. Colonial authorities and settlers almost invariably imposed administrative and political boundaries that subsequently became the basis – or at least the critical starting point – for demarking the borders of new nation-states. Within those borders, colonialism wrought economic arrangements and state machineries that structured productive activity and that affected the level of prosperity for the societies that remained. Colonial powers also introduced new cultural distinctions and modes of interest representation upon which subsequent social-stratification systems and political regimes were built.

In modern world history, colonialism is marked by a state's successful claim to sovereignty over a foreign land. Under a colonial arrangement, major actors in the interstate system at least implicitly recognize the colonizing state's patrimony over the occupied territory; and this recognition is founded in part upon the colonizing state's proven ability to implant settlers, maintain governance structures, and extract resources in the territory.[2] This definition makes colonialism a more thoroughgoing form of territorial control than what is conventionally thought of as imperialism or, for that matter, economic and political dependency. While imperialism and dependency entail asymmetrical relationships between states, they do not inherently involve a loss of sovereignty or even the insertion of governance structures under the control of a metropolis. Colonialism is set apart from these other kinds of interstate domination above all because it renders subordinate (or makes obsolete) all prior political entities that could once lay claim to – and perhaps back up through coercive means, if necessary – final authority over territorial inhabitants. So thoroughgoing is colonial domination that other international actors must treat the metropolis as the de facto political representative of the occupied land.

Though delimited in these ways, the intersocietal relationships that qualify as colonialism are nevertheless numerous and varied. According to David B. Abernethy, modern European colonialism was carried out by eight different countries and encompassed the territories of what became 125 different nation-states at one time or another.[3] Most of western Europe was sooner or later engaged in colonial projects, and most areas in the rest of the world became objects of these projects.

The undeniable, paramount importance of colonialism beckons social scientists to study the causes and consequences of this historical process. But what is the most fruitful way for researchers to proceed with their explanatory investigations? One worthy approach is to explore why colonialism occurred in certain places and at certain times;[4] another is to generalize broadly about the effects of colonialism for Europe, for the non-European regions, or for the world system as a whole.[5] Yet some scholars – especially comparatively and historically oriented social scientists – will always be drawn to questions about the sources of alternative modes of colonialism and their legacies for nation-states. Why did similar or different forms of colonialism arise within the borders of what are now sovereign states? What were the long-run consequences of particular kinds of colonialism for the national citizens of those states? Answering these questions, as this book seeks to do, requires treating territories corresponding to modern nation-states as the basic units of analysis, including during historical episodes prior to their appearance as sovereign entities.

In this introductory chapter, I develop a historical-institutional framework for studying colonialism and postcolonial development. The framework offers both principles for the causal analysis of particular cases and a general theory consisting of propositions to be applied broadly to former colonial cases.[6] The *principles of analysis* are developed in critical dialogue with the two currently dominant orientations for explaining national development: geographic and institutional perspectives. Concerning the geographic perspectives, I argue that they too often assume that features of the natural environment directly affect development; these approaches fail to give appropriate weight to mediating institutions. And they provide little guidance for theorizing the time-variant effects of geography and the ways in which geographic effects depend on the arrangement of already-existing institutions. The reigning institutional perspectives, for their part, are much too prone to treat institutions as devices that merely coordinate behavior, rather than as power-implicating instruments that unevenly distribute resources and constitute collective actors. Moreover, existing institutional work on colonialism in particular has suffered from vague conceptions of institutions and overly generalized understandings of their effects that cannot make sense of basic historical facts about particular cases. In response to the limitations of both geographic and institutional perspectives, I formulate alternative principles to be used in historically grounded, case-oriented explanation.

The *general theory* lays out orienting hypotheses that are intended to apply broadly across cases. These hypotheses seek to explain variations in both (1) levels of colonialism, which refer to the extent of settlement and institutional implantation in colonized territories, and (2) levels of postcolonial economic and social development, which capture national differences in prosperity and human well-being. To account for variations in levels of colonialism, the theory highlights the interaction between the institutions of the colonizing nation and the institutions of the precolonial society. Contrasts in the political-economic institutions of the European colonizers are hypothesized as essential for grasping why these nations often pursue quite distinct modes of colonialism in similar precolonial societies. Likewise, variations in the institutional complexity of precolonial societies are crucial for understanding why European colonizers with similar political economies follow contrasting modes of colonization. Thus, to explain levels of colonialism, I examine the "fit" between the institutions of the colonizing nation and those of the colonized territory.

To explain levels of postcolonial economic and social development, the general theory calls attention to the interaction between a territory's level of colonialism and the political economy of the colonizing European nation. The long-run consequences of a given level of colonialism vary across colonizers with different political economies. Extensive colonialism featuring heavy settlement and institution building is not expected always or usually to leave behind rich and egalitarian new countries (as some analysts suggest). Nor is extensive colonialism predicted to produce impoverished and conflict-ridden states (as others suggest). Rather, the effect of a given level of colonialism for long-run development depends on the political-economic institutions of the colonizing power.

In the core chapters of this book, both the principles of analysis and the general theory are employed in the analysis of Spanish America. My approach throughout is to start with the general theory but supplement it with ideas anticipated by the principles of analysis. In this chapter, I do not yet summarize the findings that are derived from this comparative-historical analysis; that summary is found in the concluding chapter. However, I do advance reasons why a close focus on the Spanish American countries makes good methodological sense.

The discussion so far merely anticipates arguments that need to be developed at length. I begin this task in the next section by addressing conceptual issues concerning the final outcome under investigation: relative levels of development.

RELATIVE LEVELS OF DEVELOPMENT

Why are some countries more developed than others? This basic question, posed again and again by thoughtful analysts, is an inquiry about *relative* levels of development. It asks why some nations have relatively higher (or lower) positions within the overall hierarchy of development. In a world system in which countries exhibit enormously different economic and social conditions, identifying the causes of relative levels of development must be regarded as one of the most fundamental tasks of the social sciences.

When treated as a theoretical construct, *development* may be defined, following Amartya Sen, as "the expansion of the 'capabilities' of people to lead the kind of lives they value – and have reason to value."[7] Under this definition, development is the process through which individuals are empowered to meet their objectively justifiable interests. Tangible improvements in wealth and social welfare are the crucial, though not the exclusive, components of development. Growth of the economy and real income, especially at lower and middle levels of prosperity, expands capabilities by providing access to basic goods and services that all people have reason to value. Advances in education and health enable individuals to lead longer, more informed, and enjoyable lives – things that human beings inherently and rightly seek. And both economic and social enhancements provide opportunities for individuals to engage in collective activities that are intrinsically

Table 1.1. Approximate levels of economic development for the Spanish American countries, late nineteenth century to present

Higher level	Intermediate level	Lower level
1. Argentina	5. Mexico	9. Paraguay
2. Uruguay	6. Costa Rica	10. Guatemala
3. Chile	7. Colombia	11. El Salvador
4. Venezuela	8. Peru	12. Ecuador
		13. Nicaragua
		14. Bolivia
		15. Honduras

worthy and often essential for securing other kinds of freedoms, including political democracy.[8]

Though not the only aspects of development, the expansion of wealth and human welfare – what I call economic and social development – are thus among its most important ones. They will be the focus of this book. I will inquire specifically about the causes of relative levels of economic and social development among the mainland Spanish American countries (see Tables 1.1 and 1.2) and, to a lesser degree, among the countries colonized by Britain and Portugal (see Chapter 7). The national differences in relative levels of development presented in Tables 1.1 and 1.2 will be described further and substantiated at length in light of the available data. The focus for now, however, is more general: understanding relative levels of development as an object of explanation.

The Stability of Relative Levels of Development

Relative levels of development – unlike absolute levels – tend to persist over time. Countries that now feature higher levels of wealth (or better

Table 1.2. Approximate levels of social development for the Spanish American countries, late nineteenth century to present

Higher level	Intermediate level	Lower level
1. Uruguay	5. Paraguay	9. Ecuador
2. Argentina	6. Colombia	10. Peru
3. Costa Rica	7. Venezuela	11. Honduras
4. Chile	8. Mexico	12. El Salvador
		13. Nicaragua
		14. Bolivia
		15. Guatemala

performance on health, education, and other social indicators) usually have been among the wealthier (or healthier, more educated, and so on) nations for many decades or even centuries. Similarly, countries that are now relatively poor have usually been relatively poor for a long time, often for as long as they have been countries. To be sure, there are some exceptions, such as the once-peripheral case of South Korea, which achieved a comparatively high level of development during the second half of the twentieth century.[9] But the *majority* of countries tend to fall into relatively stable positions within the global hierarchy of development.

Social scientists have long explored puzzles related to the persistence of relative levels of national development. Much of this work takes its cue from efforts to explain why certain countries – especially those in western Europe – emerged at the top of the development hierarchy in the first place. The question, "Why Europe?" (i.e., why Europe or a particular region in Europe was the birthplace of capitalist commercialization and industrialization) originally preoccupied the classical theorists and subsequently many of the most insightful scholars of macrohistorical inquiry.[10] Contemporary analysts have progressed to the point of identifying historical sources of variations in socioeconomic development among the European countries themselves, as well as other advanced capitalist nations.[11] What is more, whole paradigms of scholarship have explored the inability of less-developed countries to catch up with or replicate the developmental experiences of Europe.[12] From this work, we now know much about why nations in Africa, Asia, and Latin America have not been able to sustain high growth over extended periods of time. Likewise, we now have plausible theories to explain why a few extraordinary cases have been able to break with the pattern and achieve significant improvements in their relative level of development.

What tends to be missing, however, are well-developed theories about the origins of varying levels of development among the "non-European" countries. We still know precious little about why countries in Africa, Asia, and Latin America exhibit higher or lower levels of socioeconomic development relative to one another. Why, for example, is Uruguay so much wealthier than Bolivia? Why is Botswana a model of development in Africa whereas Sierra Leone is not? Can we identify the "original" sources of relative levels of development in such diverse countries as Uruguay, Bolivia, Botswana, and Sierra Leone using a single set of explanatory principles and a general theoretical framework? This book proposes that we can.

Rates versus Relative Levels of Development

Stability marks relative levels of national development because differences in *rates* of growth and social progress for countries are *not* stable. Episodes of *sustained* high (or low) growth rates are rare in the contemporary world economy. Likewise, dramatic improvements in social welfare (e.g., literacy, life expectancy, and education) are often held hostage by uneven growth rates – for although long-run social development is not simply a derivative of economic growth, it is shaped by such growth.[13]

Evidence for the instability of rates of economic growth is not merely impressionistic. In one important study, William Easterly and colleagues find that the correlation for national growth rates across decades ranges from 0.1 to 0.3, and hence that the performance of countries in one decade only weakly predicts performance in the next decade. These authors conclude, "With a few famous exceptions, the same countries do not do well period after period; countries are 'success stories' one period and disappointments the next."[14] In his review of the "new growth" literature, Jonathan Temple likewise warns against using growth rates as a basis for estimating long-run performance: "Frequently countries have done well for short periods, only for growth to collapse later on."[15]

A concern with *relative levels* of development differs in basic ways from a focus on *rates* of development. Whereas rates highlight variations in performance during specified intervals of time, relative levels cast the spotlight on differences that tend to endure across any given period. Precisely because relative levels of development are so persistent, one must explain them historically: their origins rest at some point *before* countries stabilize their positions in the hierarchy of development. By contrast, the causes driving rates of development are typically far less historically rooted, and they may include such short-range factors as natural disasters, business cycles, and public policies.

Nevertheless, changes do occur even for relative levels of economic development, and recent work on global inequality allows us to generalize about these changes. Most notably, richer countries tend to grow faster than poorer countries, producing national-level income divergence in the world economy.[16] Specifically, at least in the post–World War II era, upper-middle-income nations have had the highest growth rates, and lower-income nations have had the lowest. As a result, the global trend has been toward income divergence, even though there has been some convergence among the set of wealthy nations. To be sure, this trend assumes that countries are not weighted for population[17] and that control variables are not introduced that mediate the effect of the initial level of economic development.[18] Within these constraints, wealthier countries grow at higher rates than poorer countries, thereby following a pattern of divergence that has existed at least since the Industrial Revolution.

One implication of this research is that it is useful to distinguish between, on the one hand, the lower-income and middle-income countries of the world and, on the other hand, the upper-middle-income and higher-income countries. The latter group is simply pulling away from the former. However, among the lower- and middle-income countries, neither sustained convergence nor divergence appears to be taking place. Rather, when treated as a single population, the lower- and middle-income countries have been stable in their relative levels of development.

This last point is of essential importance, because the vast majority of former European colonies in Africa, Asia, and Latin America are now lower- or middle-income countries – the population that has exhibited the most stability in relative levels of development. Contemporary differences in levels of development among these countries are *not* primarily the result of diverging rates of growth or social progress since independence. Rather, the bulk of their differences can be attributed to the fact that they *started out* with different levels of development. If one wishes to explain why they have contrasting relative levels of development today, therefore, the main task is to locate the *causes of their initial differences*. This task requires one to pursue historical analysis and avoid dwelling on the ups and downs that may have occurred in more contemporary periods.

The Infrequency of Sustained Progress

A country will experience a significant change in its relative level of development only when the variables that affect rates of growth (or rates of social progress) assume atypical values for prolonged periods of time. This means that we can learn about the sources of stability in relative levels of development by elucidating why it is so rare for countries to maintain certain values on key variables (or combinations of variables) over the long run. To address this issue, however, we need to locate those specific variables that actually affect rates of progress.

Interestingly, the econometric new growth literature is not particularly useful for pinpointing factors that cause trajectories of high or low growth rates. Although well over fifty variables have been found to be significantly associated with growth, the effects of most of them are fragile.[19] Slight changes in the control variables or indicators of the regression model overturn conclusions about their impact. In addition, those variables that are robustly related to growth rates are small in number and closely tied to the phenomenon of development itself. For example, one important sensitivity analysis found that only four variables – initial level of gross domestic product (GDP) per capita, investment spending, population growth rate, and formal education – are positively correlated with growth across different model specifications. As Ross Levine and David Renelt note in reporting this finding, these factors are in part measures or symptoms of development itself; or, as Douglass C. North and Robert P. Thomas put it in reaching similar findings, “The factors we have listed (innovation, economies of scale, education, capital accumulation, etc.) are not causes of growth; they *are* growth.”[20]

Comparative-historical studies that analyze a small number of countries have arguably been more successful at explaining exceptional economic performance among the less developed countries. Various works on the "developmental state" – a research program pioneered by scholars such as Alice Amsden, Peter B. Evans, Dietrich Rueschemeyer, and Robert Wade – suggest that sustained growth is driven by a state apparatus with centralized power, corporate coherence and autonomy, and dense social ties to key domestic groups, especially capitalists.[21] Likewise, more "society-centered" analyses – formulated by scholars such as Joel S. Migdal, Robert Putnam, and Amartya Sen – argue that sustained development requires the removal of patron-client relations and other hierarchical forms of domination.[22] Together, these literatures help make sense of the economic "miracles" in Korea and Taiwan and cases of great social progress such as Kerala, India.[23] They also suggest that a prolonged development failure (e.g., movement from a relatively high per capita income to a relatively low per capita income) would be precipitated by the disintegration of existing state machineries and the degeneration of positive-sum social arrangements into zero-sum conflict.

With these insights at hand, we can begin to understand why so few countries have managed to achieve sustained high rates of growth and social progress. Most basically, a developmental state and a society lacking rigid hierarchical bonds of dependence are extremely difficult to construct. In all developing countries, powerful actors derive resources from institutional arrangements that promote inequalities. To dislodge these institutions and actors, a fundamental transformation of the basic distribution of power within society is often required. Historically, such transformations have occurred largely in conjunction with foreign interventions, land reforms, revolutions, or wars that unseat dominant economic actors.[24] Other possible conditions are copious international economic aid, an auspicious global trade environment, low levels of foreign direct investment, a delayed transition to mass politics, and powerful legitimating ideologies addressing external threats.[25] Whatever the exact recipe for developmental states and egalitarian societies, the ingredients are not often present or found in the right quantities and sequence.

In sum, the stability of relative levels of national development is a function of the rarity of those conditions that promote sustained high (or low) rates of development. Once countries arrive at their particular relative levels of development, they generally stay more or less at those levels unless something exceptional happens. If one wishes to explain differences in levels of development, therefore, one should *not* center explanatory attention on processes that take place after a country has already settled into a given level of development. The real explanatory challenge is, instead, identifying the *historical origins* of the initial levels of development.

We have now developed an understanding of what it is that this book ultimately seeks to explain. It is time to examine the ways in which different kinds of explanatory factors will or will not enter the analysis.

EXISTING PERSPECTIVES AND THEIR LIMITATIONS: NEW PRINCIPLES FOR CASE ANALYSIS

Competing theoretical perspectives emphasizing either geography or institutions strongly influence contemporary debates about the origins of development. These perspectives propose fundamentally different kinds of causes, directing attention either to physical features of the landscape or to socially constructed rules that guide behavior. Despite these differences, both perspectives as currently formulated share certain problematic ahistorical assumptions about how to explain levels of development. As a corrective, this section critiques these limitations and formulates alternatives.

Geographical Explanations

The idea that geographical endowments can explain levels of development is not new. Niccolò Machiavelli, Charles de Secondat Montesquieu, and Arnold J. Toynbee all embraced this orientation.[26] In our times, the hypothesis is associated with brilliant scholars such as Jared Diamond, David Landes, and Jeffrey D. Sachs.[27] Their work has established beyond any reasonable doubt that several features of geography are correlated with contemporary levels of national development. Even the simple variable of distance from the equator performs reasonably well as a predictor of current levels of GDP per capita: countries that are more distant from the equator tend to be richer. Yet for the purposes of actually *explaining* levels of development, as opposed to identifying features that are correlated with development, we must ask questions about *how* geography affects development, *when* geography affects development, and *what specific features* of geography affect development. For each of these questions, the existing literature provides insights, but these insights need to be enriched by a more historically contextualized approach if adequate explanation is to be achieved.

Many geographical features are virtually permanent, preceding in time all other potentially relevant causal factors. As such, geographical determinants can often be treated as fully exogenous causes – the "immovable movers" in a causal argument. One still needs to inquire, however, whether even enduring geographic features *directly* shape levels of development or whether their effects work primarily or exclusively through intermediary causal processes. In recent years, several economists have addressed this issue by exploring the effect of geographic variables, such as distance from the equator, while controlling for institutional variables (e.g., the extent of rule of law). The title of an article by Dani Rodrik, Arvind Subramanian, and Francesco Trebbi suggests the major findings: "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." Rodrik and collaborators conclude that "the quality of institutions trumps everything else. Once institutions are controlled for . . . geography has at best weak