"Truly awesome . . . brilliant in its simplicity and power."—Steven Levitt, coauthor of *Freakonomics*

THE ORIGINS OF
POWER, PROSPERITY, AND POVERTY

WHY NATIONS FAIL

DARON ACEMOGLU JAMES A. ROBINSON

PRAISE FOR *Why Nations Fail*

"Acemoglu and Robinson have made an important contribution to the debate as to why similar-looking nations differ so greatly in their economic and political development. Through a broad multiplicity of historical examples, they show how institutional developments, sometimes based on very accidental circumstances, have had enormous consequences. The openness of a society, its willingness to permit creative destruction, and the rule of law appear to be decisive for economic development."

—Kenneth J. Arrow, Nobel laureate in economics, 1972

"The authors convincingly show that countries escape poverty only when they have appropriate economic institutions, especially private property and competition. More originally, they argue countries are more likely to develop the right institutions when they have an open pluralistic political system with competition for political office, a widespread electorate, and openness to new political leaders. This intimate connection between political and economic institutions is the heart of their major contribution, and has resulted in a study of great vitality on one of the crucial questions in economics and political economy."

—Gary S. Becker, Nobel laureate in economics, 1992

"This important and insightful book, packed with historical examples, makes the case that inclusive political institutions in support of inclusive economic institutions is key to sustained prosperity. The book reviews how some good regimes got launched and then had a virtuous spiral, while bad regimes remain in a vicious spiral. This is important analysis not to be missed."

—Peter Diamond, Nobel laureate in economics, 2010

"For those who think that a nation's economic fate is determined by geography or culture, Daron Acemoglu and Jim Robinson have bad news. It's manmade institutions, not the lay of the land or the faith of our forefathers, that determine whether a country is rich or poor. Synthesizing brilliantly the work of theorists from Adam Smith to Douglass North with more recent empirical research by economic historians, Acemoglu and Robinson have produced a compelling and highly readable book."

—Niall Ferguson, author of *The Ascent of Money*

"Acemoglu and Robinson—two of the world's leading experts on development—reveal why it is not geography, disease, or culture that explain why some nations are rich and some poor, but rather a matter of institutions and politics. This highly accessible book provides welcome insight to specialists and general readers alike."

—Francis Fukuyama, author of *The End of History and the Last Man* and *The Origins of Political Order*

"A brilliant and uplifting book—yet also a deeply disturbing wake-up call. Acemoglu and Robinson lay out a convincing theory of almost everything to do with economic development. Countries rise when they put in place the right pro-growth political institutions and they fail—often spectacularly—when those institutions ossify or fail to adapt. Powerful people always and everywhere seek to grab complete control over government, undermining broader social progress for their own greed. Keep those people in check with effective democracy or watch your nation fail."

—Simon Johnson, coauthor of *13 Bankers* and professor at MIT Sloan

"Two of the world's best and most erudite economists turn to the hardest issue of all: why are some nations poor and others rich? Written with a deep knowledge of economics and political history, this is perhaps the most powerful statement made to date that 'institutions matter.' A provocative, instructive, yet thoroughly enthralling book."

—Joel Mokyr, Robert H. Strotz Professor of Arts and Sciences and Professor of Economics and History, Northwestern University

"In this delightfully readable romp through four hundred years of history, two of the giants of contemporary social science bring us an inspiring and important message: it is freedom that makes the world rich. Let tyrants everywhere tremble!"

—Ian Morris, Stanford University, author of *Why the West Rules—for Now*

"Imagine sitting around a table listening to Jared Diamond, Joseph Schumpeter, and James Madison reflect on more than two thousand years of political and economic history. Imagine that they weave their ideas into a coherent theoretical framework based on limiting extraction, promoting creative destruction, and creating strong political institutions that share power, and you begin to see the contribution of this brilliant and engagingly written book."

—Scott E. Page, University of Michigan and Santa Fe Institute

"In this stunningly wide-ranging book, Acemoglu and Robinson ask a simple but vital question, why do some nations become rich and others remain poor? Their answer is also simple—because some polities develop more inclusive political institutions. What is remarkable about the book is the crispness and clarity of the writing, the elegance of the argument, and the remarkable richness of historical detail. This book is a must-read at a moment when governments across the Western world must come up with the political will to deal with a debt crisis of unusual proportions."

—Steven Pincus, Bradford Durfee Professor of History and International and Area Studies, Yale

University

"It's the politics, stupid! That is Acemoglu and Robinson's simple yet compelling explanation for why so many countries fail to develop. From the absolutism of the Stuarts to the antebellum South, from Sierra Leone to Colombia, this magisterial work shows how powerful elites rig the rules to benefit themselves at the expense of the many. Charting a careful course between the pessimists and optimists, the authors demonstrate history and geography need not be destiny. But they also document how sensible economic ideas and policies often achieve little in the absence of fundamental political change."

—Dani Rodrik, Kennedy School of Government, Harvard University

"This is not only a fascinating and interesting book: it is a really important one. The highly original research that Professors Acemoglu and Robinson have done, and continue to do, on how economic forces, politics, and policy choices evolve together and constrain each other, and how institutions affect that evolution, is essential to understanding the successes and failures of societies and nations. And here, in this book, these insights come in a highly accessible, indeed riveting form. Those who pick this book up and start reading will have trouble putting it down."

—Michael Spence, Nobel laureate in economics, 2001

"This fascinating and readable book centers on the complex joint evolution of political and economic institutions, in good directions and bad. It strikes a delicate balance between the logic of political and economic behavior and the shifts in direction created by contingent historical events, large and small, at 'critical junctures.' Acemoglu and Robinson provide an enormous range of historical examples to show how such shifts can tilt toward favorable institutions, progressive innovation, and economic success or toward repressive institutions and eventual decay or stagnation. Somehow they can generate both excitement and reflection."

—Robert Solow, Nobel laureate in economics, 1987

Why Nations Fail

THE ORIGINS OF POWER, PROSPERITY, AND POVERTY

Daron Acemoglu and
James A. Robinson

Crown Publishers • New York

PREFACE

THIS BOOK IS about the huge differences in incomes and standards of living that separate the rich countries of the world, such as the United States, Great Britain, and Germany, from the poor, such as those in sub-Saharan Africa, Central America, and South Asia.

As we write this preface, North Africa and the Middle East have been shaken by the “Arab Spring” started by the so-called Jasmine Revolution, which was initially ignited by public outrage over the self-immolation of a street vendor, Mohamed Bouazizi, on December 17, 2010. By January 14, 2011, President Zine El Abidine Ben Ali, who had ruled Tunisia since 1987, had stepped down, but far from abating, the revolutionary fervor against the rule of privileged elites in Tunisia was getting stronger and had already spread to the rest of the Middle East. Hosni Mubarak, who had ruled Egypt with a tight grip for almost thirty years, was ousted on February 11, 2011. The fates of the regimes in Bahrain, Libya, Syria, and Yemen are unknown as we complete this preface.

The roots of discontent in these countries lie in their poverty. The average Egyptian has an income level of around 12 percent of the average citizen of the United States and can expect to live ten fewer years; 20 percent of the population is in dire poverty. Though these differences are significant, they are actually quite small compared with those between the United States and the poorest countries in the world, such as North Korea, Sierra Leone, and Zimbabwe, where well over half the population lives in poverty.

Why is Egypt so much poorer than the United States? What are the constraints that keep Egyptians from becoming more prosperous? Is the poverty of Egypt immutable, or can it be eradicated? A natural way to start thinking about this is to look at what the Egyptians themselves are saying about the problems they face and why they rose up against the Mubarak regime. Noha

Hamed, twenty-four, a worker at an advertising agency in Cairo, made her views clear as she demonstrated in Tahrir Square: "We are suffering from corruption, oppression and bad education. We are living amid a corrupt system which has to change." Another in the square, Mosaab El Shami, twenty, a pharmacy student, concurred: "I hope that by the end of this year we will have an elected government and that universal freedoms are applied and that we put an end to the corruption that has taken over this country." The protestors in Tahrir Square spoke with one voice about the corruption of the government, its inability to deliver public services, and the lack of equality of opportunity in their country. They particularly complained about repression and the absence of political rights. As Mohamed ElBaradei, former director of the International Atomic Energy Agency, wrote on Twitter on January 13, 2011, "Tunisia: repression + absence of social justice + denial of channels for peaceful change = a ticking bomb." Egyptians and Tunisians both saw their economic problems as being fundamentally caused by their lack of political rights. When the protestors started to formulate their demands more systematically, the first twelve immediate demands posted by Wael Khalil, the software engineer and blogger who emerged as one of the leaders of the Egyptian protest movement, were all focused on political change. Issues such as raising the minimum wage appeared only among the transitional demands that were to be implemented later.

To Egyptians, the things that have held them back include an ineffective and corrupt state and a society where they cannot use their talent, ambition, ingenuity, and what education they can get. But they also recognize that the roots of these problems are political. All the economic impediments they face stem from the way political power in Egypt is exercised and monopolized by a narrow elite. This, they understand, is the first thing that has to change.

Yet, in believing this, the protestors of Tahrir Square have sharply diverged from the conventional wisdom on this topic. When they reason about why a country such as Egypt is poor, most academics and commentators emphasize completely different factors. Some stress that Egypt's poverty is determined primarily by its geography, by the fact that the country is mostly a desert and lacks adequate rainfall, and that its soils and climate do not allow productive agriculture. Others instead point to cultural attributes of Egyptians that are supposedly inimical to economic development and prosperity. Egyptians, they argue, lack the same sort of work ethic and cultural traits that have allowed others to prosper, and instead have accepted Islamic beliefs that are inconsistent with economic success. A third approach, the one dominant among economists and policy pundits, is based on the notion that the rulers of Egypt simply don't know what is needed to make their country prosperous, and have followed incorrect policies and strategies in the past. If these rulers would only get the right advice from the right advisers, the thinking goes, prosperity would follow. To these academics and pundits, the fact that Egypt has been ruled by narrow elites feathering their nests at the expense of society seems irrelevant to understanding the country's economic problems.

In this book we'll argue that the Egyptians in Tahrir Square, not most academics and commentators, have the right idea. In fact, Egypt is poor precisely because it has been ruled by a narrow elite that have organized society for their own benefit at the expense of the vast mass of people. Political power has been narrowly concentrated, and has been used to create great wealth for those who possess it, such as the $70 billion fortune apparently accumulated by ex-president Mubarak. The losers have been the Egyptian people, as they only too well understand.

We'll show that this interpretation of Egyptian poverty, the people's interpretation, turns out to provide a general explanation for why poor countries are poor. Whether it is North Korea, Sierra Leone, or Zimbabwe, we'll show that poor countries are poor for the same reason that Egypt is poor. Countries such as Great Britain and the United States became rich because their citizens overthrew the elites who controlled power and created a society where political rights were much more broadly distributed, where the government was accountable and responsive to citizens, and where the great mass of people could take advantage of economic opportunities. We'll show that to understand why there is such inequality in the world today we have to delve into the past and study the historical dynamics of societies. We'll see that the reason that Britain is richer than Egypt is because in 1688, Britain (or England, to be exact) had a revolution that transformed the politics and thus the economics of the nation. People fought for and won more political rights, and they used them to expand their economic opportunities. The result was a fundamentally different political and economic trajectory, culminating in the Industrial Revolution.

The Industrial Revolution and the technologies it unleashed didn't spread to Egypt, as that country was under the control of the Ottoman Empire, which treated Egypt in rather the same way as the Mubarak family later did. Ottoman rule in Egypt was overthrown by Napoleon Bonaparte in 1798, but the country then fell under the control of British colonialism, which had as little interest as the Ottomans in promoting Egypt's prosperity. Though the Egyptians shook off the Ottoman and British empires and, in 1952, overthrew their monarchy, these were not revolutions like that of 1688 in England, and rather than fundamentally transforming politics in Egypt, they brought to power another elite as disinterested in achieving prosperity for ordinary Egyptians as the Ottoman and British had been. In consequence, the basic structure of society did not change, and Egypt stayed poor.

In this book we'll study how these patterns reproduce themselves over time and why sometimes they are altered, as they were in England in 1688 and in France with the revolution of 1789. This will help us to understand if the situation in Egypt has changed today and whether the revolution that overthrew Mubarak will lead to a new set of institutions capable of bringing prosperity to ordinary Egyptians. Egypt has had revolutions in the past that did not change things, because those who mounted the revolutions simply took over the reins from those they'd deposed and re-created a similar system. It is indeed difficult for ordinary citizens to acquire real political power and change the way their society works. But it is possible, and we'll see how this happened in England, France, and the United States, and also in Japan, Botswana, and Brazil. Fundamentally it is a political transformation of this sort that is required for a poor society to become rich. There is evidence that this may be happening in Egypt. Reda

Metwaly, another protestor in Tahrir Square, argued, “Now you see Muslims and Christians together, now you see old and young together, all wanting the same thing.” We’ll see that such a broad movement in society was a key part of what happened in these other political transformations. If we understand when and why such transitions occur, we will be in a better position to evaluate when we expect such movements to fail as they have often done in the past and when we may hope that they will succeed and improve the lives of millions.

1.

SO CLOSE AND YET SO DIFFERENT

THE ECONOMICS OF THE RIO GRANDE

THE CITY OF NOGALES is cut in half by a fence. If you stand by it and look north, you'll see Nogales, Arizona, located in Santa Cruz County. The income of the average household there is about $30,000 a year. Most teenagers are in school, and the majority of the adults are high school graduates. Despite all the arguments people make about how deficient the U.S. health care system is, the population is relatively healthy, with high life expectancy by global standards. Many of the residents are above age sixty-five and have access to Medicare. It's just one of the many services the government provides that most take for granted, such as electricity, telephones, a sewage system, public health, a road network linking them to other cities in the area and to the rest of the United States, and, last but not least, law and order. The people of Nogales, Arizona, can go about their daily activities without fear for life or safety and not constantly afraid of theft, expropriation, or other things that might jeopardize their investments in their businesses and houses. Equally important, the residents of Nogales, Arizona, take it for granted that, with all its inefficiency and occasional corruption, the government is their agent. They can vote to replace their mayor, congressmen, and senators; they vote in the presidential elections that determine who will lead their country. Democracy is second nature to them.

Life south of the fence, just a few feet away, is rather different. While the residents of Nogales, Sonora, live in a relatively prosperous part of Mexico, the income of the average household there is about one-third that in Nogales, Arizona. Most adults in Nogales, Sonora, do not have a high school degree, and many teenagers are not in school. Mothers have to worry about high rates of infant mortality.

Poor public health conditions mean it's no surprise that the residents of Nogales, Sonora, do not live as long as their northern neighbors. They also don't have access to many public amenities. Roads are in bad condition south of the fence. Law and order is in worse condition. Crime is high, and opening a business is a risky activity. Not only do you risk robbery, but getting all the permissions and greasing all the palms just to open is no easy endeavor. Residents of Nogales, Sonora, live with politicians' corruption and ineptitude every day.

In contrast to their northern neighbors, democracy is a very recent experience for them. Until the political reforms of 2000, Nogales, Sonora, just like the rest of Mexico, was under the corrupt control of the Institutional Revolutionary Party, or Partido Revolucionario Institucional (PRI).

How could the two halves of what is essentially the same city be so different? There is no difference in geography, climate, or the types of diseases prevalent in the area, since germs do not face any restrictions crossing back and forth between the United States and Mexico. Of course, health conditions are very different, but this has nothing to do with the disease environment; it is because the people south of the border live with inferior sanitary conditions and lack decent health care.

But perhaps the residents are very different. Could it be that the residents of Nogales, Arizona, are grandchildren of migrants from Europe, while those in the south are descendants of Aztecs? Not so. The backgrounds of people on both sides of the border are quite similar. After Mexico became independent from Spain in 1821, the area around "Los dos Nogales" was part of the Mexican state of Vieja California and remained so even after the Mexican-American War of 1846–1848. Indeed, it was only after the Gadsden Purchase of 1853 that the U.S. border was extended into this area. It was Lieutenant N. Michler who, while surveying the border, noted the presence of the "pretty little valley of Los Nogales." Here, on either side of the border, the two cities rose up. The inhabitants of Nogales, Arizona, and Nogales, Sonora, share ancestors, enjoy the same food and the same music, and, we would hazard to say, have the same "culture."

Of course, there is a very simple and obvious explanation for the differences between the two halves of Nogales that you've probably long since guessed: the very border that defines the two halves. Nogales, Arizona, is in the United States. Its inhabitants have access to the economic institutions of the United States, which enable them to choose their occupations freely, acquire schooling and skills, and encourage their employers to invest in the best technology, which leads to higher wages for them. They also have access to political institutions that allow them to take part in the democratic process, to elect their representatives, and replace them if they misbehave. In consequence, politicians provide the basic services (ranging from public health to roads to law and order) that the citizens demand. Those of Nogales, Sonora, are not so lucky. They live in a different world shaped by different institutions. These different institutions create very disparate incentives for the inhabitants of the two Nogaleses and for the entrepreneurs and businesses willing to invest there. These incentives created by the different institutions of the Nogaleses and the countries in which they are situated are the main reason for the differences in economic prosperity on the two sides of the border.

Why are the institutions of the United States so much more conducive to economic success than those of Mexico or, for that matter, the rest of Latin America? The answer to this question lies in the way the different societies formed during the early colonial period. An institutional divergence took place then, with implications lasting into the present day. To understand this divergence we must begin right at the foundation of the colonies in North and Latin America.

THE FOUNDING OF BUENOS AIRES

Early in 1516 the Spanish navigator Juan Díaz de Solís sailed into a wide estuary on the Eastern Seaboard of South America. Wading ashore, de Solís claimed the land for Spain, naming the river the Río de la Plata, "River of Silver," since the local people possessed silver. The indigenous peoples on either side of the estuary—the Charrúas in what is now Uruguay, and the Querandí on the plains that were to be known as the Pampas in modern

Argentina—regarded the newcomers with hostility. These locals were hunter-gatherers who lived in small groups without strong centralized political authorities. Indeed it was such a band of Charrúas who clubbed de Solís to death as he explored the new domains he had attemped to occupy for Spain.

In 1534 the Spanish, still optimistic, sent out a first mission of settlers from Spain under the leadership of Pedro de Mendoza. They founded a town on the site of Buenos Aires in the same year. It should have been an ideal place for Europeans. Buenos Aires, literally meaning "good airs," had a hospitable, temperate climate. Yet the first stay of the Spaniards there was short lived. They were not after good airs, but resources to extract and labor to coerce. The Charrúas and the Querandí were not obliging, however. They refused to provide food to the Spaniards, and refused to work when caught. They attacked the new settlement with their bows and arrows. The Spaniards grew hungry, since they had not anticipated having to provide food for themselves. Buenos Aires was not what they had dreamed of. The local people could not be forced into providing labor. The area had no silver or gold to exploit, and the silver that de Solís found had actually come all the way from the Inca state in the Andes, far to the west.

The Spaniards, while trying to survive, started sending out expeditions to find a new place that would offer greater riches and populations easier to coerce. In 1537 one of these expeditions, under the leadership of Juan de Ayolas, penetrated up the Paraná River, searching for a route to the Incas. On its way, it made contact with the Guaraní, a sedentary people with an agricultural economy based on maize and cassava. De Ayolas immediately realized that the Guaraní were a completely different proposition from the Charrúas and the Querandí. After a brief conflict, the Spanish overcame Guaraní resistance and founded a town, Nuestra Señora de Santa María de la Asunción, which remains the capital of Paraguay today. The conquistadors married the Guaraní princesses and quickly set themselves up as a new aristocracy. They adapted the existing systems of forced labor and tribute of the Guaraní, with themselves at the helm. This was the kind of colony they wanted to set up, and within four years Buenos Aires was abandoned as all the Spaniards who'd settled there moved to the new town.

Buenos Aires, the "Paris of South America," a city of wide European-style boulevards based on the great agricultural wealth of the Pampas, was not resettled until 1580. The abandonment of Buenos Aires and the conquest of the Guaraní reveals the logic of European colonization of the Americas. Early Spanish and, as we will see, English colonists were not interested in tilling the soil themselves; they wanted others to do it for them, and they wanted riches, gold and silver, to plunder.

FROM CAJAMARCA…

The expeditions of de Solís, de Mendoza, and de Ayolas came in the wake of more famous ones that followed Christopher Columbus's sighting of one of the islands of the Bahamas on October 12, 1492. Spanish expansion and colonization of the Americas began in earnest with the invasion of Mexico by Hernán Cortés in 1519, the expedition of Francisco Pizarro to Peru a decade and a half later, and the expedition of Pedro de Mendoza to the Río de la Plata just two years after that. Over the next century, Spain conquered and colonized most of central, western, and southern South America, while Portugal claimed Brazil to the east.

The Spanish strategy of colonization was highly effective. First perfected by Cortés in Mexico, it was based on the observation that the best way for the Spanish to subdue opposition was to capture the indigenous leader. This strategy enabled the Spanish to claim the accumulated wealth of the leader and coerce the indigenous peoples to give tribute and food. The next step was setting themselves up as the new elite of the indigenous society and taking control of the existing methods of taxation, tribute, and, particularly, forced labor.

When Cortés and his men arrived at the great Aztec capital of Tenochtitlan on November 8, 1519, they were welcomed by Moctezuma, the Aztec emperor, who had decided, in the face of much advice from his counselors, to welcome the Spaniards peacefully. What happened next is well described by the account compiled after 1545 by the

Franciscan priest Bernardino de Sahagún in his famous Florentine Codices.

> [At] once they [the Spanish] firmly seized Moctezuma … then each of the guns shot off … Fear prevailed. It was as if everyone had swallowed his heart. Even before it had grown dark, there was terror, there was astonishment, there was apprehension, there was a stunning of the people.
>
> And when it dawned thereupon were proclaimed all the things which [the Spaniards] required: white tortillas, roasted turkey hens, eggs, fresh water, wood, firewood, charcoal … This had Moctezuma indeed commanded.
>
> And when the Spaniards were well settled, they thereupon inquired of Moctezuma as to all the city's treasure … with great zeal they sought gold. And Moctezuma thereupon went leading the Spaniards. They went surrounding him … each holding him, each grasping him.
>
> And when they reached the storehouse, a place called Teocalco, thereupon they brought forth all the brilliant things; the quetzal feather head fan, the devices, the shields, the golden discs … the golden nose crescents, the golden leg bands, the golden arm bands, the golden forehead bands.
>
> Thereupon was detached the gold … at once they ignited, set fire to … all the precious things. They all burned. And the gold the Spaniards formed into separate bars … And the Spanish walked everywhere … They took all, all that they saw which they saw to be good.
>
> Thereupon they went to Moctezuma's own storehouse … at the place called Totocalco … they brought forth [Moctezuma's] own property … precious things all; the necklaces with pendants, the

> arm bands with tufts of quetzal feathers, the golden arm bands, the bracelets, the golden bands with shells … and the turquoise diadem, the attribute of the ruler. They took it all.

The military conquest of the Aztecs was completed by 1521. Cortés, as governor of the province of New Spain, then began dividing up the most valuable resource, the indigenous population, through the institution of the *encomienda*. The *encomienda* had first appeared in fifteenth-century Spain as part of the reconquest of the south of the country from the Moors, Arabs who had settled during and after the eighth century. In the New World, it took on a much more pernicious form: it was a grant of indigenous peoples to a Spaniard, known as the *encomendero*. The indigenous peoples had to give the *encomendero* tribute and labor services, in exchange for which the *encomendero* was charged with converting them to Christianity.

A vivid early account of the workings of the *encomienda* has come down to us from Bartolomé de las Casas, a Dominican priest who formulated the earliest and one of the most devastating critiques of the Spanish colonial system. De las Casas arrived on the Spanish island of Hispaniola in 1502 with a fleet of ships led by the new governor, Nicolás de Ovando. He became increasingly disillusioned and disturbed by the cruel and exploitative treatment of the indigenous peoples he witnessed every day. In 1513 he took part as a chaplain in the Spanish conquest of Cuba, even being granted an *encomienda* for his service. However, he renounced the grant and began a long campaign to reform Spanish colonial institutions. His efforts culminated in his book *A Short Account of the Destruction of the Indies*, written in 1542, a withering attack on the barbarity of Spanish rule. On the *encomienda* he has this to say in the case of Nicaragua:

> Each of the settlers took up residence in the town allotted to him (or encommended to him, as the legal phrase has it), put the inhabitants to work for him, stole their already

> scarce foodstuffs for himself and took over the lands owned and worked by the natives and on which they traditionally grew their own produce. The settler would treat the whole of the native population—dignitaries, old men, women and children—as members of his household and, as such, make them labor night and day in his own interests, without any rest whatsoever.

For the conquest of New Granada, modern Colombia, de las Casas reports the whole Spanish strategy in action:

> To realize their long-term purpose of seizing all the available gold, the Spaniards employed their usual strategy of apportioning among themselves (or en-commending, as they have it) the towns and their inhabitants … and then, as ever, treating them as common slaves. The man in overall command of the expedition seized the King of the whole territory for himself and held him prisoner for six or seven months, quite illicitly demanding more and more gold and emeralds from him. This King, one Bogotá, was so terrified that, in his anxiety to free himself from the clutches of his tormentors, he consented to the demand that he fill an entire house with gold and hand it over; to this end he sent his people off in search of gold, and bit by bit they brought it along with many precious stones. But still the house was not filled and the Spaniards eventually declared that they would put him to death for breaking his promise. The commander suggested they should bring the case before him, as a representative of the law, and when they did so, entering formal accusations against the King, he sentenced him to torture should he persist in not honoring the bargain. They tortured him with the strappado, put burning tallow on his belly, pinned both his

> legs to poles with iron hoops and his neck with another and then, with two men holding his hands, proceeded to burn the soles of his feet. From time to time, the commander would look in and repeat that they would torture him to death slowly unless he produced more gold, and this is what they did, the King eventually succumbing to the agonies they inflicted on him.

The strategy and institutions of conquest perfected in Mexico were eagerly adopted elsewhere in the Spanish Empire. Nowhere was this done more effectively than in Pizarro's conquest of Peru. As de las Casas begins his account:

> In 1531 another great villain journeyed with a number of men to the kingdom of Peru. He set out with every intention of imitating the strategy and tactics of his fellow adventurers in other parts of the New World.

Pizarro began on the coast near the Peruvian town of Tumbes and marched south. On November 15, 1532, he reached the mountain town of Cajamarca, where the Inca emperor Atahualpa was encamped with his army. The next day, Atahualpa, who had just vanquished his brother Huáscar in a contest over who would succeed their deceased father, Huayna Capac, came with his retinue to where the Spanish were camped. Atahualpa was irritated because news of atrocities that the Spanish had already committed, such as violating a temple of the Sun God Inti, had reached him. What transpired next is well known. The Spanish laid a trap and sprang it. They killed Atahualpa's guards and retainers, possibly as many as two thousand people, and captured the king. To gain his freedom, Atahualpa had to promise to fill one room with gold and two more of the same size with silver. He did this, but the Spanish, reneging on their promises, strangled him in July 1533. That November, the Spanish captured the Inca capital of Cusco, where the Incan aristocracy received the same treatment as Atahualpa, being imprisoned until they produced gold and silver. When they did not satisfy

Spanish demands, they were burned alive. The great artistic treasures of Cusco, such as the Temple of the Sun, had their gold stripped from them and melted down into ingots.

At this point the Spanish focused on the people of the Inca Empire. As in Mexico, citizens were divided into *encomiendas*, with one going to each of the conquistadors who had accompanied Pizarro. The *encomienda* was the main institution used for the control and organization of labor in the early colonial period, but it soon faced a vigorous contender. In 1545 a local named Diego Gualpa was searching for an indigenous shrine high in the Andes in what is today Bolivia. He was thrown to the ground by a sudden gust of wind and in front of him appeared a cache of silver ore. This was part of a vast mountain of silver, which the Spanish baptized El Cerro Rico, "The Rich Hill." Around it grew the city of Potosí, which at its height in 1650 had a population of 160,000 people, larger than Lisbon or Venice in this period.

To exploit the silver, the Spanish needed miners—a lot of miners. They sent a new viceroy, the chief Spanish colonial official, Francisco de Toledo, whose main mission was to solve the labor problem. De Toledo, arriving in Peru in 1569, first spent five years traveling around and investigating his new charge. He also commissioned a massive survey of the entire adult population. To find the labor he needed, de Toledo first moved almost the entire indigenous population, concentrating them in new towns called *reducciones*—literally "reductions"—which would facilitate the exploitation of labor by the Spanish Crown. Then he revived and adapted an Inca labor institution known as the *mita*, which, in the Incas' language, Quechua, means "a turn." Under their *mita* system, the Incas had used forced labor to run plantations designed to provide food for temples, the aristocracy, and the army. In return, the Inca elite provided famine relief and security. In de Toledo's hands the *mita*, especially the Potosí *mita*, was to become the largest and most onerous scheme of labor exploitation in the Spanish colonial period. De Toledo defined a huge catchment area, running from the middle of modern-day Peru and encompassing most of modern Bolivia. It covered about two hundred thousand square miles. In this area, one-seventh of the male inhabitants, newly arrived in their *reducciones*, were required to work in the mines at Potosí. The Potosí *mita* endured throughout the entire colonial period and was abolished only in 1825. Map 1 shows the catchment area of the *mita* superimposed on the extent of the Inca empire at the time of the Spanish conquest. It illustrates the extent to which the *mita* overlapped with the heartland of the empire, encompassing the capital Cusco.

Map 1: The Inca Empire, the Inca road network, and the mining *mita* catchment area

Remarkably, you still see the legacy of the *mita* in Peru today. Take the differences between the provinces of Calca and nearby Acomayo. There appears to be few differences among these provinces. Both are high in the mountains, and each is inhabited by the Quechua-speaking descendants of the Incas. Yet Acomayo is much poorer, with its inhabitants consuming about one-third less than those in Calca. The people know this. In Acomayo they ask intrepid foreigners, “Don’t you know that the people here are poorer than the people over there in Calca? Why would you ever want to come here?” Intrepid because it is much harder to get to Acomayo from the regional capital of Cusco, ancient center of the Inca Empire, than it is to get to Calca. The road to Calca is surfaced, the one to Acomayo is in a terrible state of disrepair. To get beyond Acomayo, you need a horse or a mule. In Calca and Acomayo, people grow the same crops, but in Calca they sell them on the market for money. In Acomayo they grow food for their own subsistence. These inequalities, apparent to the eye and to the people who live there, can be understood in terms of the institutional differences between these departments—institutional differences with historical roots going back to de Toledo and his plan for effective exploitation of indigenous labor. The major historical difference between Acomayo and Calca is that Acomayo was in the catchment area of the Potosí *mita*. Calca was not.

In addition to the concentration of labor and the *mita*, de Toledo consolidated the *encomienda* into a head tax, a fixed sum payable by each adult male every year in silver. This was another scheme designed to force people into the labor market and reduce wages for Spanish landowners. Another institution, the *repartimiento de mercancias*, also became widespread during de Toledo’s tenure. Derived from the Spanish verb *repartir*, to distribute, this *repartimiento*, literally “the distribution of goods,” involved the forced sale of goods to locals at prices determined by Spaniards. Finally, de Toledo introduced the *trajin*—meaning, literally, “the burden”—which used the indigenous people to carry heavy loads of goods, such as wine or coca leaves or textiles, as a substitute for pack animals, for the business ventures of the Spanish elite.

Throughout the Spanish colonial world in the Americas, similar institutions and social structures emerged. After an initial phase of looting, and gold and silver lust, the Spanish created a web of institutions designed to exploit the indigenous peoples. The full gamut of *encomienda*, *mita*, *repartimiento*, and *trajin* was designed to force indigenous people's living standards down to a subsistence level and thus extract all income in excess of this for Spaniards. This was achieved by expropriating their land, forcing them to work, offering low wages for labor services, imposing high taxes, and charging high prices for goods that were not even voluntarily bought. Though these institutions generated a lot of wealth for the Spanish Crown and made the conquistadors and their descendants very rich, they also turned Latin America into the most unequal continent in the world and sapped much of its economic potential.

… TO JAMESTOWN

As the Spanish began their conquest of the Americas in the 1490s, England was a minor European power recovering from the devastating effects of a civil war, the Wars of the Roses. She was in no state to take advantage of the scramble for loot and gold and the opportunity to exploit the indigenous peoples of the Americas. Nearly one hundred years later, in 1588, the lucky rout of the Spanish Armada, an attempt by King Philip II of Spain to invade England, sent political shockwaves around Europe. Fortunate though England's victory was, it was also a sign of growing English assertiveness on the seas that would enable them to finally take part in the quest for colonial empire.

It is thus no coincidence that the English began their colonization of North America at exactly the same time. But they were already latecomers. They chose North America not because it was attractive, but because it was all that was available. The "desirable" parts of the Americas, where the indigenous population to exploit was plentiful and where the gold and silver mines were located, had already been occupied. The English got the leftovers. When the eighteenth-century English writer and agriculturalist Arthur Young discussed where profitable "staple products," by which he meant exportable agricultural goods, were produced, he noted:

> It appears upon the whole, that the staple

> productions of our colonies decrease in value in proportion to their distance from the sun. In the West Indies, which are the hottest of all, they make to the amount of 8l. 12s. 1d. per head. In the southern continental ones, to the amount of 5l. 10s. In the central ones, to the amount of 9s. 6 1/2d. In the northern settlements, to that of 2s. 6d. This scale surely suggests a most important lesson—to avoid colonizing in northern latitudes.

The first English attempt to plant a colony, at Roanoke, in North Carolina, between 1585 and 1587, was a complete failure. In 1607 they tried again. Shortly before the end of 1606, three vessels, *Susan Constant*, *Godspeed*, and *Discovery*, under the command of Captain Christopher Newport, set off for Virginia. The colonists, under the auspices of the Virginia Company, sailed into Chesapeake Bay and up a river they named the James, after the ruling English monarch, James I. On May 14, 1607, they founded the settlement of Jamestown.

Though the settlers on board the ships owned by the Virginia Company were English, they had a model of colonization heavily influenced by the template set up by Cortés, Pizarro, and de Toledo. Their first plan was to capture the local chief and use him as a way to get provisions and to coerce the population into producing food and wealth for them.

When they first landed in Jamestown, the English colonists did not know that they were within the territory claimed by the Powhatan Confederacy, a coalition of some thirty polities owing allegiance to a king called Wahunsunacock. Wahunsunacock's capital was at the town of Werowocomoco, a mere twenty miles from Jamestown. The plan of the colonists was to learn more about the lay of the land. If the locals could not be induced to provide food and labor, the colonists might at least be able to trade with them. The notion that the settlers themselves would work and grow their own food seems not to have crossed their minds. That is not what conquerors of the New World did.

Wahunsunacock quickly became aware of the colonists' presence and viewed their intentions with great suspicion.

He was in charge of what for North America was quite a large empire. But he had many enemies and lacked the overwhelming centralized political control of the Incas. Wahunsunacock decided to see what the intentions of the English were, initially sending messengers saying that he desired friendly relations with them.

As the winter of 1607 closed in, the settlers in Jamestown began to run low on food, and the appointed leader of the colony's ruling council, Edward Marie Wingfield, dithered indecisively. The situation was rescued by Captain John Smith. Smith, whose writings provide one of our main sources of information about the early development of the colony, was a larger-than-life character. Born in England, in rural Lincolnshire, he disregarded his father's desires for him to go into business and instead became a soldier of fortune. He first fought with English armies in the Netherlands, after which he joined Austrian forces serving in Hungary fighting against the armies of the Ottoman Empire. Captured in Romania, he was sold as a slave and put to work as a field hand. He managed one day to overcome his master and, stealing his clothes and his horse, escape back into Austrian territory. Smith had got himself into trouble on the voyage to Virginia and was imprisoned on the *Susan Constant* for mutiny after defying the orders of Wingfield. When the ships reached the New World, the plan was to put him on trial. To the immense horror of Wingfield, Newport, and other elite colonists, however, when they opened their sealed orders, they discovered that the Virginia Company had nominated Smith to be a member of the ruling council that was to govern Jamestown.

With Newport sailing back to England for supplies and more colonists, and Wingfield uncertain about what to do, it was Smith who saved the colony. He initiated a series of trading missions that secured vital food supplies. On one of these he was captured by Opechancanough, one of Wahunsunacock's younger brothers, and was brought before the king at Werowocomoco. He was the first Englishman to meet Wahunsunacock, and it was at this initial meeting that according to some accounts Smith's life was saved only at the intervention of Wahunsunacock's young daughter Pocahontas. Freed on January 2, 1608, Smith returned to Jamestown, which was still perilously low on food, until the timely return of Newport from England later on the same day.

The colonists of Jamestown learned little from this initial experience. As 1608 proceeded, they continued their quest for gold and precious metals. They still did not seem to understand that to survive, they could not rely on the locals to feed them through either coercion or trade. It was Smith who was the first to realize that the model of colonization that had worked so well for Cortés and Pizarro simply would not work in North America. The underlying circumstances were just too different. Smith noted that, unlike the Aztecs and Incas, the peoples of Virginia did not have gold. Indeed, he noted in his diary, "Victuals you must know is all their wealth." Anas Todkill, one of the early settlers who left an extensive diary, expressed well the frustrations of Smith and the few others on which this recognition dawned:

> "There was no talke, no hope, no worke, but dig gold, refine gold, load gold."

When Newport sailed for England in April 1608 he took a cargo of pyrite, fool's gold. He returned at the end of September with orders from the Virginia Company to take firmer control over the locals. Their plan was to crown Wahunsunacock, hoping this would render him subservient to the English king James I. They invited him to Jamestown, but Wahunsunacock, still deeply suspicious of the colonists, had no intention of risking capture. John Smith recorded Wahunsunacock's reply: "If your King have sent me presents, I also am a King, and this is my land … Your father is to come to me, not I to him, nor yet to your fort, neither will I bite at such a bait."

If Wahunsunacock would not "bite at such a bait," Newport and Smith would have to go to Werowocomoco to undertake the coronation. The whole event appears to have been a complete fiasco, with the only thing coming out of it a resolve on the part of Wahunsunacock that it was time to get rid of the colony. He imposed a trade embargo. Jamestown could no longer trade for supplies. Wahunsunacock would starve them out.

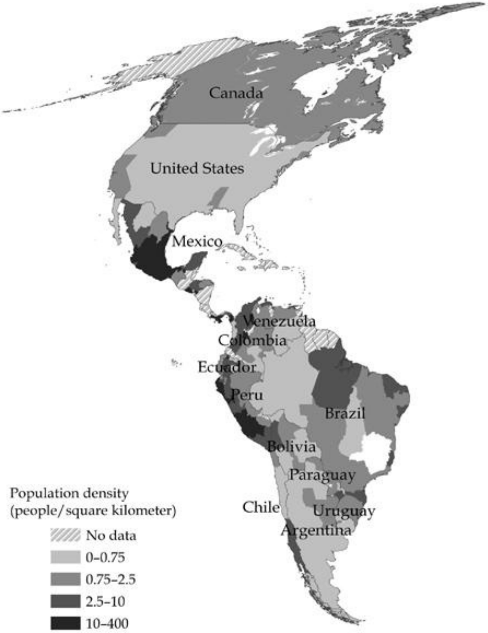
Newport set sail once more for England, in December 1608. He took with him a letter written by Smith pleading with the directors of the Virginia Company to change the way they thought about the colony. There was no possibility of a get-rich-quick exploitation of Virginia along the lines of Mexico and Peru. There were no gold or precious metals, and the indigenous people could not be forced to work or provide food. Smith realized that if there were going to be a viable colony, it was the colonists who would have to work. He therefore pleaded with the directors to send the right sort of people: "When you send againe I entreat you rather to send some thirty carpenters, husbandmen, gardeners, fishermen, blacksmiths, masons, and diggers up of trees, roots, well provided, then a thousand of such as we have."

Smith did not want any more useless goldsmiths. Once more Jamestown survived only because of his resourcefulness. He managed to cajole and bully local indigenous groups to trade with him, and when they wouldn't, he took what he could. Back in the settlement, Smith was completely in charge and imposed the rule that "he that will not worke shall not eat." Jamestown survived a second winter.

The Virginia Company was intended to be a moneymaking enterprise, and after two disastrous years, there was no whiff of profit. The directors of the company decided that they needed a new model of governance, replacing the ruling council with a single governor. The first man appointed to this position was Sir Thomas Gates. Heeding some aspects of Smith's warning, the company realized that they had to try something new. This realization was driven home by the events of the winter of 1609/1610—the so-called "starving time." The new mode of governance left no room for Smith, who, disgruntled, returned to England in the autumn of 1609. Without his resourcefulness, and with Wahunsunacock throttling the food supply, the colonists in Jamestown perished. Of the five hundred who entered the winter, only sixty were alive by March. The situation was so desperate that they resorted to cannibalism.

The "something new" that was imposed on the colony by Gates and his deputy, Sir Thomas Dale, was a work regime of draconian severity for English settlers—though not of course for the elite running the colony. It was Dale who propagated the "Lawes Divine, Morall and Martiall." This included the clauses

> No man or woman shall run away from the colony to the Indians, upon pain of death.
>
> Anyone who robs a garden, public or private, or a vineyard, or who steals ears of corn shall be punished with death.
>
> No member of the colony will sell or give any commodity of this country to a captain, mariner, master or sailor to transport out of the colony, for his own private uses, upon pain of death.



Map 2: Population density in 1500 in the Americas

If the indigenous peoples could not be exploited, reasoned the Virginia Company, perhaps the colonists could. The new model of colonial development entailed the Virginia Company owning all the land. Men were housed in barracks, and given company-determined rations. Work gangs were chosen, each one overseen by an agent of the company. It was close to martial law, with execution as the punishment of first resort. As part of the new institutions for the colony, the first clause just given is significant. The company threatened with death those who ran away. Given the new work regime, running away to live with the locals became more and more of an attractive option for the colonists who had to do the work. Also available, given the low density of even indigenous populations in Virginia at that time, was the prospect of going it alone on the frontier beyond the control of the Virginia Company. The power of the company in the face of these options was limited. It could not coerce the English settlers into hard work at subsistence rations.

Map 2 shows an estimate of the population density of different regions of the Americas at the time on the Spanish conquest. The population density of the United States, outside of a few pockets, was at most three-quarters of a person per square mile. In central Mexico or Andean Peru, the population density was as high as four hundred people per square mile, more than five hundred times higher. What was possible in Mexico or Peru was not feasible in Virginia.

It took the Virginia Company some time to recognize that its initial model of colonization did not work in Virginia, and it took a while, too, for the failure of the "Lawes Divine, Morall and Martiall" to sink in. Starting in 1618, a dramatically new strategy was adopted. Since it was possible to coerce neither the locals nor the settlers, the only alternative was to give the settlers incentives. In 1618 the company began the "headright system," which gave each male settler fifty acres of land and fifty more acres for each member of his family and for all servants that a family could bring to Virginia. Settlers were given their houses and freed from their contracts, and in 1619 a General Assembly was introduced that effectively gave all adult men a say in the laws and institutions governing the colony. It was the start of democracy in the United States.

It took the Virginia Company twelve years to learn its first lesson that what had worked for the Spanish in Mexico and in Central and South America would not work in the north. The rest of the seventeenth century saw a long series of struggles over the second lesson: that the only option for an economically viable colony was to create institutions that gave the colonists incentives to invest and to work hard.

As North America developed, English elites tried time and time again to set up institutions that would heavily restrict the economic and political rights for all but a privileged few of the inhabitants of the colony, just as the Spanish did. Yet in each case this model broke down, as it had in Virginia.

One of the most ambitious attempts began soon after the change in strategy of the Virginia Company. In 1632 ten million acres of land on the upper Chesapeake Bay were granted by the English king Charles I to Cecilius Calvert, Lord Baltimore. The Charter of Maryland gave Lord Baltimore complete freedom to create a government along any lines he wished, with clause VII noting that Baltimore had "for the good and happy Government of the said Province, free, full, and absolute Power, by the Tenor of these Presents, to Ordain, Make, and Enact Laws, of what Kind soever."

Baltimore drew up a detailed plan for creating a manorial society, a North American variant of an idealized version of seventeenth-century rural England. It entailed dividing the land into plots of thousands of acres, which would be run by lords. The lords would recruit tenants, who would work the lands and pay rents to the privileged elite controlling the land. Another similar attempt was made later in 1663, with the founding of Carolina by eight proprietors, including Sir Anthony Ashley-Cooper. Ashley-Cooper, along with his secretary, the great English philosopher John Locke, formulated the Fundamental Constitutions of Carolina. This document, like the Charter of Maryland before it, provided a blueprint for an elitist, hierarchical society based on control by a landed elite. The preamble noted that "the government of this province may be made most agreeable to the monarchy under which we live and of which this province is a part; and that we may avoid erecting a numerous democracy."

The clauses of the Fundamental Constitutions laid out a rigid social structure. At the bottom were the "leet-men," with clause 23 noting, "All the children of leet-men shall be leet-men, and so to all generations." Above the leet-men, who had no political power, were the landgraves and caziques, who were to form the aristocracy. Landgraves were to be allocated forty-eight thousand acres of land each, and caziques twenty-four thousand acres. There was to be a parliament, in which landgraves and caziques were represented, but it would be permitted to debate only those measures that had previously been approved by the eight proprietors.

Just as the attempt to impose draconian rule in Virginia failed, so did the plans for the same type of institutions in Maryland and Carolina. The reasons were similar. In all cases it proved to be impossible to force settlers into a rigid hierarchical society, because there were simply too many options open to them in the New World. Instead, they had to be provided with incentives for them to want to work. And soon they were demanding more economic freedom and further political rights. In Maryland, too, settlers insisted on getting their own land, and they forced Lord Baltimore into creating an assembly. In 1691 the assembly induced the king to declare Maryland a Crown colony, thus removing the political privileges of Baltimore and his great lords. A similar protracted struggle took place in the Carolinas, again with the proprietors losing. South Carolina became a royal colony in 1729.

By the 1720s, all the thirteen colonies of what was to become the United States had similar structures of government. In all cases there was a governor, and an assembly based on a franchise of male property holders. They were not democracies; women, slaves, and the propertyless could not vote. But political rights were very broad compared with contemporary societies elsewhere. It was these assemblies and their leaders that coalesced to form the First Continental Congress in 1774, the prelude to the independence of the United States. The assemblies believed they had the right to determine both their own membership and the right to taxation. This, as we know, created problems for the English colonial government.

A Tale of Two Constitutions

It should now be apparent that it is not a coincidence that the United States, and not Mexico, adopted and enforced a constitution that espoused democratic principles, created limitations on the use of political power, and distributed that power broadly in society. The document that the delegates sat down to write in Philadelphia in May 1787 was the outcome of a long process initiated by the formation of the General Assembly in Jamestown in 1619.

The contrast between the constitutional process that took place at the time of the independence of the United States and the one that took place a little afterward in Mexico is stark. In February 1808, Napoleon Bonaparte's French armies invaded Spain. By May they had taken Madrid, the Spanish capital. By September the Spanish king Ferdinand had been captured and had abdicated. A national junta, the Junta Central, took his place, taking the torch in the fight against the French. The Junta met first at Aranjuez, but retreated south in the face of the French armies. Finally it reached the port of Cádiz, which, though besieged by Napoleonic forces, held out. Here the Junta formed a parliament, called the Cortes. In 1812 the Cortes produced what became known as the Cádiz Constitution, which called for the introduction of a constitutional monarchy based on notions of popular sovereignty. It also called for the end of special privileges and the introduction of equality before the law. These demands were all anathema to the elites of South America, who were still ruling an institutional environment shaped by the *encomienda*, forced labor, and absolute power vested in them and the colonial state.

The collapse of the Spanish state with the Napoleonic invasion created a constitutional crisis throughout colonial Latin America. There was much dispute about whether to recognize the authority of the Junta Central, and in response, many Latin Americans began to form their own juntas. It was only a matter of time before they began to sense the possibility of becoming truly independent from Spain. The first declaration of independence took place in La Paz, Bolivia, in 1809, though it was quickly crushed by Spanish troops sent from Peru. In Mexico the political attitudes of the elite had been shaped by the 1810 Hidalgo Revolt, led by a priest, Father Miguel Hidalgo. When Hidalgo's army sacked Guanajuato on September 23, they killed the intendant, the senior colonial official, and then started indiscriminately to kill white people. It was more like class or even ethnic warfare than an independence movement, and it united all the elites in opposition. If independence allowed popular participation in politics, the local elites, not just Spaniards, were against it. Consequentially, Mexican elites viewed the Cádiz Constitution, which opened the way to popular participation, with extreme skepticism; they would never recognize its legitimacy.

In 1815, as Napoleon's European empire collapsed, King Ferdinand VII returned to power and the Cádiz Constitution was abrogated. As the Spanish Crown began trying to reclaim its American colonies, it did not face a problem with loyalist Mexico. Yet, in 1820, a Spanish army that had assembled in Cádiz to sail to the Americas to help restore Spanish authority mutinied against Ferdinand VII. They were soon joined by army units throughout the country, and Ferdinand was forced to restore the Cádiz Constitution and recall the Cortes. This Cortes was even more radical than the one that had written the Cádiz Constitution, and it proposed abolishing all forms of labor coercion. It also attacked special privileges—for example, the right of the military to be tried for crimes in their own courts. Faced finally with the imposition of this document in Mexico, the elites there decided that it was better to go it alone and declare independence.

This independence movement was led by Augustín de Iturbide, who had been an officer in the Spanish army. On February 24, 1821, he published the Plan de Iguala, his vision for an independent Mexico. The plan featured a constitutional monarchy with a Mexican emperor, and removed the provisions of the Cádiz Constitution that Mexican elites found so threatening to their status and privileges. It received instantaneous support, and Spain quickly realized that it could not stop the inevitable. But Iturbide did not just organize Mexican secession. Recognizing the power vacuum, he quickly took advantage of his military backing to have himself declared emperor, a position that the great leader of South American independence Simón Bolivar described as "by the grace of God and of bayonets." Iturbide was not constrained by the same political institutions that constrained presidents of the United States; he quickly made himself a dictator, and by October 1822 he had dismissed the constitutionally sanctioned congress and replaced it with a junta of his choosing. Though Iturbide did not last long, this pattern of events was to be repeated time and time again in nineteenth-century Mexico.

The Constitution of the United States did not create a democracy by modern standards. Who could vote in elections was left up to the individual states to determine. While northern states quickly conceded the vote to all white men irrespective of how much income they earned or property they owned, southern states did so only gradually. No state enfranchised women or slaves, and as property and wealth restrictions were lifted on white men, racial franchises explicitly disenfranchising black men were introduced. Slavery, of course, was deemed constitutional when the Constitution of the United States was written in Philadelphia, and the most sordid negotiation concerned the division of the seats in the House of Representatives among the states. These were to be allocated on the basis of a state's population, but the congressional representatives of southern states then demanded that the slaves be counted. Northerners objected. The compromise was that in apportioning seats to the House of Representatives, a slave would count as three-fifths of a free person. The conflicts between the North and South of the United States were repressed during the constitutional process as the three-fifths rule and other compromises were worked out. New fixes were added over time—for example, the Missouri Compromise, an arrangement where one proslavery and one antislavery state were always added to the union together, to keep the balance in the Senate between those for and those against slavery. These fudges kept the political institutions of the United States working peacefully until the Civil War finally resolved the conflicts in favor of the North.

The Civil War was bloody and destructive. But both before and after it there were ample economic opportunities for a large fraction of the population, especially in the northern and western United States. The situation in Mexico was very different. If the United States experienced five years of political instability between 1860 and 1865, Mexico experienced almost nonstop instability for the first fifty years of independence. This is best illustrated via the career of Antonio López de Santa Ana.

Santa Ana, son of a colonial official in Veracruz, came to prominence as a soldier fighting for the Spanish in the independence wars. In 1821 he switched sides with Iturbide and never looked back. He became president of Mexico for the first time in May of 1833, though he exercised power for less than a month, preferring to let Valentín Gómez Farías act as president. Gómez Farías's presidency lasted fifteen days, after which Santa Ana retook power. This was as brief as his first spell, however, and he was again replaced by Gómez Farías, in early July. Santa Ana and Gómez Farías continued this dance until the middle of 1835, when Santa Ana was replaced by Miguel Barragán. But Santa Ana was not a quitter. He was back as president in 1839, 1841, 1844, 1847, and, finally, between 1853 and 1855. In all, he was president eleven times, during which he presided over the loss of the Alamo and Texas and the disastrous Mexican-American War, which led to the loss of what became New Mexico and Arizona. Between 1824 and 1867 there were fifty-two presidents in Mexico, few of whom assumed power according to any constitutionally sanctioned procedure.

The consequence of this unprecedented political instability for economic institutions and incentives should be obvious. Such instability led to highly insecure property rights. It also led to a severe weakening of the Mexican state, which now had little authority and little ability to raise taxes or provide public services. Indeed, even though Santa Ana was president in Mexico, large parts of the country were not under his control, which enabled the annexation of Texas by the United States. In addition, as we just saw, the motivation behind the Mexican declaration of independence was to protect the set of economic institutions developed during the colonial period, which had made Mexico, in the words of the great German explorer and geographer of Latin America Alexander von Humbolt, "the country of inequality." These institutions, by basing the society on the exploitation of indigenous people and the creation of monopolies, blocked the economic incentives and initiatives of the great mass of the population. As the United States began to experience the Industrial Revolution in the first half of the nineteenth century, Mexico got poorer.

The Industrial Revolution started in England. Its first success was to revolutionize the production of cotton cloth using new machines powered by water wheels and later by steam engines. Mechanization of cotton production massively increased the productivity of workers in, first, textiles and, subsequently, other industries. The engine of technological breakthroughs throughout the economy was innovation, spearheaded by new entrepreneurs and businessmen eager to apply their new ideas. This initial flowering soon spread across the North Atlantic to the United States. People saw the great economic opportunities available in adopting the new technologies developed in England. They were also inspired to develop their own inventions.

We can try to understand the nature of these inventions by looking at who was granted patents. The patent system, which protects property rights in ideas, was systematized in the Statute of Monopolies legislated by the English Parliament in 1623, partially as an attempt to stop the king from arbitrarily granting "letters patent" to whomever he wanted—effectively granting exclusive rights to undertake certain activities or businesses. The striking thing about the evidence on patenting in the United States is that people who were granted patents came from all sorts of backgrounds and all walks of life, not just the rich and the elite. Many made fortunes based on their patents. Take Thomas Edison, the inventor of the phonogram and the lightbulb and the founder of General Electric, still one of the world's largest companies. Edison was the last of seven children. His father, Samuel Edison, followed many occupations, from splitting shingles for roofs to tailoring to keeping a tavern. Thomas had little formal schooling but was homeschooled by his mother.

Between 1820 and 1845, only 19 percent of patentees in the United States had parents who were professionals or were from recognizable major landowning families. During the same period, 40 percent of those who took out patents had only primary schooling or less, just like Edison. Moreover, they often exploited their patent by starting a firm, again like Edison. Just as the United States in the nineteenth century was more democratic politically than almost any other nation in the world at the time, it was also more democratic than others when it came to innovation. This was critical to its path to becoming the most economically innovative nation in the world.

If you were poor with a good idea, it was one thing to take out a patent, which was not so expensive, after all. It was another thing entirely to use that patent to make money. One way, of course, was to sell the patent to someone else. This is what Edison did early on, to raise some capital, when he sold his Quadruplex telegraph to Western Union for $10,000. But selling patents was a good idea only for someone like Edison, who had ideas faster than he could put them to practice. (He had a world-record 1,093 patents issued to him in the United States and 1,500 worldwide.) The real way to make money from a patent was to start your own business. But to start a business, you need capital, and you need banks to lend the capital to you.

Inventors in the United States were once again fortunate. During the nineteenth century there was a rapid expansion of financial intermediation and banking that was a crucial facilitator of the rapid growth and industrialization that the economy experienced. While in 1818 there were 338 banks in operation in the United States, with total assets of $160 million, by 1914 there were 27,864 banks, with total assets of $27.3 billion. Potential inventors in the United States had ready access to capital to start their businesses. Moreover, the intense competition among banks and financial institutions in the United States meant that this capital was available at fairly low interest rates.

The same was not true in Mexico. In fact, in 1910, the year in which the Mexican Revolution started, there were only forty-two banks in Mexico, and two of these controlled 60 percent of total banking assets. Unlike in the United States, where competition was fierce, there was practically no competition among Mexican banks. This lack of competition meant that the banks were able to charge their customers very high interest rates, and typically confined lending to the privileged and the already wealthy, who would then use their access to credit to increase their grip over the various sectors of the economy.

The form that the Mexican banking industry took in the nineteenth and twentieth centuries was a direct result of the postindependence political institutions of the country. The chaos of the Santa Ana era was followed by an abortive attempt by the French government of Emperor Napoleon II to create a colonial regime in Mexico under Emperor Maximilian between 1864 and 1867. The French were expelled, and a new constitution was written. But the government formed first by Benito Juárez and, after his death, by Sebastián Lerdo de Tejada was soon challenged by a young military man named Porfirio Díaz. Díaz had been a victorious general in the war against the French and had developed aspirations of power. He formed a rebel army and, in November of 1876, defeated the army of the government at the Battle of Tecoac. In May of the next year, he had himself elected president. He went on to rule Mexico in a more or less unbroken and increasingly authoritarian fashion until his overthrow at the outbreak of the revolution thirty-four years later.

Like Iturbide and Santa Ana before him, Díaz started life as a military commander. Such a career path into politics was certainly known in the United States. The first president of the United States, George Washington, was also a successful general in the War of Independence. Ulysses S. Grant, one of the victorious Union generals of the Civil War, became president in 1869, and Dwight D. Eisenhower, the supreme commander of the Allied Forces in Europe during the Second World War, was president of the United States between 1953 and 1961. Unlike Iturbide, Santa Ana, and Díaz, however, none of these military men used force to get into power. Nor did they use force to avoid having to relinquish power. They abided by the Constitution. Though Mexico had constitutions in the nineteenth century, they put few constraints on what Iturbide, Santa Ana, and Díaz could do. These men could be removed from power only the same way they had attained it: by the use of force.

Díaz violated people's property rights, facilitating the expropriation of vast amounts of land, and he granted monopolies and favors to his supporters in all lines of business, including banking. There was nothing new about this behavior. This is exactly what Spanish conquistadors had done, and what Santa Ana did in their footsteps.

The reason that the United States had a banking industry that was radically better for the economic prosperity of the country had nothing to do with differences in the motivation of those who owned the banks. Indeed, the profit motive, which underpinned the monopolistic nature of the banking industry in Mexico, was present in the United States, too. But this profit motive was channeled differently because of the radically different U.S. institutions. The bankers faced different economic institutions, institutions that subjected them to much greater competition. And this was largely because the politicians who wrote the rules for the bankers faced very different incentives themselves, forged by different political institutions. Indeed, in the late eighteenth century, shortly after the Constitution of the United States came into operation, a banking system looking similar to that which subsequently dominated Mexico began to emerge. Politicians tried to set up state banking monopolies, which they could give to their friends and partners in exchange for part of the monopoly profits. The banks also quickly got into the business of lending money to the politicians who regulated them, just as in Mexico. But this situation was not sustainable in the United States, because the politicians who attempted to create these banking monopolies, unlike their Mexican counterparts, were subject to election and reelection. Creating banking monopolies and giving loans to politicians is good business for politicians, if they can get away with it. It is not particularly good for the citizens, however. Unlike in Mexico, in the United States the citizens could keep politicians in check and get rid of ones who would use their offices to enrich themselves or create monopolies for their cronies. In consequence, the banking monopolies crumbled. The broad distribution of political rights in the United States, especially when compared to Mexico, guaranteed equal access to finance and loans. This in turn ensured that those with ideas and inventions could benefit from them.

PATH-DEPENDENT CHANGE

The world was changing in the 1870s and '80s. Latin America was no exception. The institutions that Porfirio Díaz established were not identical to those of Santa Ana or the Spanish colonial state. The world economy boomed in the second half of the nineteenth century, and innovations in transportation such as the steamship and the railway led to a huge expansion of international trade. This wave of globalization meant that resource-rich countries such as Mexico—or, more appropriately, the elites in such countries—could enrich themselves by exporting raw materials and natural resources to industrializing North America or Western Europe. Díaz and his cronies thus found themselves in a different and rapidly evolving world. They realized that Mexico had to change, too. But this didn't mean uprooting the colonial institutions and replacing them with institutions similar to those in the United States. Instead, theirs was "path-dependent" change leading only to the next stage of the institutions that had already made much of Latin America poor and unequal.

Globalization made the large open spaces of the Americas, its "open frontiers," valuable. Often these frontiers were only mythically open, since they were inhabited by indigenous peoples who were brutally dispossessed. All the same, the scramble for this newly valuable resource was one of the defining processes of the Americas in the second half of the nineteenth century. The sudden opening of this valuable frontier led not to parallel processes in the United States and Latin America, but to a further divergence, shaped by the existing institutional differences, especially those concerning who had access to the land. In the United States a long series of legislative acts, ranging from the Land Ordinance of 1785 to the Homestead Act of 1862, gave broad access to frontier lands. Though indigenous peoples had been sidelined, this created an egalitarian and economically dynamic frontier. In most Latin American countries, however, the political institutions there created a very different outcome. Frontier lands were allocated to the politically powerful and those with wealth and contacts, making such people even more powerful.

Díaz also started to dismantle many of the specific colonial institutional legacies preventing international trade, which he anticipated could greatly enrich him and his supporters. His model, however, continued to be not the type of economic development he saw north of the Rio Grande but that of Cortés, Pizarro, and de Toledo, where the elite would make huge fortunes while the rest were excluded. When the elite invested, the economy would grow a little, but such economic growth was always going to be disappointing. It also came at the expense of those lacking rights in this new order, such as the Yaqui people of Sonora, in the hinterland of Nogales. Between 1900 and 1910, possibly thirty thousand Yaqui were deported, essentially enslaved, and sent to work in the henequen plantations of Yucatán. (The fibers of the henequen plant were a valuable export, since they could be used to make rope and twine.)

The persistence into the twentieth century of a specific institutional pattern inimical to growth in Mexico and Latin America is well illustrated by the fact that, just as in the nineteenth century, the pattern generated economic stagnation and political instability, civil wars and coups, as groups struggled for the benefits of power. Díaz finally lost power to revolutionary forces in 1910. The Mexican Revolution was followed by others in Bolivia in 1952, Cuba in 1959, and Nicaragua in 1979. Meanwhile, sustained civil wars raged in Colombia, El Salvador, Guatemala, and Peru. Expropriation or the threat of expropriation of assets continued apace, with mass agrarian reforms (or attempted reforms) in Bolivia, Brazil, Chile, Colombia, Guatemala, Peru, and Venezuela. Revolutions, expropriations, and political instability came along with military governments and various types of dictatorships. Though there was also a gradual drift toward greater political rights, it was only in the 1990s that most Latin American countries became democracies, and even then they remain mired in instability.

This instability was accompanied by mass repression and murder. The 1991 National Commission for Truth and Reconciliation Report in Chile determined that 2,279 persons were killed for political reasons during the Pinochet dictatorship between 1973 and 1990. Possibly 50,000 were imprisoned and tortured, and hundreds of thousands of people were fired from their jobs. The Guatemalan Commission for Historical Clarification Report in 1999 identified a total of 42,275 named victims, though others have claimed that as many as 200,000 were murdered in Guatemala between 1962 and 1996, 70,000 during the regime of General Efrain Ríos Montt, who was able to commit these crimes with such impunity that he could run for president in 2003; fortunately he did not win. The National Commission on the Disappearance of Persons in Argentina put the number of people murdered by the military there at 9,000 persons from 1976 to 1983, although it noted that the actual number could be higher. (Estimates by human rights organizations usually place it at 30,000.)

MAKING A BILLION OR TWO

The enduring implications of the organization of colonial society and those societies' institutional legacies shape the modern differences between the United States and Mexico, and thus the two parts of Nogales. The contrast between how Bill Gates and Carlos Slim became the two richest men in the world—Warren Buffett is also a contender—illustrates the forces at work. The rise of Gates and Microsoft is well known, but Gates's status as the world's richest person and the founder of one of the most technologically innovative companies did not stop the U.S. Department of Justice from filing civil actions against the Microsoft Corporation on May 8, 1998, claiming that Microsoft had abused monopoly power. Particularly at issue was the way that Microsoft had tied its Web browser, Internet Explorer, to its Windows operating system. The government had been keeping an eye on Gates for quite some time, and as early as 1991, the Federal Trade Commission had launched an inquiry into whether Microsoft was abusing its monopoly on PC operating systems. In November 2001, Microsoft reached a deal with the Justice Department. It had its wings clipped, even if the penalties were less than many demanded.

In Mexico, Carlos Slim did not make his money by innovation. Initially he excelled in stock market deals, and in buying and revamping unprofitable firms. His major coup was the acquisition of Telmex, the Mexican telecommunications monopoly that was privatized by President Carlos Salinas in 1990. The government announced its intention to sell 51 percent of the voting stock (20.4 percent of total stock) in the company in September

1989 and received bids in November 1990. Even though Slim did not put in the highest bid, a consortium led by his Grupo Corso won the auction. Instead of paying for the shares right away, Slim managed to delay payment, using the dividends of Telmex itself to pay for the stock. What was once a public monopoly now became Slim's monopoly, and it was hugely profitable.

The economic institutions that made Carlos Slim who he is are very different from those in the United States. If you're a Mexican entrepreneur, entry barriers will play a crucial role at every stage of your career. These barriers include expensive licenses you have to obtain, red tape you have to cut through, politicians and incumbents who will stand in your way, and the difficulty of getting funding from a financial sector often in cahoots with the incumbents you're trying to compete against. These barriers can be either insurmountable, keeping you out of lucrative areas, or your greatest friend, keeping your competitors at bay. The difference between the two scenarios is of course whom you know and whom you can influence—and yes, whom you can bribe. Carlos Slim, a talented, ambitious man from a relatively modest background of Lebanese immigrants, has been a master at obtaining exclusive contracts; he managed to monopolize the lucrative telecommunications market in Mexico, and then to extend his reach to the rest of Latin America.

There have been challenges to Slim's Telmex monopoly. But they have not been successful. In 1996 Avantel, a long-distance phone provider, petitioned the Mexican Competition Commission to check whether Telmex had a dominant position in the telecommunications market. In 1997 the commission declared that Telmex had substantial monopoly power with respect to local telephony, national long-distance calls, and international long-distance calls, among other things. But attempts by the regulatory authorities in Mexico to limit these monopolies have come to nothing. One reason is that Slim and Telmex can use what is known as a *recurso de amparo*, literally an "appeal for protection." An *amparo* is in effect a petition to argue that a particular law does not apply to you. The idea of the *amparo* dates back to the Mexican constitution of 1857 and was originally intended as a safeguard of individual rights and freedoms. In the hands of Telmex and other Mexican monopolies, however, it has become a formidable tool for cementing monopoly power. Rather than protecting people's rights, the *amparo* provides a loophole in equality before the law.

Slim has made his money in the Mexican economy in large part thanks to his political connections. When he has ventured into the United States, he has not been successful. In 1999 his Grupo Curso bought the computer retailer CompUSA. At the time, CompUSA had given a franchise to a firm called COC Services to sell its merchandise in Mexico. Slim immediately violated this contract with the intention of setting up his own chain of stores, without any competition from COC. But COC sued CompUSA in a Dallas court. There are no *amparos* in Dallas, so Slim lost, and was fined $454 million. The lawyer for COC, Mark Werner, noted afterward that "the message of this verdict is that in this global economy, firms have to respect the rules of the United States if they want to come here." When Slim was subject to the institutions of the United States, his usual tactics for making money didn't work.

TOWARD A THEORY OF WORLD INEQUALITY

We live in an unequal world. The differences among nations are similar to those between the two parts of Nogales, just on a larger scale. In rich countries, individuals are healthier, live longer, and are much better educated. They also have access to a range of amenities and options in life, from vacations to career paths, that people in poor countries can only dream of. People in rich countries also drive on roads without potholes, and enjoy toilets, electricity, and running water in their houses. They also typically have governments that do not arbitrarily arrest or harass them; on the contrary, the governments provide services, including education, health care, roads, and law and order. Notable, too, is the fact that the citizens vote in elections and have some voice in the political direction their countries take.

The great differences in world inequality are evident to everyone, even to those in poor countries, though many lack access to television or the Internet. It is the perception and reality of these differences that drive people to cross the Rio Grande or the Mediterranean Sea illegally to have the chance to experience rich-country living standards and opportunities. This inequality doesn't just have consequences for the lives of individual people in poor countries; it also causes grievances and resentment, with huge political consequences in the United States and elsewhere. Understanding why these differences exist and what causes them is our focus in this book. Developing such an understanding is not just an end in itself, but also a first step toward generating better ideas about how to improve the lives of billions who still live in poverty.

The disparities on the two sides of the fence in Nogales are just the tip of the iceberg. As in the rest of northern Mexico, which benefits from trade with the United States, even if not all of it is legal, the residents of Nogales are more prosperous than other Mexicans, whose average annual household income is around $5,000. This greater relative prosperity of Nogales, Sonora, comes from maquiladora manufacturing plants centered in industrial parks, the first of which was started by Richard Campbell, Jr., a California basket manufacturer. The first tenant was Coin-Art, a musical instrument company owned by Richard Bosse, owner of the Artley flute and saxophone company in Nogales, Arizona. Coin-Art was followed by Memorex (computer wiring); Avent (hospital clothing); Grant (sunglasses); Chamberlain (a manufacturer of garage door openers for Sears); and Samsonite (suitcases). Significantly, all are U.S.-based businesses and businessmen, using U.S. capital and know-how. The greater prosperity of Nogales, Sonora, relative to the rest of Mexico, therefore, comes from outside.

The differences between the United States and Mexico are in turn small compared with those across the entire globe. The average citizen of the United States is seven times as prosperous as the average Mexican and more than ten times as the resident of Peru or Central America. She is about twenty times as prosperous as the average inhabitant of sub-Saharan Africa, and almost forty times as those living in the poorest African countries such as Mali, Ethiopia, and Sierra Leone. And it's not just the United

States. There is a small but growing group of rich countries —mostly in Europe and North America, joined by Australia, Japan, New Zealand, Singapore, South Korea, and Taiwan —whose citizens enjoy very different lives from those of the inhabitants of the rest of the globe.

The reason that Nogales, Arizona, is much richer than Nogales, Sonora, is simple; it is because of the very different institutions on the two sides of the border, which create very different incentives for the inhabitants of Nogales, Arizona, versus Nogales, Sonora. The United States is also far richer today than either Mexico or Peru because of the way its institutions, both economic and political, shape the incentives of businesses, individuals, and politicians. Each society functions with a set of economic and political rules created and enforced by the state and the citizens collectively. Economic institutions shape economic incentives: the incentives to become educated, to save and invest, to innovate and adopt new technologies, and so on. It is the political process that determines what economic institutions people live under, and it is the political institutions that determine how this process works. For example, it is the political institutions of a nation that determine the ability of citizens to control politicians and influence how they behave. This in turn determines whether politicians are agents of the citizens, albeit imperfect, or are able to abuse the power entrusted to them, or that they have usurped, to amass their own fortunes and to pursue their own agendas, ones detrimental to those of the citizens. Political institutions include but are not limited to written constitutions and to whether the society is a democracy. They include the power and capacity of the state to regulate and govern society. It is also necessary to consider more broadly the factors that determine how political power is distributed in society, particularly the ability of different groups to act collectively to pursue their objectives or to stop other people from pursuing theirs.

As institutions influence behavior and incentives in real life, they forge the success or failure of nations. Individual talent matters at every level of society, but even that needs an institutional framework to transform it into a positive force. Bill Gates, like other legendary figures in the information technology industry (such as Paul Allen, Steve Ballmer, Steve Jobs, Larry Page, Sergey Brin, and Jeff Bezos), had immense talent and ambition. But he ultimately responded to incentives. The schooling system in the United States enabled Gates and others like him to acquire a unique set of skills to complement their talents. The economic institutions in the United States enabled these men to start companies with ease, without facing insurmountable barriers. Those institutions also made the financing of their projects feasible. The U.S. labor markets enabled them to hire qualified personnel, and the relatively competitive market environment enabled them to expand their companies and market their products. These entrepreneurs were confident from the beginning that their dream projects could be implemented: they trusted the institutions and the rule of law that these generated and they did not worry about the security of their property rights. Finally, the political institutions ensured stability and continuity. For one thing, they made sure that there was no risk of a dictator taking power and changing the rules of the game, expropriating their wealth, imprisoning them, or threatening their lives and livelihoods. They also made sure that no particular interest in society could warp the government in an economically disastrous direction, because political power was both limited and distributed sufficiently broadly that a set of economic institutions that created the incentives for prosperity could emerge.

This book will show that while economic institutions are critical for determining whether a country is poor or prosperous, it is politics and political institutions that determine what economic institutions a country has. Ultimately the good economic institutions of the United States resulted from the political institutions that gradually emerged after 1619. Our theory for world inequality shows how political and economic institutions interact in causing poverty or prosperity, and how different parts of the world ended up with such different sets of institutions. Our brief review of the history of the Americas begins to give a sense of the forces that shape political and economic institutions. Different patterns of institutions today are deeply rooted in the past because once society gets organized in a particular way, this tends to persist. We'll show that this fact comes from the way that political and economic institutions interact.

This persistence and the forces that create it also explain why it is so difficult to remove world inequality and to make poor countries prosperous. Though institutions are the key to the differences between the two Nogaleses and between Mexico and the United States, that doesn't mean there will be a consensus in Mexico to change institutions. There is no necessity for a society to develop or adopt the institutions that are best for economic growth or the welfare of its citizens, because other institutions may be even better for those who control politics and political institutions. The powerful and the rest of society will often disagree about which set of institutions should remain in place and which ones should be changed. Carlos Slim would not have been happy to see his political connections disappear and the entry barriers protecting his businesses fizzle—no matter that the entry of new businesses would enrich millions of Mexicans. Because there is no such consensus, what rules society ends up with is determined by politics: who has power and how this power can be exercised. Carlos Slim has the power to get what he wants. Bill Gates's power is far more limited. That's why our theory is about not just economics but also politics. It is about the effects of institutions on the success and failure of nations—thus the economics of poverty and prosperity; it is also about how institutions are determined and change over time, and how they fail to change even when they create poverty and misery for millions—thus the politics of poverty and prosperity.

2.

THEORIES THAT DON'T WORK

THE LAY OF THE LAND

THE FOCUS OF our book is on explaining world inequality and also some of the easily visible broad patterns that nest within it. The first country to experience sustained economic growth was England—or Great Britain, usually just Britain, as the union of England, Wales, and Scotland after 1707 is known. Growth emerged slowly in the second half of the eighteenth century as the Industrial Revolution, based on major technological breakthroughs and their application in industry, took root. Industrialization in England was soon followed by industrialization in most of Western Europe and the United States. English prosperity also spread rapidly to Britain's "settler colonies" of Canada, Australia, and New Zealand. A list of the thirty richest countries today would include them, plus Japan, Singapore, and South Korea. The prosperity of these latter three is in turn part of a broader pattern in which many East Asian nations, including Taiwan and subsequently China, have experienced recent rapid growth.

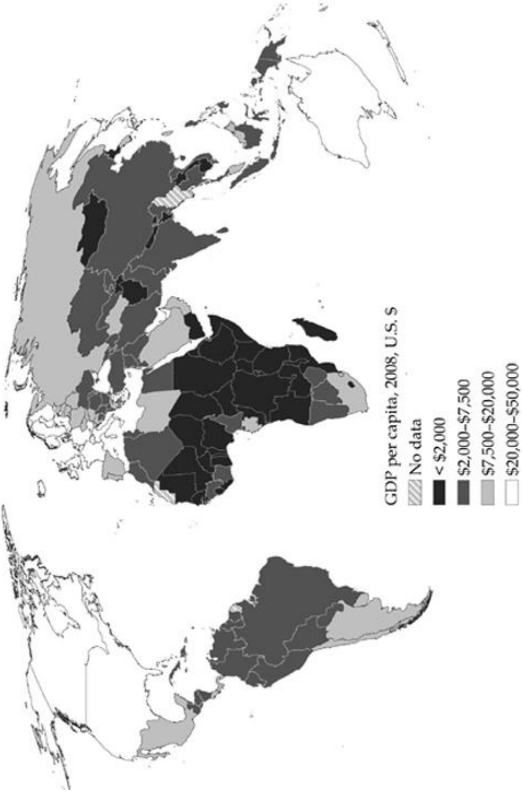
The bottom of the world income distribution paints as sharp and as distinctive a picture as the top. If you instead make a list of the poorest thirty countries in the world today, you will find almost all of them in sub-Saharan Africa. They are joined by countries such as Afghanistan, Haiti, and Nepal, which, though not in Africa, all share something critical with African nations, as we'll explain. If you went back fifty years, the countries in the top and bottom thirty wouldn't be greatly different. Singapore and South Korea would not be among the richest countries, and there would be several different countries in the bottom thirty, but the overall picture that emerged would be remarkably consistent with what we see today. Go back one hundred years, or a hundred and fifty, and you'd find nearly the same countries in the same groups.

Map 3 shows the lay of the land in 2008. The countries shaded in the darkest color are the poorest in the world, those where average per-capita incomes (called by economists GDP, gross domestic product) are less than \$2,000 annually. Most of Africa is in this color, as are Afghanistan, Haiti, and parts of Southeast Asia (for example, Cambodia and Laos). North Korea is also among this group of countries. The countries in white are the richest, those with annual income per-capita of \$20,000 or more. Here we find the usual suspects: North America, western Europe, Australasia, and Japan.

Another interesting pattern can be discerned in the Americas. Make a list of the nations in the Americas from richest to poorest. You will find that at the top are the United States and Canada, followed by Chile, Argentina, Brazil, Mexico, and Uruguay, and maybe also Venezuela, depending on the price of oil. After that you have Colombia, the Dominican Republic, Ecuador, and Peru. At the bottom there is another distinct, much poorer group, comprising Bolivia, Guatemala, and Paraguay. Go back fifty years, and you'll find an identical ranking. One hundred years: same thing. One hundred and fifty years: again the same. So it is not just that the United States and Canada are richer than Latin America; there is also a definite and persistent divide between the rich and poor nations within Latin America.

A final interesting pattern is in the Middle East. There we find oil-rich nations such as Saudi Arabia and Kuwait, which have income levels close to those of our top thirty. Yet if the oil price fell, they would quickly fall back down the table. Middle Eastern countries with little or no oil, such as Egypt, Jordan, and Syria, all cluster around a level of income similar to that of Guatemala or Peru. Without oil, Middle Eastern countries are also all poor, though, like those in Central America and the Andes, not so poor as those in sub-Saharan Africa.

While there is a lot of persistence in the patterns of prosperity we see around us today, these patterns are not unchanging or immutable. First, as we have already emphasized, most of current world inequality emerged since the late eighteenth century, following on the tails of the Industrial Revolution. Not only were gaps in prosperity much smaller as late as the middle of the eighteenth century, but the rankings which have been so stable since then are not the same when we go further back in history. In the Americas, for example, the ranking we see for the last hundred and fifty years was completely different five hundred years ago. Second, many nations have experienced several decades of rapid growth, such as much of East Asia since the Second World War and, more recently, China. Many of these subsequently saw that growth go into reverse. Argentina, for example, grew rapidly for five decades up until 1920, becoming one of the richest countries in the world, but then started a long slide. The Soviet Union is an even more noteworthy example, growing rapidly between 1930 and 1970, but subsequently experiencing a rapid collapse.



Map 3: Prosperity around the world in 2008

What explains these major differences in poverty and prosperity and the patterns of growth? Why did Western European nations and their colonial offshoots filled with

European settlers start growing in the nineteenth century, scarcely looking back? What explains the persistent ranking of inequality within the Americas? Why have sub-Saharan African and Middle Eastern nations failed to achieve the type of economic growth seen in Western Europe, while much of East Asia has experienced breakneck rates of economic growth?

One might think that the fact that world inequality is so huge and consequential and has such sharply drawn patterns would mean that it would have a well-accepted explanation. Not so. Most hypotheses that social scientists have proposed for the origins of poverty and prosperity just don't work and fail to convincingly explain the lay of the land.

THE GEOGRAPHY HYPOTHESIS

One widely accepted theory of the causes of world inequality is the geography hypothesis, which claims that the great divide between rich and poor countries is created by geographical differences. Many poor countries, such as those of Africa, Central America, and South Asia, are between the tropics of Cancer and Capricorn. Rich nations, in contrast, tend to be in temperate latitudes. This geographic concentration of poverty and prosperity gives a superficial appeal to the geography hypothesis, which is the starting point of the theories and views of many social scientists and pundits alike. But this doesn't make it any less wrong.

As early as the late eighteenth century, the great French political philosopher Montesquieu noted the geographic concentration of prosperity and poverty, and proposed an explanation for it. He argued that people in tropical climates tended to be lazy and to lack inquisitiveness. As a consequence, they didn't work hard and were not innovative, and this was the reason why they were poor. Montesquieu also speculated that lazy people tended to be ruled by despots, suggesting that a tropical location could explain not just poverty but also some of the political phenomena associated with economic failure, such as dictatorship.

The theory that hot countries are intrinsically poor, though contradicted by the recent rapid economic advance of countries such as Singapore, Malaysia, and Botswana, is still forcefully advocated by some, such as the economist Jeffrey Sachs. The modern version of this view emphasizes not the direct effects of climate on work effort or thought processes, but two additional arguments: first, that tropical diseases, particularly malaria, have very adverse consequences for health and therefore labor productivity; and second, that tropical soils do not allow for productive agriculture. The conclusion, though, is the same: temperate climates have a relative advantage over tropical and semitropical areas.

World inequality, however, cannot be explained by climate or diseases, or any version of the geography hypothesis. Just think of Nogales. What separates the two parts is not climate, geography, or disease environment, but the U.S.-Mexico border.

If the geography hypothesis cannot explain differences between the north and south of Nogales, or North and South Korea, or those between East and West Germany before the fall of the Berlin Wall, could it still be a useful theory for explaining differences between North and South America? Between Europe and Africa? Simply, no.

History illustrates that there is no simple or enduring connection between climate or geography and economic success. For instance, it is not true that the tropics have always been poorer than temperate latitudes. As we saw in the last chapter, at the time of the conquest of the Americas by Columbus, the areas south of the Tropic of Cancer and north of the Tropic of Capricorn, which today include Mexico, Central America, Peru, and Bolivia, held the great Aztec and Inca civilizations. These empires were politically centralized and complex, built roads, and provided famine relief. The Aztecs had both money and writing, and the Incas, even though they lacked both these two key technologies, recorded vast amounts of information on knotted ropes called quipus. In sharp contrast, at the time of the Aztecs and Incas, the north and south of the area inhabited by the Aztecs and Incas, which today includes the United States, Canada, Argentina, and Chile, were mostly inhabited by Stone Age civilizations lacking these technologies. The tropics in the Americas were thus much richer than the temperate zones, suggesting that the "obvious fact" of tropical poverty is neither obvious nor a fact. Instead, the greater riches in the United States and Canada represent a stark reversal of fortune relative to what was there when the Europeans arrived.

This reversal clearly had nothing to do with geography and, as we have already seen, something to do with the way these areas were colonized. This reversal was not confined to the Americas. People in South Asia, especially the Indian subcontinent, and in China were more prosperous than those in many other parts of Asia and certainly more than the peoples inhabiting Australia and New Zealand. This, too, was reversed, with South Korea, Singapore, and Japan emerging as the richest nations in Asia, and Australia and New Zealand surpassing almost all of Asia in terms of prosperity. Even within sub-Saharan Africa there was a similar reversal. More recently, before the start of intense European contact with Africa, the southern Africa region was the most sparsely settled and the farthest from having developed states with any kind of control over their territories. Yet South Africa is now one of the most prosperous nations in sub-Saharan Africa. Further back in history we again see much prosperity in the tropics; some of the great premodern civilizations, such as Angkor in modern Cambodia, Vijayanagara in southern India, and Aksum in Ethiopia, flourished in the tropics, as did the great Indus Valley civilizations of Mohenjo Daro and Harappa in modern Pakistan. History thus leaves little doubt that there is no simple connection between a tropical location and economic success.

Tropical diseases obviously cause much suffering and high rates of infant mortality in Africa, but they are not the reason Africa is poor. Disease is largely a consequence of poverty and of governments being unable or unwilling to undertake the public health measures necessary to eradicate them. England in the nineteenth century was also a very unhealthy place, but the government gradually invested in clean water, in the proper treatment of sewage and effluent, and, eventually, in an effective health service. Improved health and life expectancy were not the cause of England's economic success but one of the fruits of its previous political and economic changes. The same is true for Nogales, Arizona.

The other part of the geography hypothesis is that the tropics are poor because tropical agriculture is intrinsically unproductive. Tropical soils are thin and unable to maintain nutrients, the argument goes, and emphasizes how quickly these soils are eroded by torrential rains. There certainly is some merit in this argument, but as we'll show, the prime determinant of why agricultural productivity—agricultural output per acre—is so low in many poor countries, particularly in sub-Saharan Africa, has little to do with soil quality. Rather, it is a consequence of the ownership structure of the land and the incentives that are created for farmers by the governments and institutions under which they live. We will also show that world inequality cannot be explained by differences in agricultural productivity. The great inequality of the modern world that emerged in the nineteenth century was caused by the uneven dissemination of industrial technologies and manufacturing production. It was not caused by divergence in agricultural performance.

Another influential version of the geography hypothesis is advanced by the ecologist and evolutionary biologist Jared Diamond. He argues that the origins of intercontinental inequality at the start of the modern period, five hundred years ago, rested in different historical endowments of plant and animal species, which subsequently influenced agricultural productivity. In some places, such as the Fertile Crescent in the modern Middle East, there were a large number of species that could be domesticated by humans. Elsewhere, such as the Americas, there were not. Having many species capable of being domesticated made it very attractive for societies to make the transition from a hunter-gatherer to a farming lifestyle. As a consequence, farming developed earlier in the Fertile Crescent than in the Americas. Population density grew, allowing specialization of labor, trade, urbanization, and political development. Crucially, in places where farming dominated, technological innovation took place much more rapidly than in other parts of the world. Thus, according to Diamond, the differential availability of animal and plant species created differential intensities of farming, which led to different paths of technological change and prosperity across different continents.

Though Diamond's thesis is a powerful approach to the puzzle on which he focuses, it cannot be extended to explain modern world inequality. For example, Diamond argues that the Spanish were able to dominate the civilizations of the Americas because of their longer history of farming and consequent superior technology. But we now need to explain why the Mexicans and Peruvians inhabiting the former lands of the Aztecs and Incas are poor. While having access to wheat, barley, and horses might have made the Spanish richer than the Incas, the gap in incomes between the two was not very large. The average income of a Spaniard was probably less than double that of a citizen of the Inca Empire. Diamond's thesis implies that once the Incas had been exposed to all the species and resulting technologies that they had not been able to develop themselves, they ought quickly to have attained the living standards of the Spanish. Yet nothing of the sort happened. On the contrary, in the nineteenth and twentieth centuries, a much larger gap in incomes between Spain and Peru emerged. Today the average Spaniard is more than six times richer than the average Peruvian. This gap in incomes is closely connected to the uneven dissemination of modern industrial technologies, but this has little to do either with the potential for animal and plant domestication or with intrinsic agricultural productivity differences between Spain and Peru.

While Spain, albeit with a lag, adopted the technologies of steam power, railroads, electricity, mechanization, and factory production, Peru did not, or at best did so very slowly and imperfectly. This technological gap persists today and reproduces itself on a bigger scale as new technologies, in particular those related to information technology, fuel further growth in many developed and some rapidly developing nations. Diamond's thesis does not tell us why these crucial technologies are not diffusing and equalizing incomes across the world and does not explain why the northern half of Nogales is so much richer than its twin just to the south of the fence, even though both were part of the same civilization five hundred years ago.

The story of Nogales highlights another major problem in adapting Diamond's thesis: as we have already seen, whatever the drawbacks of the Inca and Aztec empires were in 1532, Peru and Mexico were undoubtedly more prosperous than those parts of the Americas that went on to become the United States and Canada. North America became more prosperous precisely because it enthusiastically adopted the technologies and advances of the Industrial Revolution. The population became educated and railways spread out across the Great Plains in stark contrast to what happened in South America. This cannot be explained by pointing to differential geographic endowments of North and South America, which, if anything, favored South America.

Inequality in the modern world largely results from the uneven dissemination and adoption of technologies, and Diamond's thesis does include important arguments about this. For instance, he argues, following the historian William McNeill, that the east–west orientation of Eurasia enabled crops, animals, and innovations to spread from the Fertile Crescent into Western Europe, while the north–south orientation of the Americas accounts for why writing systems, which were created in Mexico, did not spread to the Andes or North America. Yet the orientation of continents cannot provide an explanation for today's world inequality. Consider Africa. Though the Sahara Desert did present a significant barrier to the movement of goods and ideas from the north to sub-Saharan Africa, this was not insurmountable. The Portuguese, and then other Europeans, sailed around the coast and eliminated differences in knowledge at a time when gaps in incomes were very small compared with what they are today. Since then, Africa has not caught up with Europe; on the contrary, there is now a much larger income gap between most African and European countries.

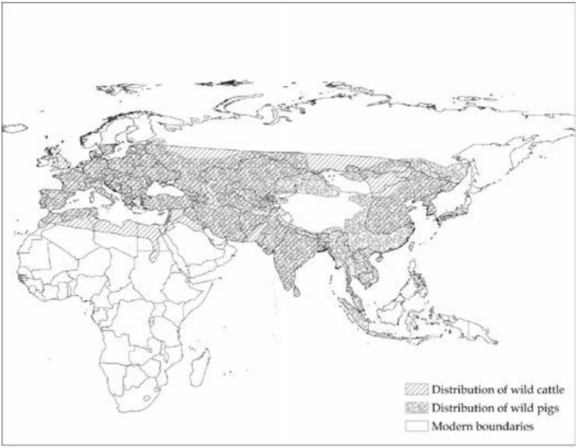
It should also be clear that Diamond's argument, which is about continental inequality, is not well equipped to explain variation within continents—an essential part of modern world inequality. For example, while the orientation of the Eurasian landmass might explain how England managed to benefit from the innovations of the Middle East without having to reinvent them, it doesn't explain why the Industrial Revolution happened in England rather than, say, Moldova.

In addition, as Diamond himself points out, China and India benefited greatly from very rich suites of animals and plants, and from the orientation of Eurasia. But most of the poor people of the world today are in those two countries.

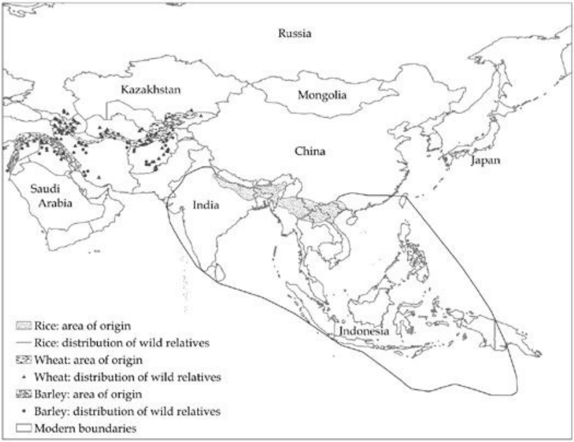
In fact, the best way to see the scope of Diamond's thesis is in terms of his own explanatory variables. Map 4 shows data on the distribution of *Sus scrofa*, the ancestor of the modern pig, and the aurochs, ancestor of the modern cow. Both species were widely distributed throughout Eurasia and even North Africa. Map 5 (this page) shows the distribution of some of the wild ancestors of modern domesticated crops, such as *Oryza sativa*, the ancestor of Asian cultivated rice, and the ancestors of modern wheat and barley. It demonstrates that the wild ancestor of rice was distributed widely across south and southeast Asia, while the ancestors of barley and wheat were distributed along a long arc from the Levant, reaching through Iran and into Afghanistan and the cluster of "stans" (Turkmenistan, Tajikistan, and Krgyzistan). These ancestral species are present in much of Eurasia. But their wide distribution suggests that inequality within Eurasia cannot be explained by a theory based on the incidence of the species.

The geography hypothesis is not only unhelpful for explaining the origins of prosperity throughout history, and mostly incorrect in its emphasis, but also unable to account for the lay of the land we started this chapter with. One might argue that any persistent pattern, such as the hierarchy of incomes within the Americas or the sharp and long-ranging differences between Europe and the Middle East, can be explained by unchanging geography. But this is not so. We have already seen that the patterns within the Americas are highly unlikely to have been driven by geographical factors. Before 1492 it was the civilizations in the central valley of Mexico, Central America, and the Andes that had superior technology and living standards to North America or places such as Argentina and Chile. While the geography stayed the same, the institutions imposed by European colonists created a "reversal of fortune." Geography is also unlikely to explain the poverty of the Middle East for similar reasons. After all, the Middle East led the world in the Neolithic Revolution, and the first towns developed in modern Iraq. Iron was first smelted in

Turkey, and as late as the Middle Ages the Middle East was technologically dynamic. It was not the geography of the Middle East that made the Neolithic Revolution flourish in that part of the world, as we will see in chapter 5, and it was, again, not geography that made the Middle East poor. Instead, it was the expansion and consolidation of the Ottoman Empire, and it is the institutional legacy of this empire that keeps the Middle East poor today.



Map 4: The historical distribution of wild cattle and pigs



Map 5: The historical distribution of wild rice, wheat, and barley

Finally, geographic factors are unhelpful for explaining not only the differences we see across various parts of the world today but also why many nations such as Japan or China stagnate for long periods and then start a rapid growth process. We need another, better theory.

THE CULTURE HYPOTHESIS

The second widely accepted theory, the culture hypothesis, relates prosperity to culture. The culture hypothesis, just like the geography hypothesis, has a distinguished lineage, going back at least to the great German sociologist Max Weber, who argued that the Protestant Reformation and the Protestant ethic it spurred played a key role in facilitating the rise of modern industrial society in Western Europe. The culture hypothesis no longer relies solely on religion, but stresses other types of beliefs, values, and ethics as well.

Though it is not politically correct to articulate in public, many people still maintain that Africans are poor because they lack a good work ethic, still believe in witchcraft and magic, or resist new Western technologies. Many also believe that Latin America will never be rich because its people are intrinsically profligate and impecunious, and because they suffer from some "Iberian" or *"mañana"* culture. Of course, many once believed that the Chinese culture and Confucian values were inimical to economic growth, though now the importance of the Chinese work ethic as the engine of growth in China, Hong Kong, and Singapore is trumpeted.

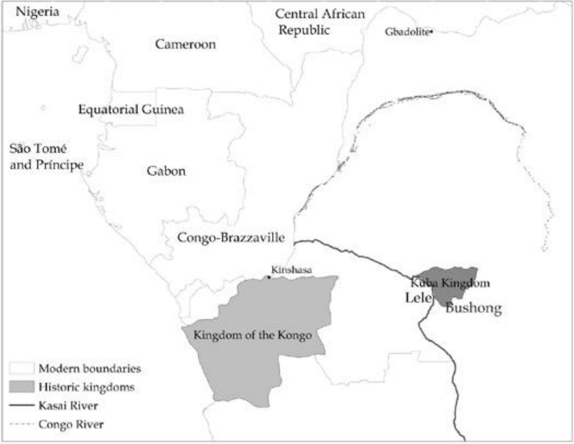
Is the culture hypothesis useful for understanding world inequality? Yes and no. Yes, in the sense that social norms, which are related to culture, matter and can be hard to change, and they also sometimes support institutional differences, this book's explanation for world inequality. But mostly no, because those aspects of culture often emphasized—religion, national ethics, African or Latin values—are just not important for understanding how we got here and why the inequalities in the world persist. Other aspects, such as the extent to which people trust each other or are able to cooperate, are important but they are mostly an outcome of institutions, not an independent cause.

Let us go back to Nogales. As we noted earlier, many aspects of culture are the same north and south of the fence. Nevertheless, there may be some marked differences in practices, norms, and values, though these are not causes but outcomes of the two places' divergent development paths. For example, in surveys Mexicans typically say they trust other people less than the citizens of the United States say they trust others. But it is not a surprise that Mexicans lack trust when their government cannot eliminate drug cartels or provide a functioning unbiased legal system. The same is true with North and South Korea, as we discuss in the next chapter. The South is one of the richest countries in the world, while the North grapples with periodic famine and abject poverty. While "culture" is very different between the South and the North today, it played no role in causing the diverging economic fortunes of these two half nations. The Korean peninsula has a long period of common history. Before the Korean War and the division at the 38th parallel, it had an unprecedented homogeneity in terms of language, ethnicity, and culture. Just as in Nogales, what matters is the border. To the north is a different regime, imposing different institutions, creating different incentives. Any difference in culture between south and north of the border cutting through the two parts of Nogales or the two parts of Korea is thus not a cause of the differences in prosperity but, rather, a consequence.

What about Africa and African culture? Historically, sub-Saharan Africa was poorer than most other parts of the world, and its ancient civilizations did not develop the wheel, writing (with the exception of Ethiopia and Somalia), or the plow. Though these technologies were not widely used until the advent of formal European colonization in the late nineteenth and early twentieth century, African societies knew about them much earlier. Europeans began sailing around the west coast in the late fifteenth century, and Asians were continually sailing to East Africa from much earlier times.

We can understand why these technologies were not adopted from the history of the Kingdom of Kongo at the mouth of the Congo River, which has given its name to the modern Democratic Republic of Congo. Map 6 shows where the Kongo was along with another important central African state, the Kuba Kingdom, which we discuss later in the book.

Kongo came into intense contact with the Portuguese after it was first visited by the mariner Diogo Cão in 1483. At the time, Kongo was a highly centralized polity by African standards, whose capital, Mbanza, had a population of sixty thousand, which made it about the same size as the Portuguese capital of Lisbon and larger than London, which had a population of about fifty thousand in 1500. The king of Kongo, Nzinga a Nkuwu, converted to Catholicism and changed his name to João I. Later Mbanza's name was changed to São Salvador. Thanks to the Portuguese, the Kongolese learned about the wheel and the plow, and the Portuguese even encouraged their adoption with agricultural missions in 1491 and 1512. But all these initiatives failed. Still, the Kongolese were far from averse to modern technologies in general. They were very quick to adopt one venerable Western innovation: the gun. They used this new and powerful tool to respond to market incentives: to capture and export slaves. There is no sign here that African values or culture prevented the adoption of new technologies and practices. As their contacts with Europeans deepened, the Kongolese adopted other Western practices: literacy, dress styles, and house designs. In the nineteenth century, many African societies also took advantage of the rising economic opportunities created by the Industrial Revolution by changing their production patterns. In West Africa there was rapid economic development based on the export of palm oil and ground nuts; throughout southern Africa, Africans developed exports to the rapidly expanding industrial and mining areas of the Rand in South Africa. Yet these promising economic experiments were obliterated not by African culture or the inability of ordinary Africans to act in their own self-interest, but first by European colonialism and then by postindependence African governments.



Map 6: Kingdom of the Kongo, Kuba Kingdom, the Bushong, and the Lele

The real reason that the Kongolese did not adopt superior technology was because they lacked any incentives to do so. They faced a high risk of all their output being expropriated and taxed by the all-powerful king, whether or not he had converted to Catholicism. In fact, it wasn't only their property that was insecure. Their continued existence was held by a thread. Many of them were captured and sold as slaves—hardly the environment to encourage investment to increase long-term productivity. Neither did the king have incentives to adopt the plow on a large scale or to make increasing agricultural productivity his main priority; exporting slaves was so much more profitable.

It might be true today that Africans trust each other less than people in other parts of the world. But this is an outcome of a long history of institutions which have undermined human and property rights in Africa. The potential to be captured and sold as a slave no doubt influenced the extent to which Africans trusted others historically.

What about Max Weber's Protestant ethic? Though it may be true that predominantly Protestant countries, such as the Netherlands and England, were the first economic successes of the modern era, there is little relationship between religion and economic success. France, a predominantly Catholic country, quickly mimicked the economic performance of the Dutch and English in the nineteenth century, and Italy is as prosperous as any of these nations today. Looking farther east, you'll see that none of the economic successes of East Asia have anything to do with any form of Christian religion, so there is not much support for a special relationship between Protestantism and economic success there, either.

Let's turn to a favorite area for the enthusiasts of the culture hypothesis: the Middle East. Middle Eastern countries are primarily Islamic, and the non–oil producers among them are very poor, as we have already noted. Oil producers are richer, but this windfall of wealth has done little to create diversified modern economies in Saudi Arabia or Kuwait. Don't these facts show convincingly that religion matters? Though plausible, this argument is not right, either. Yes, countries such as Syria and Egypt are poor, and their populations are primarily Muslim. But these countries also systemically differ in other ways that are far more important for prosperity. For one, they were all provinces of the Ottoman Empire, which heavily, and adversely, shaped the way they developed. After Ottoman rule collapsed, the Middle East was absorbed into the English and French colonial empires, which, again, stunted their possibilities. After independence, they followed much of the former colonial world by developing hierarchical, authoritarian political regimes with few of the political and economic institutions that, we will argue, are crucial for generating economic success. This development path was forged largely by the history of Ottoman and European rule. The relationship between the Islamic religion and poverty in the Middle East is largely spurious.

The role of these historical events, rather than cultural factors, in shaping the Middle East's economic trajectory is also seen in the fact that the parts of the Middle East that temporarily broke away from the hold of the Ottoman Empire and the European powers, such as Egypt between 1805 and 1848 under Muhammad Ali, could embark on a path of rapid economic change. Muhammad Ali usurped power following the withdrawal of the French forces that had occupied Egypt under Napoleon Bonaparte. Exploiting the weakness of the Ottoman hold over the Egyptian territory at the time, he was able to found his own dynasty, which would, in one form or another, rule until the Egyptian Revolution under Nasser in 1952. Muhammad Ali's reforms, though coercive, did bring growth to Egypt as the state bureaucracy, the army, and the tax system were modernized and there was growth in agriculture and industry. Nevertheless, this process of modernization and growth came to an end after Ali's death, as Egypt fell under European influence.

But perhaps this is the wrong way to think about culture. Maybe the cultural factors that matter are not tied to religion but rather to particular "national cultures." Perhaps it is the influence of English culture that is important and explains why countries such as the United States, Canada, and Australia are so prosperous? Though this idea sounds initially appealing, it doesn't work, either. Yes, Canada and the United States were English colonies, but so were Sierra Leone and Nigeria. The variation in prosperity within former English colonies is as great as that in the entire world. The English legacy is not the reason for the success of North America.

There is yet one more version of the culture hypothesis: perhaps it is not English versus non-English that matters but, rather, European versus non-European. Could it be that Europeans are superior somehow because of their work ethic, outlook on life, Judeo-Christian values, or Roman heritage? It is true that Western Europe and North America, filled primarily by people of European descent, are the most prosperous parts of the world. Perhaps it is the superior European cultural legacy that is at the root of prosperity—and the last refuge of the culture hypothesis. Alas, this version of the culture hypothesis has as little explanatory potential as the others. A greater proportion of the population of Argentina and Uruguay, compared with the population of Canada and the United States, is of

European descent, but Argentina's and Uruguay's economic performance leaves much to be desired. Japan and Singapore never had more than a sprinkling of inhabitants of European descent, yet they are as prosperous as many parts of Western Europe.

China, despite many imperfections in its economic and political system, has been the most rapidly growing nation of the past three decades. Chinese poverty until Mao Zedong's death had nothing to do with Chinese culture; it was due to the disastrous way Mao organized the economy and conducted politics. In the 1950s, he promoted the Great Leap Forward, a drastic industrialization policy that led to mass starvation and famine. In the 1960s, he propagated the Cultural Revolution, which led to the mass persecution of intellectuals and educated people—anyone whose party loyalty might be doubted. This again led to terror and a huge waste of the society's talent and resources. In the same way, current Chinese growth has nothing to do with Chinese values or changes in Chinese culture; it results from a process of economic transformation unleashed by the reforms implemented by Deng Xiaoping and his allies, who, after Mao Zedong's death, gradually abandoned socialist economic policies and institutions, first in agriculture and then in industry.

Just like the geography hypothesis, the culture hypothesis is also unhelpful for explaining other aspects of the lay of the land around us today. There are of course differences in beliefs, cultural attitudes, and values between the United States and Latin America, but just like those that exist between Nogales, Arizona, and Nogales, Sonora, or those between South and North Korea, these differences are a consequence of the two places' different institutions and institutional histories. Cultural factors that emphasize how "Hispanic" or "Latin" culture molded the Spanish Empire can't explain the differences within Latin America—for example, why Argentina and Chile are more prosperous than Peru and Bolivia. Other types of cultural arguments—for instance, those that stress contemporary indigenous culture—fare equally badly. Argentina and Chile have few indigenous people compared with Peru and Bolivia. Though this is true, indigenous culture as an explanation does not work, either. Colombia, Ecuador, and Peru have similar income levels, but Colombia has very few indigenous people today, while Ecuador and Peru have many. Finally, cultural attitudes, which are in general slow to change, are unlikely to account by themselves for the growth miracles in East Asia and China. Though institutions are persistent, too, in certain circumstances they do change rapidly, as we'll see.

The Ignorance Hypothesis

The final popular theory for why some nations are poor and some are rich is the ignorance hypothesis, which asserts that world inequality exists because we or our rulers do not know how to make poor countries rich. This idea is the one held by most economists, who take their cue from the famous definition proposed by the English economist Lionel Robbins in 1935 that "economics is a science which studies human behavior as a relationship between ends and scarce means which have alternative uses."

It is then a small step to conclude that the science of economics should focus on the best use of scarce means to satisfy social ends. Indeed, the most famous theoretical result in economics, the so-called First Welfare Theorem, identifies the circumstances under which the allocation of resources in a "market economy" is socially desirable from an economic point of view. A market economy is an abstraction that is meant to capture a situation in which all individuals and firms can freely produce, buy, and sell any products or services that they wish. When these circumstances are not present there is a "market failure." Such failures provide the basis for a theory of world inequality, since the more that market failures go unaddressed, the poorer a country is likely to be. The ignorance hypothesis maintains that poor countries are poor because they have a lot of market failures and because economists and policymakers do not know how to get rid of them and have heeded the wrong advice in the past. Rich countries are rich because they have figured out better policies and have successfully eliminated these failures.

Could the ignorance hypothesis explain world inequality? Could it be that African countries are poorer than the rest of the world because their leaders tend to have the same mistaken views of how to run their countries, leading to the poverty there, while Western European leaders are better informed or better advised, which explains their relative success? While there are famous examples of leaders adopting disastrous policies because they were mistaken about those policies' consequences, ignorance can explain at best a small part of world inequality.

On the face of it, the sustained economic decline that soon set in in Ghana after independence from Britain was caused by ignorance. The British economist Tony Killick, then working as an adviser for the government of Kwame Nkrumah, recorded many of the problems in great detail. Nkrumah's policies focused on developing state industry, which turned out to be very inefficient. Killick recalled:

> The footwear factory … that would have linked the meat factory in the North through transportation of the hides to the South (for a distance of over 500 miles) to a tannery (now abandoned); the leather was to have been backhauled to the footwear factory in Kumasi, in the center of the country and about 200 miles north of the tannery. Since the major footwear market is in the Accra metropolitan area, the shoes would then have to be transported an additional 200 miles back to the South.

Killick somewhat understatedly remarks that this was an enterprise "whose viability was undermined by poor siting." The footwear factory was one of many such projects, joined by the mango canning plant situated in a part of Ghana which did not grow mangos and whose output was to be more than the entire world demand for the product. This endless stream of economically irrational developments was not caused by the fact that Nkrumah or his advisers were badly informed or ignorant of the right economic policies. They had people like Killick and had even been advised by Nobel laureate Sir Arthur Lewis, who knew the policies were not good. What drove the form the economic policies took was the fact that Nkrumah needed to use them to buy political support and sustain his undemocratic regime.

Neither Ghana's disappointing performance after independence nor the countless other cases of apparent economic mismanagement can simply be blamed on ignorance. After all, if ignorance were the problem, well-meaning leaders would quickly learn what types of policies increased their citizens' incomes and welfare, and would gravitate toward those policies.

Consider the divergent paths of the United States and Mexico. Blaming this disparity on the ignorance of the leaders of the two nations is, at best, highly implausible. It wasn't differences in knowledge or intentions between John Smith and Cortés that laid the seeds of divergence during the colonial period, and it wasn't differences in knowledge between later U.S. presidents, such as Teddy Roosevelt or Woodrow Wilson, and Porfirio Díaz that made Mexico choose economic institutions that enriched elites at the expense of the rest of society at the end of the nineteenth and beginning of the twentieth centuries while Roosevelt and Wilson did the opposite. Rather, it was the differences in the institutional constraints the countries' presidents and elites were facing. Similarly, leaders of African nations that have languished over the last half century under insecure property rights and economic institutions, impoverishing much of their populations, did not allow this to happen because they thought it was good economics; they did so because they could get away with it and enrich themselves at the expense of the rest, or because they thought it was good politics, a way of keeping themselves in power by buying the support of crucial groups or elites.

The experience of Ghana's prime minister in 1971, Kofi Busia, illustrates how misleading the ignorance hypothesis can be. Busia faced a dangerous economic crisis. After coming to power in 1969, he, like Nkrumah before him, pursued unsustainable expansionary economic policies and maintained various price controls through marketing boards and an overvalued exchange rate. Though Busia had been an opponent of Nkrumah, and led a democratic government, he faced many of the same political constraints. As with Nkrumah, his economic policies were adopted not because he was "ignorant" and believed that these policies were good economics or an ideal way to develop the country. The policies were chosen because they were good politics, enabling Busia to transfer resources to politically powerful groups, for example in urban areas, who needed to be kept contented. Price controls squeezed agriculture, delivering cheap food to the urban constituencies and generating revenues to finance government spending. But these controls were unsustainable. Ghana was soon suffering from a series of balance-of-payment crises and foreign exchange shortages. Faced with these dilemmas, on December 27, 1971, Busia signed an agreement with the International Monetary Fund that included a massive devaluation of the currency.

The IMF, the World Bank, and the entire international community put pressure on Busia to implement the reforms contained in the agreement. Though the international institutions were blissfully unaware, Busia knew he was taking a huge political gamble. The immediate consequence of the currency's devaluation was rioting and discontent in Accra, Ghana's capital, that mounted uncontrollably until Busia was overthrown by the military, led by Lieutenant Colonel Acheampong, who immediately reversed the devaluation.

The ignorance hypothesis differs from the geography and culture hypotheses in that it comes readily with a suggestion about how to "solve" the problem of poverty: if ignorance got us here, enlightened and informed rulers and policymakers can get us out and we should be able to "engineer" prosperity around the world by providing the right advice and by convincing politicians of what is good economics. Yet Busia's experience underscores the fact that the main obstacle to the adoption of policies that would reduce market failures and encourage economic growth is not the ignorance of politicians but the incentives and constraints they face from the political and economic institutions in their societies.

Although the ignorance hypothesis still rules supreme among most economists and in Western policymaking circles—which, almost to the exclusion of anything else, focus on how to engineer prosperity—it is just another hypothesis that doesn't work. It explains neither the origins of prosperity around the world nor the lay of the land around us—for example, why some nations, such as Mexico and Peru, but not the United States or England, adopted institutions and policies that would impoverish the majority of their citizens, or why almost all sub-Saharan Africa and most of Central America are so much poorer than Western Europe or East Asia.

When nations break out of institutional patterns condemning them to poverty and manage to embark on a path to economic growth, this is not because their ignorant leaders suddenly have become better informed or less self-interested or because they've received advice from better economists. China, for example, is one of the countries that made the switch from economic policies that caused poverty and the starvation of millions to those encouraging economic growth. But, as we will discuss in greater detail later, this did not happen because the Chinese Communist Party finally understood that the collective ownership of agricultural land and industry created terrible economic incentives. Instead, Deng Xiaoping and his allies, who were no less self-interested than their rivals but who had different interests and political objectives, defeated their powerful opponents in the Communist Party and masterminded a political revolution of sorts, radically changing the leadership and direction of the party. Their economic reforms, which created market incentives in agriculture and then subsequently in industry, followed from this political revolution. It was politics that determined the switch from communism and toward market incentives in China, not better advice or a better understanding of how the economy worked.

WE WILL ARGUE that to understand world inequality we have to understand why some societies are organized in very inefficient and socially undesirable ways. Nations sometimes do manage to adopt efficient institutions and achieve prosperity, but alas, these are the rare cases. Most economists and policymakers have focused on "getting it right," while what is really needed is an explanation for why poor nations "get it wrong." Getting it wrong is mostly not about ignorance or culture. As we will show, poor countries are poor because those who have power make choices that create poverty. They get it wrong not by mistake or ignorance but on purpose. To understand this, you have to go beyond economics and expert advice on the best thing to do and, instead, study how decisions actually get made, who gets to make them, and why those people decide to do what they do. This is the study of politics and political processes. Traditionally economics has ignored politics, but understanding politics is crucial for explaining world inequality. As the economist Abba Lerner noted in the 1970s, “Economics has gained the title Queen of the Social Sciences by choosing solved political problems as its domain.”

We will argue that achieving prosperity depends on solving some basic political problems. It is precisely because economics has assumed that political problems are solved that it has not been able to come up with a convincing explanation for world inequality. Explaining world inequality still needs economics to understand how different types of policies and social arrangements affect economic incentives and behavior. But it also needs politics.

3.

THE MAKING OF PROSPERITY AND POVERTY

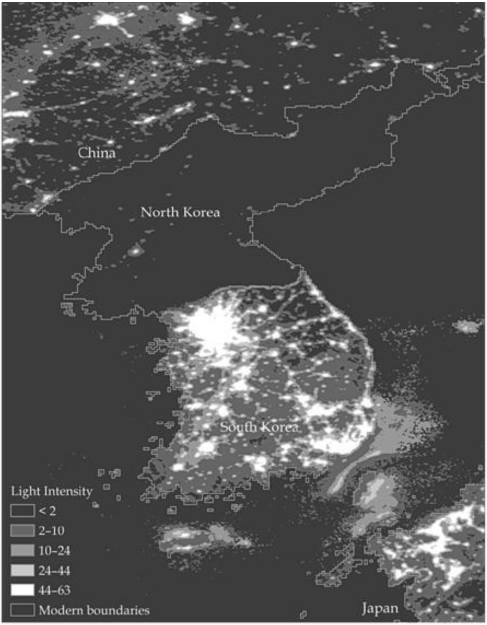
THE ECONOMICS OF THE 38TH PARALLEL

IN THE SUMMER OF 1945, as the Second World War was drawing to a close, the Japanese colony in Korea began to collapse. Within a month of Japan's August 15 unconditional surrender, Korea was divided at the 38th parallel into two spheres of influence. The South was administered by the United States. The North, by Russia. The uneasy peace of the cold war was shattered in June 1950 when the North Korean army invaded the South. Though initially the North Koreans made large inroads, capturing the capital city, Seoul, by the autumn, they were in full retreat. It was then that Hwang Pyŏng-Wŏn and his brother were separated. Hwang Pyŏng-Wŏn managed to hide and avoid being drafted into the North Korean army. He stayed in the South and worked as a pharmacist. His brother, a doctor working in Seoul treating wounded soldiers from the South Korean army, was taken north as the North Korean army retreated. Dragged apart in 1950, they met again in 2000 in Seoul for the first time in fifty years, after the two governments finally agreed to initiate a limited program of family reunification.

As a doctor, Hwang Pyŏng-Wŏn's brother had ended up working for the air force, a good job in a military dictatorship. But even those with privileges in North Korea don't do that well. When the brothers met, Hwang Pyŏng-Wŏn asked about how life was north of the 38th parallel. He had a car, but his brother didn't. "Do you have a telephone?" he asked his brother. "No," said his brother. "My daughter, who works at the Foreign Ministry, has a phone, but if you don't know the code you can't call." Hwang Pyŏng-Wŏn recalled how all the people from the North at the reunion were asking for money, so he offered some to his brother. But his brother said, "If I go back with money the government will say, 'Give that money to us,' so keep it." Hwang Pyŏng-Wŏn noticed his brother's coat was threadbare: "Take off that coat and leave it, and when you go back wear this one," he suggested. "I can't do that," his brother replied. "This is just borrowed from the government to come here." Hwang Pyŏng-Wŏn recalled how when they parted, his brother was ill at ease and always nervous as though someone were listening. He was poorer than Hwang Pyŏng-Wŏn imagined. His brother said he lived well, but Hwang Pyŏng-Wŏn thought he looked awful and was thin as a rake.

The people of South Korea have living standards similar to those of Portugal and Spain. To the north, in the so-called Democratic People's Republic of Korea, or North Korea, living standards are akin to those of a sub-Saharan African country, about one-tenth of average living standards in South Korea. The health of North Koreans is in an even worse state; the average North Korean can expect to live ten years less than his cousins south of the 38th parallel. Map 7 illustrates in a dramatic way the economic gap between the Koreas. It plots data on the intensity of light at night from satellite images. North Korea is almost completely dark due to lack of electricity; South Korea is blazing with light.

These striking differences are not ancient. In fact, they did not exist prior to the end of the Second World War. But after 1945, the different governments in the North and the South adopted very different ways of organizing their economies. South Korea was led, and its early economic and political institutions were shaped, by the Harvard- and Princeton-educated, staunchly anticommunist Syngman Rhee, with significant support from the United States. Rhee was elected president in 1948. Forged in the midst of the Korean War and against the threat of communism spreading to the south of the 38th parallel, South Korea was no democracy. Both Rhee and his equally famous successor, General Park Chung-Hee, secured their places in history as authoritarian presidents. But both governed a market economy where private property was recognized, and after 1961, Park effectively threw the weight of the state behind rapid economic growth, channeling credit and subsidies to firms that were successful.



Map 7: Lights in South Korea and darkness in the North

The situation north of the 38th parallel was different. Kim Il-Sung, a leader of anti-Japanese communist partisans during the Second World War, established himself as dictator by 1947 and, with the help of the Soviet Union, introduced a rigid form of centrally planned economy as part of the so-called Juche system. Private property was outlawed, and markets were banned. Freedoms were curtailed not only in the marketplace, but in every sphere of

North Koreans' lives—except for those who happened to be part of the very small ruling elite around Kim Il-Sung and, later, his son and successor Kim Jong-Il.

It should not surprise us that the economic fortunes of South and North Korea diverged sharply. Kim Il-Sung's command economy and the Juche system soon proved to be a disaster. Detailed statistics are not available from North Korea, which is a secretive state, to say the least. Nonetheless, available evidence confirms what we know from the all-too-often recurring famines: not only did industrial production fail to take off, but North Korea in fact experienced a collapse in agricultural productivity. Lack of private property meant that few people had incentives to invest or to exert effort to increase or even maintain productivity. The stifling, repressive regime was inimical to innovation and the adoption of new technologies. But Kim Il-Sung, Kim Jong-Il, and their cronies had no intention of reforming the system, or introducing private property, markets, private contracts, or changing economic and political institutions. North Korea continues to stagnate economically.

Meanwhile, in the South, economic institutions encouraged investment and trade. South Korean politicians invested in education, achieving high rates of literacy and schooling. South Korean companies were quick to take advantage of the relatively educated population, the policies encouraging investment and industrialization, exports, and the transfer of technology. South Korea quickly became one of East Asia's "Miracle Economies," one of the most rapidly growing nations in the world.

By the late 1990s, in just about half a century, South Korean growth and North Korean stagnation led to a tenfold gap between the two halves of this once-united country—imagine what a difference a couple of centuries could make. The economic disaster of North Korea, which led to the starvation of millions, when placed against the South Korean economic success, is striking: neither culture nor geography nor ignorance can explain the divergent paths of North and South Korea. We have to look at institutions for an answer.

Extractive and Inclusive Economic Institutions

Countries differ in their economic success because of their different institutions, the rules influencing how the economy works, and the incentives that motivate people. Imagine teenagers in North and South Korea and what they expect from life. Those in the North grow up in poverty, without entrepreneurial initiative, creativity, or adequate education to prepare them for skilled work. Much of the education they receive at school is pure propaganda, meant to shore up the legitimacy of the regime; there are few books, let alone computers. After finishing school, everyone has to go into the army for ten years. These teenagers know that they will not be able to own property, start a business, or become more prosperous even if many people engage illegally in private economic activities to make a living. They also know that they will not have legal access to markets where they can use their skills or their earnings to purchase the goods they need and desire. They are even unsure about what kind of human rights they will have.

Those in the South obtain a good education, and face incentives that encourage them to exert effort and excel in their chosen vocation. South Korea is a market economy, built on private property. South Korean teenagers know that, if successful as entrepreneurs or workers, they can one day enjoy the fruits of their investments and efforts; they can improve their standard of living and buy cars, houses, and health care.

In the South the state supports economic activity. So it is possible for entrepreneurs to borrow money from banks and financial markets, for foreign companies to enter into partnerships with South Korean firms, for individuals to take up mortgages to buy houses. In the South, by and large, you are free to open any business you like. In the North, you are not. In the South, you can hire workers, sell your products or services, and spend your money in the marketplace in whichever way you want. In the North, there are only black markets. These different rules are the institutions under which North and South Koreans live.

Inclusive economic institutions, such as those in South Korea or in the United States, are those that allow and encourage participation by the great mass of people in economic activities that make best use of their talents and skills and that enable individuals to make the choices they wish. To be inclusive, economic institutions must feature secure private property, an unbiased system of law, and a provision of public services that provides a level playing field in which people can exchange and contract; it also must permit the entry of new businesses and allow people to choose their careers.

THE CONTRAST OF South and North Korea, and of the United States and Latin America, illustrates a general principle. Inclusive economic institutions foster economic activity, productivity growth, and economic prosperity. Secure private property rights are central, since only those with such rights will be willing to invest and increase productivity. A businessman who expects his output to be stolen, expropriated, or entirely taxed away will have little incentive to work, let alone any incentive to undertake investments and innovations. But such rights must exist for the majority of people in society.

In 1680 the English government conducted a census of the population of its West Indian colony of Barbados. The census revealed that of the total population on the island of around 60,000, almost 39,000 were African slaves who were the property of the remaining one-third of the population. Indeed, they were mostly the property of the largest 175 sugar planters, who also owned most of the land. These large planters had secure and well-enforced property rights over their land and even over their slaves. If one planter wanted to sell slaves to another, he could do so and expect a court to enforce such a sale or any other contract he wrote. Why? Of the forty judges and justices of the peace on the island, twenty-nine of them were large planters. Also, the eight most senior military officials were all large planters. Despite well-defined, secure, and enforced property rights and contracts for the island's elite, Barbados did not have inclusive economic institutions, since two-thirds of the population were slaves with no access to education or economic opportunities, and no ability or incentive to use their talents or skills. Inclusive economic institutions require secure property rights and economic opportunities not just for the elite but for a broad cross-section of society.

Secure property rights, the law, public services, and the freedom to contract and exchange all rely on the state, the institution with the coercive capacity to impose order, prevent theft and fraud, and enforce contracts between private parties. To function well, society also needs other public services: roads and a transport network so that goods can be transported; a public infrastructure so that economic activity can flourish; and some type of basic regulation to prevent fraud and malfeasance. Though many of these public services can be provided by markets and private citizens, the degree of coordination necessary to do so on a large scale often eludes all but a central authority. The state is thus inexorably intertwined with economic institutions, as the enforcer of law and order, private property, and contracts, and often as a key provider of public services. Inclusive economic institutions need and use the state.

The economic institutions of North Korea or of colonial Latin America—the *mita*, *encomienda*, or *repartimiento* described earlier—do not have these properties. Private property is nonexistent in North Korea. In colonial Latin America there was private property for Spaniards, but the property of the indigenous peoples was highly insecure. In neither type of society was the vast mass of people able to make the economic decisions they wanted to; they were subject to mass coercion. In neither type of society was the power of the state used to provide key public services that promoted prosperity. In North Korea, the state built an education system to inculcate propaganda, but was unable to prevent famine. In colonial Latin America, the state focused on coercing indigenous peoples. In neither type of society was there a level playing field or an unbiased legal system. In North Korea, the legal system is an arm of the ruling Communist Party, and in Latin America it was a tool of discrimination against the mass of people. We call such institutions, which have opposite properties to those we call inclusive, extractive economic institutions—extractive because such institutions are designed to extract incomes and wealth from one subset of society to benefit a different subset.

ENGINES OF PROSPERITY

Inclusive economic institutions create inclusive markets, which not only give people freedom to pursue the vocations in life that best suit their talents but also provide a level playing field that gives them the opportunity to do so. Those who have good ideas will be able to start businesses, workers will tend to go to activities where their productivity is greater, and less efficient firms can be replaced by more efficient ones. Contrast how people choose their occupations under inclusive markets to colonial Peru and Bolivia, where under the *mita*, many were forced to work in silver and mercury mines, regardless of their skills or whether they wanted to. Inclusive markets are not just free markets. Barbados in the seventeenth century also had markets. But in the same way that it lacked property rights for all but the narrow planter elite, its markets were far from inclusive; markets in slaves were in fact one part of the economic institutions systematically coercing the majority of the population and robbing them of the ability to choose their occupations and how they should utilize their talents.

Inclusive economic institutions also pave the way for two other engines of prosperity: technology and education. Sustained economic growth is almost always accompanied by technological improvements that enable people (labor), land, and existing capital (buildings, existing machines, and so on) to become more productive. Think of our great-great-grandparents, just over a century ago, who did not have access to planes or automobiles or most of the drugs and health care we now take for granted, not to mention indoor plumbing, air-conditioning, shopping malls, radio, or motion pictures; let alone information technology, robotics, or computer-controlled machinery. And going back a few more generations, the technological know-how and living standards were even more backward, so much so that we would find it hard to imagine how most people struggled through life. These improvements follow from science and from entrepreneurs such as Thomas Edison, who applied science to create profitable businesses. This process of innovation is made possible by economic institutions that encourage private property, uphold contracts, create a level playing field, and encourage and allow the entry of new businesses that can bring new technologies to life. It should therefore be no surprise that it was U.S. society, not Mexico or Peru, that produced Thomas Edison, and that it was South Korea, not North Korea, that today produces technologically innovative companies such as Samsung and Hyundai.

Intimately linked to technology are the education, skills, competencies, and know-how of the workforce, acquired in schools, at home, and on the job. We are so much more productive than a century ago not just because of better technology embodied in machines but also because of the greater know-how that workers possess. All the technology in the world would be of little use without workers who knew how to operate it. But there is more to skills and competencies than just the ability to run machines. It is the education and skills of the workforce that generate the scientific knowledge upon which our progress is built and that enable the adaptation and adoption of these technologies in diverse lines of business. Though we saw in chapter 1 that many of the innovators of the Industrial Revolution and afterward, like Thomas Edison, were not highly educated, these innovations were much simpler than modern technology. Today technological change requires education both for the innovator and the worker. And here we see the importance of economic institutions that create a level playing field. The United States could produce, or attract from foreign lands, the likes of Bill Gates, Steve Jobs, Sergey Brin, Larry Page, and Jeff Bezos, and the hundreds of scientists who made fundamental discoveries in information technology, nuclear power, biotech, and other fields upon which these entrepreneurs built their businesses. The supply of talent was there to be harnessed because most teenagers in the United States have access to as much schooling as they wish or are capable of attaining. Now imagine a different society, for example the Congo or Haiti, where a large fraction of the population has no means of attending school, or where, if they manage to go to school, the quality of teaching is lamentable, where teachers do not show up for work, and even if they do, there may not be any books.

The low education level of poor countries is caused by economic institutions that fail to create incentives for parents to educate their children and by political institutions that fail to induce the government to build, finance, and support schools and the wishes of parents and children. The price these nations pay for low education of their population and lack of inclusive markets is high. They fail to mobilize their nascent talent. They have many potential Bill Gateses and perhaps one or two Albert Einsteins who are now working as poor, uneducated farmers, being coerced to do what they don't want to do or being drafted into the army, because they never had the opportunity to realize their vocation in life.

The ability of economic institutions to harness the potential of inclusive markets, encourage technological innovation, invest in people, and mobilize the talents and skills of a large number of individuals is critical for economic growth. Explaining why so many economic institutions fail to meet these simple objectives is the central theme of this book.

EXTRACTIVE AND INCLUSIVE POLITICAL INSTITUTIONS

All economic institutions are created by society. Those of North Korea, for example, were forced on its citizens by the communists who took over the country in the 1940s, while those of colonial Latin America were imposed by Spanish conquistadors. South Korea ended up with very different economic institutions than the North because different people with different interests and objectives made the decisions about how to structure society. In other words, South Korea had different politics.

Politics is the process by which a society chooses the rules that will govern it. Politics surrounds institutions for the simple reason that while inclusive institutions may be good for the economic prosperity of a nation, some people or groups, such as the elite of the Communist Party of North Korea or the sugar planters of colonial Barbados, will be much better off by setting up institutions that are extractive. When there is conflict over institutions, what happens depends on which people or group wins out in the game of politics—who can get more support, obtain additional resources, and form more effective alliances. In short, who wins depends on the distribution of political power in society.

The political institutions of a society are a key determinant of the outcome of this game. They are the rules that govern incentives in politics. They determine how the government is chosen and which part of the government has the right to do what. Political institutions determine who has power in society and to what ends that power can be used. If the distribution of power is narrow and unconstrained, then the political institutions are absolutist, as exemplified by the absolutist monarchies reigning throughout the world during much of history. Under absolutist political institutions such as those in North Korea and colonial Latin America, those who can wield this power will be able to set up economic institutions to enrich themselves and augment their power at the expense of society. In contrast, political institutions that distribute power broadly in society and subject it to constraints are pluralistic. Instead of being vested in a single individual or a narrow group, political power rests with a broad coalition or a plurality of groups.

There is obviously a close connection between pluralism and inclusive economic institutions. But the key to understanding why South Korea and the United States have inclusive economic institutions is not just their pluralistic political institutions but also their sufficiently centralized and powerful states. A telling contrast is with the East African nation of Somalia. As we will see later in the book, political power in Somalia has long been widely distributed—almost pluralistic. Indeed there is no real authority that can control or sanction what anyone does. Society is divided into deeply antagonistic clans that cannot dominate one another. The power of one clan is constrained only by the guns of another. This distribution of power leads not to inclusive institutions but to chaos, and at the root of it is the Somali state's lack of any kind of political centralization, or state centralization, and its inability to enforce even the minimal amount of law and order to support economic activity, trade, or even the basic security of its citizens.

Max Weber, who we met in the previous chapter, provided the most famous and widely accepted definition of the state, identifying it with the “monopoly of legitimate violence” in society. Without such a monopoly and the degree of centralization that it entails, the state cannot play its role as enforcer of law and order, let alone provide public services and encourage and regulate economic activity. When the state fails to achieve almost any political centralization, society sooner or later descends into chaos, as did Somalia.

We will refer to political institutions that are sufficiently centralized and pluralistic as inclusive political institutions. When either of these conditions fails, we will refer to the institutions as extractive political institutions.

There is strong synergy between economic and political institutions. Extractive political institutions concentrate power in the hands of a narrow elite and place few constraints on the exercise of this power. Economic institutions are then often structured by this elite to extract resources from the rest of the society. Extractive economic institutions thus naturally accompany extractive political institutions. In fact, they must inherently depend on extractive political institutions for their survival. Inclusive political institutions, vesting power broadly, would tend to uproot economic institutions that expropriate the resources of the many, erect entry barriers, and suppress the functioning of markets so that only a few benefit.

In Barbados, for example, the plantation system based on the exploitation of slaves could not have survived without political institutions that suppressed and completely excluded the slaves from the political process. The economic system impoverishing millions for the benefit of a narrow communist elite in North Korea would also be unthinkable without the total political domination of the Communist Party.

This synergistic relationship between extractive economic and political institutions introduces a strong feedback loop: political institutions enable the elites controlling political power to choose economic institutions with few constraints or opposing forces. They also enable the elites to structure future political institutions and their evolution. Extractive economic institutions, in turn, enrich the same elites, and their economic wealth and power help consolidate their political dominance. In Barbados or in

Latin America, for example, the colonists were able to use their political power to impose a set of economic institutions that made them huge fortunes at the expense of the rest of the population. The resources these economic institutions generated enabled these elites to build armies and security forces to defend their absolutist monopoly of political power. The implication of course is that extractive political and economic institutions support each other and tend to persist.

There is in fact more to the synergy between extractive economic and political institutions. When existing elites are challenged under extractive political institutions and the newcomers break through, the newcomers are likewise subject to only a few constraints. They thus have incentives to maintain these political institutions and create a similar set of economic institutions, as Porfirio Díaz and the elite surrounding him did at the end of the nineteenth century in Mexico.

Inclusive economic institutions, in turn, are forged on foundations laid by inclusive political institutions, which make power broadly distributed in society and constrain its arbitrary exercise. Such political institutions also make it harder for others to usurp power and undermine the foundations of inclusive institutions. Those controlling political power cannot easily use it to set up extractive economic institutions for their own benefit. Inclusive economic institutions, in turn, create a more equitable distribution of resources, facilitating the persistence of inclusive political institutions.

It was not a coincidence that when, in 1618, the Virginia Company gave land, and freedom from their draconian contracts, to the colonists it had previously tried to coerce, the General Assembly in the following year allowed the colonists to begin governing themselves. Economic rights without political rights would not have been trusted by the colonists, who had seen the persistent efforts of the Virginia Company to coerce them. Neither would these economies have been stable and durable. In fact, combinations of extractive and inclusive institutions are generally unstable. Extractive economic institutions under inclusive political institutions are unlikely to survive for long, as our discussion of Barbados suggests.

Similarly, inclusive economic institutions will neither support nor be supported by extractive political ones. Either they will be transformed into extractive economic institutions to the benefit of the narrow interests that hold power, or the economic dynamism they create will destabilize the extractive political institutions, opening the way for the emergence of inclusive political institutions. Inclusive economic institutions also tend to reduce the benefits the elites can enjoy by ruling over extractive political institutions, since those institutions face competition in the marketplace and are constrained by the contracts and property rights of the rest of society.

WHY NOT ALWAYS CHOOSE PROSPERITY?

Political and economic institutions, which are ultimately the choice of society, can be inclusive and encourage economic growth. Or they can be extractive and become impediments to economic growth. Nations fail when they have extractive economic institutions, supported by extractive political institutions that impede and even block economic growth. But this means that the choice of institutions—that is, the politics of institutions—is central to our quest for understanding the reasons for the success and failure of nations. We have to understand why the politics of some societies lead to inclusive institutions that foster economic growth, while the politics of the vast majority of societies throughout history has led, and still leads today, to extractive institutions that hamper economic growth.

It might seem obvious that everyone should have an interest in creating the type of economic institutions that will bring prosperity. Wouldn't every citizen, every politician, and even a predatory dictator want to make his country as wealthy as possible?

Let's return to the Kingdom of Kongo we discussed earlier. Though this kingdom collapsed in the seventeenth century, it provided the name for the modern country that became independent from Belgian colonial rule in 1960. As an independent polity, Congo experienced almost unbroken economic decline and mounting poverty under the rule of Joseph Mobutu between 1965 and 1997. This decline continued after Mobutu was overthrown by Laurent Kabila. Mobutu created a highly extractive set of economic institutions. The citizens were impoverished, but Mobutu and the elite surrounding him, known as Les Grosses Legumes (the Big Vegetables), became fabulously wealthy. Mobutu built himself a palace at his birthplace, Gbadolite, in the north of the country, with an airport large enough to land a supersonic Concord jet, a plane he frequently rented from Air France for travel to Europe. In Europe he bought castles and owned large tracts of the Belgian capital of Brussels.

Wouldn't it have been better for Mobutu to set up economic institutions that increased the wealth of the Congolese rather than deepening their poverty? If Mobutu had managed to increase the prosperity of his nation, would he not have been able to appropriate even more money, buy a Concord instead of renting one, have more castles and mansions, possibly a bigger and more powerful army? Unfortunately for the citizens of many countries in the world, the answer is no. Economic institutions that create incentives for economic progress may simultaneously redistribute income and power in such a way that a predatory dictator and others with political power may become worse off.

The fundamental problem is that there will necessarily be disputes and conflict over economic institutions. Different institutions have different consequences for the prosperity of a nation, how that prosperity is distributed, and who has power. The economic growth which can be induced by institutions creates both winners and losers. This was clear during the Industrial Revolution in England, which laid the foundations of the prosperity we see in the rich countries of the world today. It centered on a series of pathbreaking technological changes in steam power, transportation, and textile production. Even though mechanization led to enormous increases in total incomes and ultimately became the foundation of modern industrial society, it was bitterly opposed by many. Not because of ignorance or shortsightedness; quite the opposite. Rather, such opposition to economic growth has its own, unfortunately coherent, logic. Economic growth and technological change are accompanied by what the great economist

Joseph Schumpeter called creative destruction. They replace the old with the new. New sectors attract resources away from old ones. New firms take business away from established ones. New technologies make existing skills and machines obsolete. The process of economic growth and the inclusive institutions upon which it is based create losers as well as winners in the political arena and in the economic marketplace. Fear of creative destruction is often at the root of the opposition to inclusive economic and political institutions.

European history provides a vivid example of the consequences of creative destruction. On the eve of the Industrial Revolution in the eighteenth century, the governments of most European countries were controlled by aristocracies and traditional elites, whose major source of income was from landholdings or from trading privileges they enjoyed thanks to monopolies granted and entry barriers imposed by monarchs. Consistent with the idea of creative destruction, the spread of industries, factories, and towns took resources away from the land, reduced land rents, and increased the wages that landowners had to pay their workers. These elites also saw the emergence of new businessmen and merchants eroding their trading privileges. All in all, they were the clear economic losers from industrialization. Urbanization and the emergence of a socially conscious middle and working class also challenged the political monopoly of landed aristocracies. So with the spread of the Industrial Revolution the aristocracies weren't just the economic losers; they also risked becoming political losers, losing their hold on political power. With their economic and political power under threat, these elites often formed a formidable opposition against industrialization.

The aristocracy was not the only loser from industrialization. Artisans whose manual skills were being replaced by mechanization likewise opposed the spread of industry. Many organized against it, rioting and destroying the machines they saw as responsible for the decline of their livelihood. They were the Luddites, a word that has today become synonymous with resistance to technological change. John Kay, English inventor of the "flying shuttle" in 1733, one of the first significant improvements in the mechanization of weaving, had his house burned down by Luddites in 1753. James Hargreaves, inventor of the "spinning jenny," a complementary revolutionary improvement in spinning, got similar treatment.

In reality, the artisans were much less effective than the landowners and elites in opposing industrialization. The Luddites did not possess the political power—the ability to affect political outcomes against the wishes of other groups—of the landed aristocracy. In England, industrialization marched on, despite the Luddites' opposition, because aristocratic opposition, though real, was muted. In the Austro-Hungarian and the Russian empires, where the absolutist monarchs and aristocrats had far more to lose, industrialization was blocked. In consequence, the economies of Austria-Hungary and Russia stalled. They fell behind other European nations, where economic growth took off during the nineteenth century.

The success and failure of specific groups notwithstanding, one lesson is clear: powerful groups often stand against economic progress and against the engines of prosperity. Economic growth is not just a process of more and better machines, and more and better educated people, but also a transformative and destabilizing process associated with widespread creative destruction. Growth thus moves forward only if not blocked by the economic losers who anticipate that their economic privileges will be lost and by the political losers who fear that their political power will be eroded.

Conflict over scarce resources, income and power, translates into conflict over the rules of the game, the economic institutions, which will determine the economic activities and who will benefit from them. When there is a conflict, the wishes of all parties cannot be simultaneously met. Some will be defeated and frustrated, while others will succeed in securing outcomes they like. Who the winners of this conflict are has fundamental implications for a nation's economic trajectory. If the groups standing against growth are the winners, they can successfully block economic growth, and the economy will stagnate.

The logic of why the powerful would not necessarily want to set up the economic institutions that promote economic success extends easily to the choice of political institutions.

In an absolutist regime, some elites can wield power to set up economic institutions they prefer. Would they be interested in changing political institutions to make them more pluralistic? In general not, since this would only dilute their political power, making it more difficult, maybe impossible, for them to structure economic institutions to further their own interests. Here again we see a ready source of conflict. The people who suffer from the extractive economic institutions cannot hope for absolutist rulers to voluntarily change political institutions and redistribute power in society. The only way to change these political institutions is to force the elite to create more pluralistic institutions.

In the same way that there is no reason why political institutions should automatically become pluralistic, there is no natural tendency toward political centralization. There would certainly be incentives to create more centralized state institutions in any society, particularly in those with no such centralization whatsoever. For example, in Somalia, if one clan created a centralized state capable of imposing order on the country, this could lead to economic benefits and make this clan richer. What stops this? The main barrier to political centralization is again a form of fear from change: any clan, group, or politician attempting to centralize power in the state will also be centralizing power in their own hands, and this is likely to meet the ire of other clans, groups, and individuals, who would be the political losers of this process. Lack of political centralization means not only lack of law and order in much of a territory but also there being many actors with sufficient powers to block or disrupt things, and the fear of their opposition and violent reaction will often deter many would-be centralizers. Political centralization is likely only when one group of people is sufficiently more powerful than others to build a state. In Somalia, power is evenly balanced, and no one clan can impose its will on any other. Therefore, the lack of political centralization persists.

THE LONG AGONY OF THE CONGO

There are few better, or more depressing, examples of the forces that explain the logic of why economic prosperity is so persistently rare under extractive institutions or that illustrate the synergy between extractive economic and political institutions than the Congo. Portuguese and Dutch visitors to Kongo in the fifteenth and sixteenth centuries remarked on the "miserable poverty" there. Technology was rudimentary by European standards, with the Kongolese having neither writing, the wheel, nor the plow. The reason for this poverty, and the reluctance of Kongolese farmers to adopt better technologies when they learned of them, is clear from existing historical accounts. It was due to the extractive nature of the country's economic institutions.

As we have seen, the Kingdom of Kongo was governed by the king in Mbanza, subsequently São Salvador. Areas away from the capital were ruled by an elite who played the roles of governors of different parts of the kingdom. The wealth of this elite was based on slave plantations around São Salvador and the extraction of taxes from the rest of the country. Slavery was central to the economy, used by the elite to supply their own plantations and by Europeans on the coast. Taxes were arbitrary; one tax was even collected every time the king's beret fell off. To become more prosperous, the Kongolese people would have had to save and invest—for example, by buying plows. But it would not have been worthwhile, since any extra output that they produced using better technology would have been subject to expropriation by the king and his elite. Instead of investing to increase their productivity and selling their products in markets, the Kongolese moved their villages away from the market; they were trying to be as far away from the roads as possible, in order to reduce the incidence of plunder and to escape the reach of slave traders.

The poverty of the Kongo was therefore the result of extractive economic institutions that blocked all the engines of prosperity or even made them work in reverse. The Kongo's government provided very few public services to its citizens, not even basic ones, such as secure property rights or law and order. On the contrary, the government was itself the biggest threat to its subjects' property and human rights. The institution of slavery meant that the most fundamental market of all, an inclusive labor market where people can choose their occupation or jobs in ways that are so crucial for a prosperous economy, did not exist. Moreover, long-distance trade and mercantile activities were controlled by the king and were open only to those associated with him. Though the elite quickly became literate after the Portuguese introduced writing, the king made no attempt to spread literacy to the great mass of the population.

Nevertheless, though "miserable poverty" was widespread, the Kongolese extractive institutions had their own impeccable logic: they made a few people, those with political power, very rich. In the sixteenth century, the king of Kongo and the aristocracy were able to import European luxury goods and were surrounded by servants and slaves.

The roots of the economic institutions of Kongolese society flowed from the distribution of political power in society and thus from the nature of political institutions. There was nothing to stop the king from taking people's possessions or bodies, other than the threat of revolt. Though this threat was real, it was not enough to make people or their wealth secure. The political institutions of Kongo were truly absolutist, making the king and the elite subject to essentially no constraints, and it gave no say to the citizens in the way their society was organized.

Of course, it is not difficult to see that the political institutions of Kongo contrast sharply with inclusive political institutions where power is constrained and broadly distributed. The absolutist institutions of Kongo were kept in place by the army. The king had a standing army of five thousand troops in the mid-seventeenth century, with a core of five hundred musketeers—a formidable force for its time. Why the king and the aristocracy so eagerly adopted European firearms is thus easy to understand.

There was no chance of sustained economic growth under this set of economic institutions and even incentives for generating temporary growth were highly limited. Reforming economic institutions to improve individual property rights would have made the Kongolese society at large more prosperous. But it is unlikely that the elite would have benefited from this wider prosperity. First, such reforms would have made the elite economic losers, by undermining the wealth that the slave trade and slave plantations brought them. Second, such reforms would have been possible only if the political power of the king and the elite were curtailed. For instance, if the king continued to command his five hundred musketeers, who would have believed an announcement that slavery had been abolished? What would have stopped the king from changing his mind later on? The only real guarantee would have been a change in political institutions so that citizens gained some countervailing political power, giving them some say over taxation or what the musketeers did. But in this case it is dubious that sustaining the consumption and lifestyle of the king and the elite would have been high on their list of priorities. In this scenario, changes that would have created better economic institutions in society would have made the king and aristocracy political as well as economic losers.

The interaction of economic and political institutions five hundred years ago is still relevant for understanding why the modern state of Congo is still miserably poor today. The advent of European rule in this area, and deeper into the basin of the River Congo at the time of the "scramble for Africa" in the late nineteenth century, led to an insecurity of human and property rights even more egregious than that which characterized the precolonial Kongo. In addition, it reproduced the pattern of extractive institutions and political absolutism that empowered and enriched a few at the expense of the masses, though the few now were Belgian colonialists, most notably King Leopold II.

When Congo became independent in 1960, the same pattern of economic institutions, incentives, and performance reproduced itself. These Congolese extractive economic institutions were again supported by highly extractive political institutions. The situation was worsened because European colonialism created a polity, Congo, made up of many different precolonial states and societies that the national state, run from Kinshasa, had little control over. Though President Mobutu used the state to enrich himself and his cronies—for example, through the Zairianization program of 1973, which involved the mass expropriation of foreign economic interests—he presided over a noncentralized state with little authority over much of the country, and had to appeal to foreign assistance to stop the provinces of Katanga and Kasai from seceding in the

1960s. This lack of political centralization, almost to the point of total collapse of the state, is a feature that Congo shares with much of sub-Saharan Africa.

The modern Democratic Republic of Congo remains poor because its citizens still lack the economic institutions that create the basic incentives that make a society prosperous. It is not geography, culture, or the ignorance of its citizens or politicians that keep the Congo poor, but its extractive economic institutions. These are still in place after all these centuries because political power continues to be narrowly concentrated in the hands of an elite who have little incentive to enforce secure property rights for the people, to provide the basic public services that would improve the quality of life, or to encourage economic progress. Rather, their interests are to extract income and sustain their power. They have not used this power to build a centralized state, for to do so would create the same problems of opposition and political challenges that promoting economic growth would. Moreover, as in much of the rest of sub-Saharan Africa, infighting triggered by rival groups attempting to take control of extractive institutions destroyed any tendency for state centralization that might have existed.

The history of the Kingdom of Kongo, and the more recent history of the Congo, vividly illustrates how political institutions determine economic institutions and, through these, the economic incentives and the scope for economic growth. It also illustrates the symbiotic relationship between political absolutism and economic institutions that empower and enrich a few at the expense of many.

GROWTH UNDER EXTRACTIVE POLITICAL INSTITUTIONS

Congo today is an extreme example, with lawlessness and highly insecure property rights. However, in most cases such extremism would not serve the interest of the elite, since it would destroy all economic incentives and generate few resources to be extracted. The central thesis of this book is that economic growth and prosperity are associated with inclusive economic and political institutions, while extractive institutions typically lead to stagnation and poverty. But this implies neither that extractive institutions can never generate growth nor that all extractive institutions are created equal.

There are two distinct but complementary ways in which growth under extractive political institutions can emerge. First, even if economic institutions are extractive, growth is possible when elites can directly allocate resources to high-productivity activities that they themselves control. A prominent example of this type of growth under extractive institutions was the Caribbean Islands between the sixteenth and eighteenth centuries. Most people were slaves, working under gruesome conditions in plantations, living barely above subsistence level. Many died from malnutrition and exhaustion. In Barbados, Cuba, Haiti, and Jamaica in the seventeenth and eighteenth centuries, a small minority, the planter elite, controlled all political power and owned all the assets, including all the slaves. While the majority had no rights, the planter elite's property and assets were well protected. Despite the extractive economic institutions that savagely exploited the majority of the population, these islands were among the richest places in the world, because they could produce sugar and sell it in world markets. The economy of the islands stagnated only when there was a need to shift to new economic activities, which threatened both the incomes and the political power of the planter elite.

Another example is the economic growth and industrialization of the Soviet Union from the first Five-Year Plan in 1928 until the 1970s. Political and economic institutions were highly extractive, and markets were heavily constrained. Nevertheless, the Soviet Union was able to achieve rapid economic growth because it could use the power of the state to move resources from agriculture, where they were very inefficiently used, into industry.

The second type of growth under extractive political institutions arises when the institutions permit the development of somewhat, even if not completely, inclusive economic institutions. Many societies with extractive political institutions will shy away from inclusive economic institutions because of fear of creative destruction. But the degree to which the elite manage to monopolize power varies across societies. In some, the position of the elite could be sufficiently secure that they may permit some moves toward inclusive economic institutions when they are fairly certain that this will not threaten their political power. Alternatively, the historical situation could be such as to endow an extractive political regime with rather inclusive economic institutions, which they decide not to block. These provide the second way in which growth can take place under extractive political institutions.

The rapid industrialization of South Korea under General Park is an example. Park came to power via a military coup in 1961, but he did so in a society heavily supported by the United States and with an economy where economic institutions were essentially inclusive. Though Park's regime was authoritarian, it felt secure enough to promote economic growth, and in fact did so very actively—perhaps partly because the regime was not directly supported by extractive economic institutions. Differently from the Soviet Union and most other cases of growth under extractive institutions, South Korea transitioned from extractive political institutions toward inclusive political institutions in the 1980s. This successful transition was due to a confluence of factors.

By the 1970s, economic institutions in South Korea had become sufficiently inclusive that they reduced one of the strong rationales for extractive political institutions—the economic elite had little to gain from their own or the military's dominance of politics. The relative equality of income in South Korea also meant that the elite had less to fear from pluralism and democracy. The key influence of the United States, particularly given the threat from North Korea, also meant that the strong democracy movement that challenged the military dictatorship could not be repressed for long. Though General Park's assassination in 1979 was followed by another military coup, led by Chun Doo-hwan, Chun's chosen successor, Roh Tae-woo, initiated a process of political reforms that led to the consolidation of a pluralistic democracy after 1992. Of course, no transition of this sort took place in the Soviet Union. In consequence, Soviet growth ran out of steam, and the economy began to collapse in the 1980s and then totally fell apart in the 1990s.

Chinese economic growth today also has several commonalities with both the Soviet and South Korean experiences. While the early stages of Chinese growth were spearheaded by radical market reforms in the agricultural sector, reforms in the industrial sector have been more muted. Even today, the state and the Communist Party play a central role in deciding which sectors and which companies will receive additional capital and will expand—in the process, making and breaking fortunes. As in the Soviet Union in its heyday, China is growing rapidly, but this is still growth under extractive institutions, under the control of the state, with little sign of a transition to inclusive political institutions. The fact that Chinese economic institutions are still far from fully inclusive also suggests that a South Korean–style transition is less likely, though of course not impossible.

It is worth noting that political centralization is key to both ways in which growth under extractive political institutions can occur. Without some degree of political centralization, the planter elite in Barbados, Cuba, Haiti, and Jamaica would not have been able to keep law and order and defend their own assets and property. Without significant political centralization and a firm grip on political power, neither the South Korean military elites nor the Chinese Communist Party would have felt secure enough to manufacture significant economic reforms and still manage to cling to power. And without such centralization, the state in the Soviet Union or China could not have been able to coordinate economic activity to channel resources toward high productivity areas. A major dividing line between extractive political institutions is therefore their degree of political centralization. Those without it, such as many in sub-Saharan Africa, will find it difficult to achieve even limited growth.

Even though extractive institutions can generate some growth, they will usually not generate sustained economic growth, and certainly not the type of growth that is accompanied by creative destruction. When both political and economic institutions are extractive, the incentives will not be there for creative destruction and technological change. For a while the state may be able to create rapid economic growth by allocating resources and people by fiat, but this process is intrinsically limited. When the limits are hit, growth stops, as it did in the Soviet Union in the

1970s. Even when the Soviets achieved rapid economic growth, there was little technological change in most of the economy, though by pouring massive resources into the military they were able to develop military technologies and even pull ahead of the United States in the space and nuclear race for a short while. But this growth without creative destruction and without broad-based technological innovation was not sustainable and came to an abrupt end.

In addition, the arrangements that support economic growth under extractive political institutions are, by their nature, fragile—they can collapse or can be easily destroyed by the infighting that the extractive institutions themselves generate. In fact, extractive political and economic institutions create a general tendency for infighting, because they lead to the concentration of wealth and power in the hands of a narrow elite. If another group can overwhelm and outmaneuver this elite and take control of the state, they will be the ones enjoying this wealth and power. Consequently, as our discussion of the collapse of the later Roman Empire and the Maya cities will illustrate (this page and this page), fighting to control the all-powerful state is always latent, and it will periodically intensify and bring the undoing of these regimes, as it turns into civil war and sometimes into total breakdown and collapse of the state. One implication of this is that even if a society under extractive institutions initially achieves some degree of state centralization, it will not last. In fact, the infighting to take control of extractive institutions often leads to civil wars and widespread lawlessness, enshrining a persistent absence of state centralization as in many nations in sub-Saharan Africa and some in Latin America and South Asia.

Finally, when growth comes under extractive political institutions but where economic institutions have inclusive aspects, as they did in South Korea, there is always the danger that economic institutions become more extractive and growth stops. Those controlling political power will eventually find it more beneficial to use their power to limit competition, to increase their share of the pie, or even to steal and loot from others rather than support economic progress. The distribution and ability to exercise power will ultimately undermine the very foundations of economic prosperity, unless political institutions are transformed from extractive to inclusive.

11.

THE VIRTUOUS CIRCLE

THE BLACK ACT

WINDSOR CASTLE, located just west of London, is one of the great royal residencies of England. In the early eighteenth century, the castle was surrounded by a great forest, full of deer, though little of this remains today. One of the keepers of the forest in 1722, Baptist Nunn, was locked in to a violent conflict. On June 27 he recorded,

> Blacks came in the night shot at me 3 times 2 bullets into my chamber window and [I] agreed to pay them 5 guineas at Crowthorne on the 30th.

Another entry in Nunn's diary read, "A fresh surprise. One appeared disguised with a message of destruction."

Who were these mysterious "Blacks" making threats, shooting at Nunn, and demanding money? The Blacks were groups of local men who had their faces "blacked" to conceal their appearance at night. They appeared widely across southern England in this period, killing and maiming deer and other animals, burning down haystacks and barns, and destroying fences and fish ponds. On the surface it was sheer lawlessness, but it wasn't. Illegal hunting (poaching) deer in lands owned by the king or other members of the aristocracy had been going on for a long time. In the 1640s, during the Civil War, the entire population of deer at Windsor Castle was killed. After the Restoration in 1660, when Charles II came to the throne, the deer park was restocked. But the Blacks were not just poaching deer to eat; they also engaged in wanton destruction. To what end?

A crucial building block of the Glorious Revolution of 1688 was the pluralistic nature of interests represented in

Parliament. None of the merchants, industrialists, gentry, or aristocracy allied with William of Orange and then with the Hanoverian monarchs, who succeeded Queen Anne in 1714, were strong enough to impose their will unilaterally.

Attempts at restoring the Stuart monarchy continued throughout much of the eighteenth century. After James II's death in 1701, his son, James Francis Edward Stuart, the "Old Pretender," was recognized as the lawful heir to the English Crown by France, Spain, the pope, and supporters of the Stuart monarchy in England and Scotland, the so-called Jacobites. In 1708 the Old Pretender attempted to take back the throne with support of French troops, but was unsuccessful. In the ensuing decades there would be several Jacobite revolts, including major ones in 1715 and 1719. In 1745–46, the Old Pretender's son, Charles Edward Stuart, the "Young Pretender," made an attempt to take back the throne, but his forces were defeated by the British army.

The Whig political party, which as we saw (this page–this page) was founded in the 1670s to represent the new mercantile and economic interests, was the main organization behind the Glorious Revolution, and the Whigs dominated Parliament from 1714 to 1760. Once in power, they were tempted to use their newly found position to prey on the rights of others, to have their cake and eat it, too. They were no different from the Stuart kings, but their power was far from absolute. It was constrained both by competing groups in Parliament, particularly the Tory Party which had formed to oppose the Whigs, and by the very institutions that they had fought to introduce to strengthen Parliament and to prevent the emergence of a new absolutism and the return of the Stuarts. The pluralistic nature of society that emerged from the Glorious Revolution also meant that the population at large, even those without formal representation in Parliament, had been empowered, and "blacking" was precisely a response by the common people to perceptions that the Whigs were exploiting their position.

The case of William Cadogan, a successful general in the War of the Spanish Succession between 1701 and 1714 and in the suppression of the Jacobite revolts, illustrates the sort of encroachment of common people's rights by the Whigs that led to blacking. George I made Cadogan a baron in 1716 and then an earl in 1718. He was also an influential member of the Regency Council of Lords Justices, which presided over major affairs of state, and he served as the acting commander in chief. He bought a large property of about a thousand acres at Caversham, about twenty miles west of Windsor. There he built a grand house and ornate gardens and laid out a 240-acre deer park. Yet this property was consolidated by encroaching on the rights of those around the estate. People were evicted, and their traditional rights to graze animals and collect peat and firewood were abrogated. Cadogan faced the wrath of the Blacks. On January 1, 1722, and again in July, the park was raided by mounted and armed Blacks. The first attack killed sixteen deer. Earl Cadogan was not alone. The estates of many notable landowners and politicians were also raided by the Blacks.

The Whig government was not going to take this lying down. In May 1723, Parliament passed the Black Act, which created an extraordinary fifty new offenses that were punishable by hanging. The Black Act made it a crime not only to carry weapons but to have a blackened face. The law in fact was soon amended to make blacking punishable by hanging. The Whig elites went about implementing the law with gusto. Baptist Nunn set up a network of informers in Windsor Forest to discover the identity of the Blacks. Soon several were arrested. The transition from arrest to hanging ought to have been a straightforward affair. After all, the Black Act had already been enacted, the Whigs were in charge of Parliament, Parliament was in charge of the country, and the Blacks were acting directly contrary to the interests of some powerful Whigs. Even Sir Robert Walpole, secretary of state, then prime minister—and like Cadogan, another influential member of the Regency Council of the Lords Justices—was involved. He had a vested interest in Richmond Park in southwest London, which had been created out of common land by Charles I. This park also encroached upon the traditional rights of local residents to graze their animals, hunt hares and rabbits, and collect firewood. But the ending of these rights appears to have been rather laxly enforced, and grazing and hunting continued, until Walpole arranged for his son to become the park ranger. At this time, the park was closed off, a new wall was constructed, and man traps were installed. Walpole liked hunting deer, and he had a lodge built for himself at Houghton, within the park. The animosity of local Blacks was soon ignited.

On November 10, 1724, a local resident outside the park, John Huntridge, was accused of aiding deer stealers and abetting known Blacks, both crimes punishable by hanging. The prosecution of Huntridge came right from the top, initiated by the Regency Council of Lords Justices, which Walpole and Cadogan dominated. Walpole went so far as to extract evidence himself as to Huntridge's guilt from an informant, Richard Blackburn. Conviction ought to have been a foregone conclusion, but it wasn't. After a trial of eight or nine hours, the jury found Huntridge innocent, partly on procedural grounds, since there were irregularities with the way the evidence had been collected.

Not all Blacks or those who sympathized with them were as lucky as Huntridge. Though some others were also acquitted or had their convictions commuted, many were hanged or transported to the penal colony of choice at the time, North America; the law in fact stayed on the statute books until it was repealed in 1824. Yet Huntridge's victory is remarkable. The jury was made up not of Huntridge's peers, but of major landowners and gentry, who ought to have sympathized with Walpole. But this was no longer the seventeenth century, where the Court of Star Chamber would simply follow the wishes of Stuart monarchs and act as an open tool of repression against their opponents, and where kings could remove judges whose decisions they did not like. Now the Whigs also had to abide by the rule of law, the principle that laws should not be applied selectively or arbitrarily and that nobody is above the law.

THE EVENTS SURROUNDING the Black Act would show that the Glorious Revolution had created the rule of law, and that this notion was stronger in England and Britain, and the elites were far more constrained by it than they themselves imagined. Notably, the rule of law is not the same as rule by law. Even if the Whigs could pass a harsh, repressive law to quash obstacles from common people, they had to contend with additional constraints because of the rule of law. Their law violated the rights that the Glorious Revolution and the changes in political institutions that followed from it had already established for everybody by tearing down the "divine" rights of kings and the privileges of elites. The rule of law then implied that both elites and nonelites alike would resist its implementation.

The rule of law is a very strange concept when you think about it in historical perspective. Why should laws be applied equally to all? If the king and the aristocracy have political power and the rest don't, it's only natural that whatever is fair game for the king and the aristocracy should be banned and punishable for the rest. Indeed, the rule of law is not imaginable under absolutist political institutions. It is a creation of pluralist political institutions and of the broad coalitions that support such pluralism. It's only when many individuals and groups have a say in decisions, and the political power to have a seat at the table, that the idea that they should all be treated fairly starts making sense. By the early eighteenth century, Britain was becoming sufficiently pluralistic, and the Whig elites would discover that, as enshrined in the notion of the rule of law, laws and institutions would constrain them, too.

But why did the Whigs and parliamentarians abide by such restraints? Why didn't they use their control over Parliament and the state to force an uncompromising implementation of the Black Act and overturn the courts when the decisions didn't go their way? The answer reveals much about the nature of the Glorious Revolution—why it didn't just replace an old absolutism with a new version—the link between pluralism and the rule of law, and the dynamics of virtuous circles. As we saw in chapter 7, the Glorious Revolution was not the overthrow of one elite by another, but a revolution against absolutism by a broad coalition made up of the gentry, merchants, and manufacturers as well as groupings of Whigs and Tories. The emergence of pluralist political institutions was a consequence of this revolution. The rule of law also emerged as a by-product of this process. With many parties at the table sharing power, it was natural to have laws and constraints apply to all of them, lest one party start amassing too much power and ultimately undermine the very foundations of pluralism. Thus the notion that there were limits and restraints on rulers, the essence of the rule of law, was part of the logic of pluralism engendered by the broad coalition that made up the opposition to Stuart absolutism.

In this light, it should be no surprise that the principle of the rule of law, coupled with the notion that monarchs did not have divine rights, was in fact a key argument against Stuart absolutism. As the British historian E. P. Thompson put it, in the struggle against the Stuart monarchs:

> immense efforts were made … to project the image of a ruling class which was itself subject to the rule of law, and whose legitimacy rested upon the equity and universality of those legal forms. And the rulers were, in serious senses, whether willingly or unwillingly, the prisoners of their own rhetoric; they played games of power according to rules which suited them, but they could not break those rules or the whole game would be thrown away.

Throwing the game away would destabilize the system and open the way for absolutism by a subset of the broad coalition or even risk the return of the Stuarts. In Thompson's words, what inhibited Parliament from creating a new absolutism was that

> take away law, and the royal prerogative … might flood back upon their properties and lives.

Moreover,

> it was inherent in the very nature of the medium which they [those aristocrats, merchants etc. fighting the Crown] had selected for their own self-defense that it could not be reserved for the exclusive use only of their own class. The law, in its forms and traditions, entailed principles of equity

> and universality which … had to be extended to all sorts and degrees of men.

Once in place, the notion of the rule of law not only kept absolutism at bay but also created a type of virtuous circle: if the laws applied equally to everybody, then no individual or group, not even Cadogan or Walpole, could rise above the law, and common people accused of encroaching on private property still had the right to a fair trial.

WE SAW HOW INCLUSIVE economic and political institutions emerge. But why do they persist over time? The history of the Black Act and the limits to its implementation illustrate the virtuous circle, a powerful process of positive feedback that preserves these institutions in the face of attempts at undermining them and, in fact, sets in motion forces that lead to greater inclusiveness. The logic of virtuous circles stems partly from the fact that inclusive institutions are based on constraints on the exercise of power and on a pluralistic distribution of political power in society, enshrined in the rule of law. The ability of a subset to impose its will on others without any constraints, even if those others are ordinary citizens, as Huntridge was, threatens this very balance. If it were temporarily suspended in the case of the peasants protesting against elites encroaching on their communal lands, what was there to guarantee that it would not be suspended again? And the next time it was suspended, what would prevent the Crown and aristocracy from taking back what the merchants, businessmen, and the gentry had gained in the intervening half century? In fact, the next time it was suspended, perhaps the entire project of pluralism would come crumbling down, because a narrow set of interests would take control at the expense of the broad coalition. The political system would not risk this. But this made pluralism, and the rule of law that it implied, persistent features of British political institutions. And we will see that once pluralism and the rule of law were established, there would be demand for even greater pluralism and greater participation in the political process.

The virtuous circle arises not only from the inherent logic of pluralism and the rule of law, but also because inclusive political institutions tend to support inclusive economic institutions. This then leads to a more equal distribution of income, empowering a broad segment of society and making the political playing field even more level. This limits what one can achieve by usurping political power and reduces the incentives to re-create extractive political institutions. These factors were important in the emergence of truly democratic political institutions in Britain.

Pluralism also creates a more open system and allows independent media to flourish, making it easier for groups that have an interest in the continuation of inclusive institutions to become aware and organize against threats to these institutions. It is highly significant that the English state stopped censoring the media after 1688. The media played a similarly important role in empowering the population at large and in the continuation of the virtuous circle of institutional development in the United States, as we will see in this chapter.

While the virtuous circle creates a tendency for inclusive institutions to persist, it is neither inevitable nor irreversible. Both in Britain and the United States, inclusive economic and political institutions were subject to many challenges. In 1745 the Young Pretender got all the way to Derby, a mere hundred miles from London, with an army to unseat the political institutions forged during the Glorious Revolution. But he was defeated. More important than the challenges from without were potential challenges from within that might also have led to the unraveling of inclusive institutions. As we saw in the context of the Peterloo Massacre in Manchester in 1819 (this page), and as we will see in more detail next, British political elites thought of using repression to avoid having to further open the political system, but they pulled back from the brink. Similarly, inclusive economic and political institutions in the United States faced serious challenges, which could have conceivably succeeded, but didn't. And of course it was not preordained that these challenges should be defeated. It is due to not only the virtuous circle but also to the realization of the contingent path of history that British and U.S. inclusive institutions survived and became substantially stronger over time.

THE SLOW MARCH OF DEMOCRACY

The response to the Black Act showed ordinary British people that they had more rights than they previously realized. They could defend their traditional rights and economic interests in the courts and in Parliament through the use of petitions and lobbying. But this pluralism had not yet delivered effective democracy. Most adult men could not vote; neither could women; and there were many inequities in the existing democratic structures. All this was to change. The virtuous circle of inclusive institutions not only preserves what has already been achieved but also opens the door to greater inclusiveness. The odds were against the British elite of the eighteenth century maintaining their grip on political power without serious challenges. This elite had come to power by challenging the divine right of kings and opening the door to participation by the people in politics, but then they gave this right only to a small minority. It was only a matter of time until more and more of the population demanded the right to participate in the political process. And in the years leading up to 1831, they did.

The first three decades of the nineteenth century witnessed increasing social unrest in Britain, mostly in response to increasing economic inequities and demands from the disenfranchised masses for greater political representation. The Luddite Riots of 1811–1816, where workers fought against the introduction of new technologies they believed would reduce their wages, were followed by riots explicitly demanding political rights, the Spa Fields Riots of 1816 in London and the Peterloo Massacre of 1819 in Manchester. In the Swing Riots of 1830, agricultural workers protested against falling living standards as well as the introduction of new technology. Meanwhile, in Paris, the July Revolution of 1830 exploded. A consensus among elites was starting to form that the discontent was reaching the boiling point, and the only way to defuse social unrest, and turn back a revolution, was by meeting the demands of the masses and undertaking parliamentary reform.

It was no surprise then that the 1831 election was mostly about a single issue: political reform. The Whigs, almost one hundred years after Sir Robert Walpole, were much more responsive to the wishes of the common man and campaigned to extend voting rights. But this meant only a small increase in the electorate. Universal suffrage, even only for men, was not on the table. The Whigs won the election, and their leader, Earl Grey, became the prime minister. Earl Grey was no radical—far from it. He and the Whigs pushed for reform not because they thought a broader voting franchise was more just or because they wanted to share power. British democracy was not given by the elite. It was largely taken by the masses, who were empowered by the political processes that had been ongoing in England and the rest of Britain for the last several centuries. They had become emboldened by the changes in the nature of political institutions unleashed by the Glorious Revolution. Reforms were granted because the elite thought that reform was the only way to secure the continuation of their rule, albeit in a somewhat lessened form. Earl Grey, in his famous speech to Parliament in favor of political reform, said this very clearly:

> There is no-one more decided against annual Parliaments, universal suffrage and the ballot, than I am. My object is not to favour, but to put an end to such hopes and projects … The principle of my reform is, to prevent the necessity of revolution … reforming to preserve and not to overthrow.

The masses did not just want the vote for its own sake but to have a seat at the table to be able to defend their interests. This was well understood by the Chartist movement, which led the campaign for universal suffrage after 1838, taking its name from its adoption of the People's Charter, named to evoke a parallel with the Magna Carta. Chartist J. R. Stephens articulated why universal suffrage, and the vote for all citizens, was key for the masses:

> The question of universal suffrage … is a knife and fork question, a bread and cheese question … by universal suffrage I mean to say that every working man in the land has a

> right to a good coat on his back, a good hat on his head, a good roof for the shelter of his household, a good dinner upon his table.

Stephens had well understood that universal suffrage was the most durable way of empowering the British masses further and guaranteeing a coat, a hat, a roof, and a good dinner for the working man.

Ultimately, Earl Grey was successful both in ensuring the passage of the First Reform Act and in defusing the revolutionary tides without taking any major strides toward universal mass suffrage. The 1832 reforms were modest, only doubling the voting franchise from 8 percent to about 16 percent of the adult male population (from about 2 to 4 percent of all the population). They also got rid of rotten boroughs and gave independent representation to the new industrializing cities such as Manchester, Leeds, and Sheffield. But this still left many issues unresolved. Hence there were soon further demands for greater voting rights and further social unrest. In response, further reform would follow.

Why did the British elites give in to the demands? Why did Earl Grey feel that partial—indeed, very partial—reform was the only way to preserve the system? Why did they have to put up with the lesser of the two evils, reform or revolution, rather than maintaining their power without any reform? Couldn't they just have done what the Spanish conquistadors did in South America, what Austria-Hungarian and Russian monarchs would do in the next several decades when the demands for reform reached those lands, and what the British themselves did in the Caribbean and in India: use force to put down the demands? The answer to this question comes from the virtuous circle. The economic and political changes that had already taken place in Britain made using force to repress these demands both unattractive for the elite and increasingly infeasible. As E. P. Thompson wrote:

> When the struggles of 1790–1832 signalled that this equilibrium had changed, the rulers of England were faced with alarming alternatives. They could either dispense with

> the rule of law, dismantle their elaborate constitutional structures, countermand their own rhetoric and rule by force; or they could submit to their own rules and surrender their hegemony … they took halting steps in the first direction. But in the end, rather than shatter their own self-image and repudiate 150 years of constitutional legality, they surrendered to the law.

Put differently, the same forces that made the British elite not wish to tear down the edifice of the rule of law during the Black Act also made them shun repression and rule by force, which would again risk the stability of the entire system. If undermining the law in trying to implement the Black Act would have weakened the system that merchants, businessmen, and the gentry had built in the Glorious Revolution, setting up a repressive dictatorship in 1832 would have entirely undermined it. In fact, the organizers of the protests for parliamentary reform were well aware of the importance of the rule of law and its symbolism to the British political institutions during this period. They used its rhetoric to bring home this point. One of the first organizations seeking parliamentary reform was called the Hampden Club, after the member of Parliament who had first resisted Charles I over the ship money tax, a crucial event leading up to the first major uprising against Stuart absolutism, as we saw in chapter 7.

There was also dynamic positive feedback between inclusive economic and political institutions making such a course of action attractive. Inclusive economic institutions led to the development of inclusive markets, inducing a more efficient allocation of resources, greater encouragement to acquire education and skills, and further innovations in technology. All of these forces were in play in Britain by 1831. Clamping down on popular demands and undertaking a coup against inclusive political institutions would also destroy these gains, and the elites opposing greater democratization and greater inclusiveness might find themselves among those losing their fortunes from this destruction.

Another aspect of this positive feedback is that under inclusive economic and political institutions, controlling power became less central. In Austria-Hungary and in Russia, as we saw in chapter 8, the monarchs and the aristocracy had much to lose from industrialization and reform. In contrast, in Britain at the beginning of the nineteenth century, thanks to the development of inclusive economic institutions, there was much less at stake: there were no serfs, relatively little coercion in the labor market, and few monopolies protected by entry barriers. Clinging to power was thus much less valuable for the British elite.

The logic of the virtuous circle also meant that such repressive steps would be increasingly infeasible, again because of the positive feedback between inclusive economic and political institutions. Inclusive economic institutions lead to a more equitable distribution of resources than extractive institutions. As such, they empower the citizens at large and thus create a more level playing field, even when it comes to the fight for power. This makes it more difficult for a small elite to crush the masses rather than to give in to their demands, or at least to some of them. The British inclusive institutions had also already unleashed the Industrial Revolution, and Britain was highly urbanized. Using repression against an urban, concentrated, and partially organized and empowered group of people would have been much harder than repressing a peasantry or dependent serfs.

The virtuous circle thus brought the First Reform Act to Britain in 1832. But this was just the beginning. There was still a long road to travel toward real democracy, because in 1832 the elite had only offered what they thought they had to and no more. The issue of parliamentary reform was taken up by the Chartist movement, whose People's Charter of 1838 included the clauses

> A vote for every man twenty-one years of age, of sound mind, and not undergoing punishment for crime.
>
> The ballot.—To protect the elector in the exercise of his vote.
>
> No property qualification for members of Parliament—thus enabling the constituencies to return the man of their choice, be he rich or poor.
>
> Payment of members, thus enabling an honest

> tradesman, working man, or other person, to serve a constituency, when taken from his business to attend to the interests of the Country.
>
> Equal Constituencies, securing the same amount of representation for the same number of electors, instead of allowing small constituencies to swamp the votes of large ones.
>
> Annual Parliaments, thus presenting the most effectual check to bribery and intimidation, since though a constituency might be bought once in seven years (even with the ballot), no purse could buy a constituency (under a system of universal suffrage) in each ensuing twelve-month; and since members, when elected for a year only, would not be able to defy and betray their constituents as now.

By the "ballot," they meant the secret ballot and the end of open voting, which had facilitated the buying of votes and the coercion of voters.

The Chartist movement organized a series of mass demonstrations, and throughout this period Parliament continually discussed the potential for further reforms. Though the Chartists disintegrated after 1848, they were followed by the National Reform Union, founded in 1864, and the Reform League, which was founded in 1865. In July 1866, major pro-reform riots in Hyde Park brought reform right to the top of the political agenda once more. This pressure bore dividends in the form of the Second Reform Act of 1867, in which the total electorate was doubled and working-class voters became the majority in all urban constituencies. Shortly afterward the secret ballot was introduced and moves were made to eliminate corrupt electoral practices such as "treating" (essentially buying votes in exchange for which the voter received a treat, usually money, food, or alcohol). The electorate was doubled again by the Third Reform Act of 1884, when 60 percent of adult males were enfranchised. Following the First World War, the Representation of the People Act of 1918 gave the vote to all adult males over the age of twenty-one, and to women over the age of thirty who were taxpayers or married to taxpayers. Ultimately, all women also received the vote on the same terms as men in 1928.

The measures of 1918 were negotiated during the war and reflected a quid pro quo between the government and the working classes, who were needed to fight and produce munitions. The government may also have taken note of the radicalism of the Russian Revolution.

Parallel with the gradual development of more inclusive political institutions was a movement toward even more inclusive economic institutions. One major consequence of the First Reform Act was the repeal of the Corn Laws in 1846. As we saw in chapter 7, the Corn Laws banned the import of grains and cereals, keeping their prices high and ensuring lucrative profits for large landowners. The new parliamentarians from Manchester and Birmingham wanted cheap corn and low wages. They won, and the landed interests suffered a major defeat.

The changes in the electorate and other dimensions of political institutions taking place during the course of the nineteenth century were followed by further reforms. In 1871 the Liberal prime minister Gladstone opened up the civil service to public examination, making it meritocratic, and thus continuing the process of political centralization and the building of state institutions that started during the Tudor period. Liberal and Tory governments during this period introduced a considerable amount of labor market legislation. For example, the Masters and Servants Acts, which allowed employers to use the law to reduce the mobility of their workers, was repealed, changing the nature of labor relations in favor of workers. During 1906–1914, the Liberal Party, under the leadership of H. H. Asquith and David Lloyd George, began to use the state to provide far more public services, including health and unemployment insurance, government-financed pensions, minimum wages, and a commitment to redistributive taxation. As a result of these fiscal changes, taxes as a proportion of national product more than doubled in the last three decades of the nineteenth century, and then doubled again in the first three decades of the twentieth. The tax system also became more "progressive," so that wealthier people bore a heavier burden.

Meanwhile, the education system, which was previously either primarily for the elite, run by religious denominations, or required poor people to pay fees, was made more accessible to the masses; the Education Act of 1870 committed the government to the systematic provision of universal education for the first time. Education became free of charge in 1891. The school-leaving age was set at eleven in 1893. In 1899 it was increased to twelve, and special provisions for the children of needy families were introduced. As a result of these changes, the proportion of ten-year-olds enrolled in school, which stood at a disappointing 40 percent in 1870, increased to 100 percent in 1900. Finally, the Education Act of 1902 led to a large expansion in resources for schools and introduced the grammar schools, which subsequently became the foundation of secondary education in Britain.

In fact, the British example, an illustration of the virtuous circle of inclusive institutions, provides an example of a "gradual virtuous circle." The political changes were unmistakably toward more inclusive political institutions and were the result of demands from empowered masses. But they were also gradual. Every decade another step, sometimes smaller, sometimes larger, was taken toward democracy. There was conflict over each step, and the outcome of each was contingent. But the virtuous circle created forces that reduced the stakes involved in clinging to power. It also spurred the rule of law, making it harder to use force against those who were demanding what these elites had themselves demanded from Stuart monarchs. It became less likely that this conflict would turn into an all-out revolution and more likely that it would be resolved in favor of greater inclusiveness. There is great virtue in this sort of gradual change. It is less threatening to the elite than the wholesale overthrow of the system. Each step is small, and it makes sense to give in to a small demand rather than create a major showdown. This partly explains how the Corn Law was repealed without more serious conflict. By 1846 landowners could no longer control legislation in Parliament. This was an outcome of the First Reform Act. However, if in 1832 the expansion of the electorate, the reform of the rotten boroughs, and the repeal of the Corn Laws had all been on the table, landowners would have put up much more resistance. The fact that there were first limited political reforms and that repeal of the Corn Laws came on the agenda only later defused conflict.

Gradual change also prevented ventures into uncharted territories. A violent overthrow of the system means that something entirely new has to be built in place of what has been removed. This was the case with the French Revolution, when the first experiment with democracy led to the Terror and then back to a monarchy twice before finally leading to the French Third Republic in 1870. It was the case in the Russian Revolution, where the desires of many for a more equal system than that of the Russian Empire led to a one-party dictatorship that was much more violent, bloody, and vicious than what it had replaced. Gradual reform was difficult in these societies precisely because they lacked pluralism and were highly extractive. It was the pluralism emerging from the Glorious Revolution, and the rule of law that it introduced, that made gradual change feasible, and desirable, in Britain.

The conservative English commentator Edmund Burke, who steadfastly opposed the French Revolution, wrote in 1790, "It is with infinite caution that any man should venture upon pulling down an edifice, which has answered in any tolerable degree for ages the common purposes of society, or on building it up again without having models and patterns of approved utility before his eyes." Burke was wrong on the big picture. The French Revolution had replaced a rotten edifice and opened the way for inclusive institutions not only in France, but throughout much of Western Europe. But Burke's caution was not entirely off the mark. The gradual process of British political reform, which had started in 1688 and would pick up pace three decades after Burke's death, would be more effective because its gradual nature made it more powerful, harder to resist, and ultimately more durable.

BUSTING TRUSTS

Inclusive institutions in the United States had their roots in the struggles in Virginia, Maryland, and the Carolinas during the colonial period (this page–this page). These institutions were reinforced by the Constitution of the United States, with its system of constraints and its separation of powers. But the Constitution did not mark the end of the development of inclusive institutions. Just as in Britain, these were strengthened by a process of positive feedback, based on the virtuous circle.

By the middle of the nineteenth century, all white males, though not women or blacks, could vote in the United States. Economic institutions became more inclusive—for example, with the passage of the Homestead Act in 1862 (this page), which made frontier land available to potential settlers rather than allocating these lands to political elites. But just as in Britain, challenges to inclusive institutions were never entirely absent. The end of the U.S. Civil War initiated a rapid spurt of economic growth in the North. As railways, industry, and commerce expanded, a few people made vast fortunes. Emboldened by their economic success, these men and their companies became increasingly unscrupulous. They were called the Robber Barons because of their hard-nosed business practices aimed at consolidating monopolies and preventing any potential competitor from entering the market or doing business on an equal footing. One of the most notorious of these was Cornelius Vanderbilt, who famously remarked, "What do I care about the Law? Hain't I got the power?"

Another was John D. Rockefeller, who started the Standard Oil Company in 1870. He quickly eliminated rivals in Cleveland and attempted to monopolize the transportation and retailing of oil and oil products. By 1882 he had created a massive monopoly—in the language of the day, a trust. By 1890 Standard Oil controlled 88 percent of the refined oil flows in the United States, and Rockefeller became the world's first billionaire in 1916. Contemporary cartoons depict Standard Oil as an octopus wrapping itself around not just the oil industry but also Capitol Hill.

Almost as infamous was John Pierpont Morgan, the founder of the modern banking conglomerate J.P. Morgan, which later, after many mergers over decades, eventually became JPMorgan Chase. Along with Andrew Carnegie, Morgan founded the U.S. Steel Company in 1901, the first corporation with a capitalized value of more than $1 billion and by far the largest steel corporation in the world. In the 1890s, large trusts began to emerge in nearly every sector of the economy, and many of them controlled more than 70 percent of the market in their sector. These included several household names, such as Du Pont, Eastman

Kodak, and International Harvester. Historically the United States, at least the northern and midwestern United States, had relatively competitive markets and had been more egalitarian than other parts of the country, particularly the South. But during this period, competition gave way to monopoly, and wealth inequality rapidly increased.

The pluralistic U.S. political system already empowered a broad segment of society that could stand up against such encroachments. Those who were the victims of the monopolistic practices of the Robber Barons, or who objected to their unscrupulous domination of their industries, began to organize against them. They formed the Populist and then subsequently the Progressive movements.

The Populist movement emerged out of a long-running agrarian crisis, which afflicted the Midwest from the late 1860s onward. The National Grange of the Order of Patrons of Husbandry, known as the Grangers, was founded in 1867 and began to mobilize farmers against unfair and discriminatory business practices. In 1873 and 1874, the Grangers won control of eleven midwestern state legislatures, and rural discontent culminated in the formation of the People's Party in 1892, which got 8.5 percent of the popular vote in the 1892 presidential election. In the next two elections, the Populists fell in behind the two unsuccessful Democratic campaigns by William Jennings Bryan, who made many of their issues his own. Grass-roots opposition to the spread of the trusts had now organized to try to counteract the influence that Rockefeller and other Robber Barons were exerting over national politics.

These political movements slowly began to have an impact on political attitudes and then on legislation, particularly concerning the role of the state in the regulation of monopoly. The first important piece of legislation was the Interstate Commerce Act of 1887, which created the Interstate Commerce Commission and initiated the development of the federal regulation of industry. This was quickly followed by the Sherman Antitrust Act of 1890. The Sherman Act, which is still a major part of U.S. antitrust regulation, would become the basis for attacks on the Robber Barons' trusts. Major action against the trusts came after the election of presidents committed to reform and to limiting the power of the Robber Barons: Theodore Roosevelt, 1901–1909; William Taft, 1909–1913; and Woodrow Wilson, 1913–1921.

A key political force behind antitrust and the move to impose federal regulation of industry was again the farm vote. Early attempts by individual states in the 1870s to regulate railroads came from farmers' organizations. Indeed, nearly all the fifty-nine petitions that concerned trusts sent to Congress prior to the enactment of the Sherman Act came from farming states and emanated from organizations such as the Farmers' Union, Farmers' Alliance, Farmers' Mutual Benefit Association, and Patrons of Animal Husbandry. Farmers found a collective interest in opposing the monopolistic practices of industry.

From the ashes of the Populists, who seriously declined after throwing their weight behind the Democrats, came the Progressives, a heterogeneous reform movement concerned with many of the same issues. The Progressive movement initially gelled around the figure of Teddy Roosevelt, who was William McKinley's vice president and who assumed the presidency following McKinley's assassination in 1901. Prior to his rise to national office, Roosevelt had been an uncompromising governor of New York and had worked hard to eliminate political corruption and "machine politics." In his first address to Congress, Roosevelt turned his attention to the trusts. He argued that the prosperity of the United States was based on market economy and the ingenuity of businessmen, but at the same time,

> there are real and grave evils … and a … widespread conviction in the minds of the American people that the great corporations known as trusts are in certain of their features and tendencies hurtful to the general welfare. This springs from no spirit of envy or un-charitableness, nor lack of pride in the great industrial achievements that have placed this country at the head of the nations struggling for commercial supremacy. It does not rest upon a lack of intelligent appreciation

> of the necessity of meeting changing and changed conditions of trade with new methods, nor upon ignorance of the fact that combination of capital in the effort to accomplish great things is necessary when the world's progress demands that great things be done. It is based upon sincere conviction that combination and concentration should be, not prohibited, but supervised and within reasonable limits controlled; and in my judgment this conviction is right.

He continued: "It should be as much the aim of those who seek for social betterment to rid the business world of crimes of cunning as to rid the entire body politic of crimes of violence." His conclusion was that

> in the interest of the whole people, the nation should, without interfering with the power of the states in the matter itself, also assume power of supervision and regulation over all corporations doing an interstate business. This is especially true where the corporation derives a portion of its wealth from the existence of some monopolistic element or tendency in its business.

Roosevelt proposed that Congress establish a federal agency with power to investigate the affairs of the great corporations and that, if necessary, a constitutional amendment could be used to create such an agency. By 1902 Roosevelt had used the Sherman Act to break up the Northern Securities Company, affecting the interests of J.P. Morgan, and subsequent suits had been brought against Du Pont, the American Tobacco Company, and the Standard Oil Company. Roosevelt strengthened the Interstate Commerce Act with the Hepburn Act of 1906, which increased the powers of the Interstate Commerce Commission, particularly allowing it to inspect the financial accounts of railways and extending its authority into new spheres. Roosevelt's successor, William Taft, prosecuted trusts even more assiduously, the high point of this being the breakup of the Standard Oil Company in 1911. Taft also promoted other important reforms, such as the introduction of a federal income tax, which came with the ratification of the Sixteenth Amendment in 1913.

The apogee of Progressive reforms came with the election of Woodrow Wilson in 1912. Wilson noted in his 1913 book, *The New Freedom*, "If monopoly persists, monopoly will always sit at the helm of government. I do not expect to see monopoly restrain itself. If there are men in this country big enough to own the government of the United States, they are going to own it."

Wilson worked to pass the Clayton Antitrust Act in 1914, strengthening the Sherman Act, and he created the Federal Trade Commission, which enforced the Clayton Act. In addition, under the impetus of the investigation of the Pujo Committee, led by Louisiana congressman Arsene Pujo, into the "money trust," the spread of monopoly into the financial industry, Wilson moved to increase regulation of the financial sector. In 1913 he created the Federal Reserve Board, which would regulate monopolistic activities in the financial sector.

The rise of Robber Barons and their monopoly trusts in the late nineteenth and early twentieth centuries underscores that, as we already emphasized in chapter 3, the presence of markets is not by itself a guarantee of inclusive institutions. Markets can be dominated by a few firms, charging exorbitant prices and blocking the entry of more efficient rivals and new technologies. Markets, left to their own devices, can cease to be inclusive, becoming increasingly dominated by the economically and politically powerful. Inclusive economic institutions require not just markets, but inclusive markets that create a level playing field and economic opportunities for the majority of the people. Widespread monopoly, backed by the political power of the elite, contradicts this. But the reaction to the monopoly trusts also illustrates that when political institutions are inclusive, they create a countervailing force against movements away from inclusive markets. This is the virtuous circle in action. Inclusive economic institutions provide foundations upon which inclusive political institutions can flourish, while inclusive political institutions restrict deviations away from inclusive economic institutions. Trust busting in the United States, in contrast to what we have seen in Mexico (this page–this page), illustrates this facet of the virtuous circle. While there is no political body in Mexico restricting Carlos Slim's monopoly, the Sherman and Clayton Acts have been used repeatedly in the United States over the past century to restrict trusts, monopolies, and cartels, and to ensure that markets remain inclusive.

The U.S. experience in the first half of the twentieth century also emphasizes the important role of free media in empowering broad segments of society and thus in the virtuous circle. In 1906 Roosevelt coined the term *muckraker*, based on a literary character, the man with the muckrake in Bunyan's *Pilgrim's Progress*, to describe what he regarded as intrusive journalism. The term stuck and came to symbolize journalists who were intrusively, but also effectively, exposing the excesses of Robber Barons as well as corruption in local and federal politics. Perhaps the most famous muckraker was Ida Tarbell, whose 1904 book, *History of the Standard Oil Company*, played a key role in moving public opinion against Rockefeller and his business interests, culminating in the breakup of Standard Oil in 1911. Another key muckraker was lawyer and author Louis Brandeis, who would later be named Supreme Court justice by President Wilson. Brandeis outlined a series of financial scandals in his book *Other People's Money and How Bankers Use It*, and was highly influential on the Pujo Committee. The newspaper magnate William Randolph Hearst also played a salient role as muckraker. His serialization in his magazine *The Cosmopolitan* in 1906 of articles by David Graham Phillips, called "The Treason of the Senate," galvanized the campaign to introduce direct elections for the Senate, another key Progressive reform that happened with the enactment of the Seventeenth Amendment to the U.S. constitution in 1913.

The muckrakers played a major role in inducing politicians to take action against the trusts. The Robber Barons hated the muckrakers, but the political institutions of the United States made it impossible for them to stamp out and silence them. Inclusive political institutions allow a free media to flourish, and a free media, in turn, makes it more likely that threats against inclusive economic and political institutions will be widely known and resisted. In contrast, such freedom is impossible under extractive political institutions, under absolutism, or under dictatorships, which helps extractive regimes to prevent serious opposition from forming in the first place. The information that the free media provided was clearly key during the first half of the twentieth century in the United States. Without this information, the U.S. public would not have known the true extent of the power and abuses of the Robber Barons and would not have mobilized against their trusts.

PACKING THE COURT

Franklin D. Roosevelt, the Democratic Party candidate and cousin of Teddy Roosevelt, was elected president in 1932 in the midst of the Great Depression. He came to power with a popular mandate to implement an ambitious set of policies for combating the Great Depression. At the time of his inauguration in early 1933, one-quarter of the labor force was unemployed, with many thrown into poverty. Industrial production had fallen by over half since the Depression hit in 1929, and investment had collapsed. The policies Roosevelt proposed to counteract this situation were collectively known as the New Deal. Roosevelt had won a solid victory, with 57 percent of the popular vote, and the Democratic Party had majorities in both the Congress and Senate, enough to pass New Deal legislation. However, some of the legislation raised constitutional issues and ended up in the Supreme Court, where Roosevelt's electoral mandate cut much less ice.

One of the key pillars of the New Deal was the National Industrial Recovery Act. Title I focused on industrial recovery. President Roosevelt and his team believed that restraining industrial competition, giving workers greater rights to form trade unions, and regulating working standards were crucial to the recovery effort. Title II established the Public Works Administration, whose infrastructure projects include such landmarks as the Thirtieth Street railroad station in Philadelphia, the Triborough Bridge, the Grand Coulee Dam, and the Overseas Highway connecting Key West, Florida, with the mainland. President Roosevelt signed the bill into law on

June 16, 1933, and the National Industrial Recovery Act was put into operation. However, it immediately faced challenges in the courts. On May 27, 1935, the Supreme Court unanimously ruled that Title I of the act was unconstitutional. Their verdict noted solemnly, "Extraordinary conditions may call for extraordinary remedies. But … extraordinary conditions do not create or enlarge constitutional power."

Before the Court's ruling came in, Roosevelt had moved to the next step of his agenda and had signed the Social Security Act, which introduced the modern welfare state into the United States: pensions at retirement, unemployment benefits, aid to families with dependent children, and some public health care and disability benefits. He also signed the National Labor Relations Act, which further strengthened the rights of workers to organize unions, engage in collective bargaining, and conduct strikes against their employers. These measures also faced challenges in the Supreme Court. As these were making their way through the judiciary, Roosevelt was reelected in 1936 with a strong mandate, receiving 61 percent of the popular vote.

With his popularity at record highs, Roosevelt had no intention of letting the Supreme Court derail more of his policy agenda. He laid out his plans in one of his regular Fireside Chats, which was broadcast live on the radio on March 9, 1937. He started by pointing out that in his first term, much-needed policies had only cleared the Supreme Court by a whisker. He went on:

> I am reminded of that evening in March, four years ago, when I made my first radio report to you. We were then in the midst of the great banking crisis. Soon after, with the authority of the Congress, we asked the nation to turn over all of its privately held gold, dollar for dollar, to the government of the United States. Today's recovery proves how right that policy was. But when, almost two years later, it came before the Supreme Court its constitutionality was upheld only by a five-to-four vote. The change of one vote would have

> thrown all the affairs of this great nation back into hopeless chaos. In effect, four justices ruled that the right under a private contract to exact a pound of flesh was more sacred than the main objectives of the Constitution to establish an enduring nation.

Obviously, this should not be risked again. Roosevelt continued:

> Last Thursday I described the American form of government as a three-horse team provided by the Constitution to the American people so that their field might be plowed. The three horses are, of course, the three branches of government—the Congress, the executive, and the courts. Two of the horses, the Congress and the executive, are pulling in unison today; the third is not.

Roosevelt then pointed out that the U.S. Constitution had not actually endowed the Supreme Court with the right to challenge the constitutionality of legislation, but that it had assumed this role in 1803. At the time, Justice Bushrod Washington had stipulated that the Supreme Court should "presume in favor of [a law's] validity until its violation of the Constitution is proved beyond all reasonable doubt." Roosevelt then charged:

> In the last four years the sound rule of giving statutes the benefit of all reasonable doubt has been cast aside. The Court has been acting not as a judicial body, but as a policymaking body.

Roosevelt claimed that he had an electoral mandate to change this situation and that "after consideration of what reform to propose the only method which was clearly constitutional … was to infuse new blood into all our courts." He also argued that the Supreme Court judges were overworked, and the load was just too much for the older justices—who happened to be the ones striking down his legislation. He then proposed that all judges should face compulsory retirement at the age of seventy and that he should be allowed to appoint up to six new justices. This plan, which Roosevelt presented as the Judiciary Reorganization Bill, would have sufficed to remove the justices who had been appointed earlier by more conservative administrations and who had most strenuously opposed the New Deal.

Though Roosevelt skillfully tried to win popular support for the measure, opinion polls suggested that only about 40 percent of the population was in favor of the plan. Louis Brandeis was now a Supreme Court justice. Though Brandeis sympathized with much of Roosevelt's legislation, he spoke against the president's attempts to erode the power of the Supreme Court and his allegations that the justices were overworked. Roosevelt's Democratic Party had large majorities in both houses of Congress. But the House of Representatives more or less refused to deal with Roosevelt's bill. Roosevelt then tried the Senate. The bill was sent to the Senate Judiciary Committee, which then held highly contentious meetings, soliciting various opinions on the bill. They ultimately sent it back to the Senate floor with a negative report, arguing that the bill was a "needless, futile and utterly dangerous abandonment of constitutional principle … without precedent or justification." The Senate voted 70 to 20 to send it back to committee to be rewritten. All the "court packing" elements were stripped away. Roosevelt would be unable to remove the constraints placed on his power by the Supreme Court. Even though Roosevelt's powers remained constrained, there were compromises, and the Social Security and the National Labor Relations Acts were both ruled constitutional by the Court.

More important than the fate of these two acts was the general lesson from this episode. Inclusive political institutions not only check major deviations from inclusive economic institutions, but they also resist attempts to undermine their own continuation. It was in the immediate interests of the Democratic Congress and Senate to pack the court and ensure that all New Deal legislation survived. But in the same way that British political elites in the early eighteenth century understood that suspending the rule of law would endanger the gains they had wrested from the monarchy, congressmen and senators understood that if the president could undermine the independence of the judiciary, then this would undermine the balance of power in the system that protected them from the president and ensured the continuity of pluralistic political institutions.

Perhaps Roosevelt would have decided next that obtaining legislative majorities took too much compromise and time and that he would instead rule by decree, totally undermining pluralism and the U.S. political system. Congress certainly would not have approved this, but then Roosevelt could have appealed to the nation, asserting that Congress was impeding the necessary measures to fight the Depression. He could have used the police to close Congress. Sound farfetched? This is exactly what happened in Peru and Venezuela in the 1990s. Presidents Fujimori and Chávez appealed to their popular mandate to close uncooperative congresses and subsequently rewrote their constitutions to massively strengthen the powers of the president. The fear of this slippery slope by those sharing power under pluralistic political institutions is exactly what stopped Walpole from fixing British courts in the 1720s, and it is what stopped the U.S. Congress from backing Roosevelt's court-packing plan. Roosevelt had encountered the power of virtuous circles.

But this logic does not always play out, particularly in societies that may have some inclusive features but that are broadly extractive. We have already seen these dynamics in Rome and Venice. Another illustration comes from comparing Roosevelt's failed attempt to pack the Court with similar efforts in Argentina, where crucially the same struggles took place in the context of predominantly extractive economic and political institutions.

The 1853 constitution of Argentina created a Supreme Court with duties similar to those of the U.S. Supreme Court. An 1887 decision allowed the Argentine court to assume the same role as that of the U.S. Supreme Court in deciding whether specific laws were constitutional. In theory, the Supreme Court could have developed as one of the important elements of inclusive political institutions in Argentina, but the rest of the political and economic system remained highly extractive, and there was neither empowerment of broad segments of society nor pluralism in Argentina. As in the United States, the constitutional role of the Supreme Court would also be challenged in Argentina. In 1946 Juan Domingo Perón was democratically elected president of Argentina. Perón was a former colonel and had first come to national prominence after a military coup in 1943, which had appointed him minister of labor. In this post, he built a political coalition with trade unions and the labor movement, which would be crucial for his presidential bid.

Shortly after Perón's victory, his supporters in the Chamber of Deputies proposed the impeachment of four of the five members of the Court. The charges leveled against the Court were several. One involved unconstitutionally accepting the legality of two military regimes in 1930 and 1943—rather ironic, since Perón had played a key role in the latter coup. The other focused on legislation that the court had struck down, just as its U.S. counterpart had done. In particular, just prior to Perón's election as president, the Court had issued a decision ruling that Perón's new national labor relations board was unconstitutional. Just as Roosevelt heavily criticized the Supreme Court in his 1936 reelection campaign, Perón did the same in his 1946 campaign. Nine months after initiating the impeachment process, the Chamber of Deputies impeached three of the judges, the fourth having already resigned. The Senate approved the motion. Perón then appointed four new justices. The undermining of the Court clearly had the effect of freeing Perón from political constraints. He could now exercise unchecked power, in much the same way the military regimes in Argentina did before and after his presidency. His newly appointed judges, for example, ruled as constitutional the conviction of Ricardo Balbín, the leader of the main opposition party to Perón, the Radical Party, for disrespecting Perón. Perón could effectively rule as a dictator.

Since Perón successfully packed the Court, it has become the norm in Argentina for any new president to handpick his own Supreme Court justices. So a political institution that might have exercised some constraints on the power of the executive is gone. Perón's regime was removed from power by another coup in 1955, and was followed by a long sequence of transitions between military and civilian rule. Both new military and civilian regimes picked their own justices. But picking Supreme Court justices in Argentina was not an activity confined to transitions between military and civilian rule. In 1990 Argentina finally experienced a transition between democratically elected governments—one democratic government followed by another. Yet, by this time democratic governments did not behave much differently from military ones when it came to the Supreme Court. The incoming president was Carlos Saúl Menem of the Perónist Party. The sitting Supreme Court had been appointed after the transition to democracy in 1983 by the Radical Party president Raúl Alfonsín. Since this was a democratic transition, there should have been no reason for Menem to appoint his own court. But in the run-up to the election, Menem had already shown his colors. He continually, though not successfully, tried to encourage (or even intimidate) members of the court to resign. He famously offered Justice Carlos Fayt an ambassadorship. But he was rebuked, and Fayt responded by sending him a copy of his book *Law and Ethics*, with the note "Beware I wrote this" inscribed. Undeterred, within three months of taking office, Menem sent a law to the Chamber of Deputies proposing to expand the Court from five to nine members. One argument was the same Roosevelt used in 1937: the court was overworked. The law quickly passed the Senate and Chamber, and this allowed Menem to name four new judges. He had his majority.

Menem's victory against the Supreme Court set in motion the type of slippery-slope dynamics we mentioned earlier. His next step was to rewrite the constitution to remove the term limit so he could run for president again. After being reelected, Menem moved to rewrite the constitution again, but was stopped not by Argentina political institutions but by factions within his own Perónist Party, who fought back against his personal domination.

Since independence, Argentina has suffered from most of the institutional problems that have plagued Latin America. It has been trapped in a vicious, not a virtuous, circle. As a consequence, positive developments, such as first steps toward the creation of an independent Supreme Court, never gained a foothold. With pluralism, no group wants or dares to overthrow the power of another, for fear that its own power will be subsequently challenged. At the same time, the broad distribution of power makes such an overthrow difficult. A Supreme Court can have power if it receives significant support from broad segments of society willing to push back attempts to vitiate the Court's independence. That has been the case in the United States, but not Argentina. Legislators there were happy to undermine the Court even if they anticipated that this could jeopardize their own position. One reason is that with extractive institutions there is much to gain from overthrowing the Supreme Court, and the potential benefits are worth the risks.

POSITIVE FEEDBACK AND VIRTUOUS CIRCLES

Inclusive economic and political institutions do not emerge by themselves. They are often the outcome of significant conflict between elites resisting economic growth and political change and those wishing to limit the economic and political power of existing elites. Inclusive institutions emerge during critical junctures, such as during the Glorious Revolution in England or the foundation of the Jamestown colony in North America, when a series of factors weaken the hold of the elites in power, make their opponents stronger, and create incentives for the formation of a pluralistic society. The outcome of political conflict is never certain, and even if in hindsight we see many historical events as inevitable, the path of history is contingent. Nevertheless, once in place, inclusive economic and political institutions tend to create a virtuous circle, a process of positive feedback, making it more likely that these institutions will persist and even expand.

The virtuous circle works through several mechanisms. First, the logic of pluralistic political institutions makes usurpation of power by a dictator, a faction within the government, or even a well-meaning president much more difficult, as Franklin Roosevelt discovered when he tried to remove the checks on his power imposed by the Supreme Court, and as Sir Robert Walpole discovered when he attempted to summarily implement the Black Act. In both cases, concentrating power further in the hands of an individual or a narrow group would have started undermining the foundations of pluralistic political institutions, and the true measure of pluralism is precisely its ability to resist such attempts. Pluralism also enshrines the notion of the rule of law, the principle that laws should be applied equally to everybody—something that is naturally impossible under an absolutist monarchy. But the rule of law, in turn, implies that laws cannot simply be used by one group to encroach upon the rights of another. What's more, the principle of the rule of law opens the door for greater participation in the political process and greater inclusivity, as it powerfully introduces the idea that people should be equal not only before the law but also in the political system. This was one of the principles that made it difficult for the British political system to resist the forceful calls for greater democracy throughout the nineteenth century, opening the way to the gradual extension of the franchise to all adults.

Second, as we have seen several times before, inclusive political institutions support and are supported by inclusive economic institutions. This creates another mechanism of the virtuous circle. Inclusive economic institutions remove the most egregious extractive economic relations, such as slavery and serfdom, reduce the importance of monopolies, and create a dynamic economy, all of which reduces the economic benefits that one can secure, at least in the short run, by usurping political power. Because economic institutions had already become sufficiently inclusive in Britain by the eighteenth century, the elite had less to gain by clinging to power and, in fact, much to lose by using widespread repression against those demanding greater democracy. This facet of the virtuous circle made the gradual march of democracy in nineteenth-century Britain both less threatening to the elite and more likely to succeed. This contrasts with the situation in absolutist regimes such as the Austro-Hungarian or Russian empires, where economic institutions were still highly extractive and, in consequence, where calls for greater political inclusion later in the nineteenth century would be met by repression because the elite had too much to lose from sharing power.

Finally, inclusive political institutions allow a free media to flourish, and a free media often provides information about and mobilizes opposition to threats against inclusive institutions, as it did during the last quarter of the nineteenth century and first quarter of the twentieth century, when the increasing economic domination of the Robber Barons was threatening the essence of inclusive economic institutions in the United States.

Though the outcome of the ever-present conflicts continues to be contingent, through these mechanisms the virtuous circle creates a powerful tendency for inclusive institutions to persist, to resist challenges, and to expand as they did in both Britain and the United States. Unfortunately, as we will see in the next chapter, extractive institutions create equally strong forces toward their persistence—the process of the vicious circle.

12.

THE VICIOUS CIRCLE

You Can't Take the Train to Bo Anymore

ALL OF THE WEST AFRICAN nation of Sierra Leone became a British colony in 1896. The capital city, Freetown, had originally been founded in the late eighteenth century as a home for repatriated and freed slaves. But when Freetown became a British colony, the interior of Sierra Leone was still made up of many small African kingdoms. Gradually, in the second half of the nineteenth century, the British extended their rule into the interior through a long series of treaties with African rulers. On August 31, 1896, the British government declared the colony a protectorate on the basis of these treaties. The British identified important rulers and gave them a new title, paramount chief. In eastern Sierra Leone, for example, in the modern diamond-mining district of Kono, they encountered Suluku, a powerful warrior king. King Suluku was made Paramount Chief Suluku, and the chieftaincy of Sandor was created as an administrative unit in the protectorate.

Though kings such as Suluku had signed treaties with a British administrator, they had not understood that these treaties would be interpreted as carte blanche to set up a colony. When the British tried to levy a hut tax—a tax of five shillings to be raised from every house—in January 1898, the chiefs rose up in a civil war that became known as the Hut Tax Rebellion. It started in the north, but was strongest and lasted longer in the south, particularly in Mendeland, dominated by the Mende ethnic group. The Hut Tax Rebellion was soon defeated, but it warned the British about the challenges of controlling the Sierra Leonean hinterland. The British had already started to build a railway from Freetown into the interior. Work began in March 1896, and the line reached Songo Town in December 1898, in the midst of the Hut Tax Rebellion. British parliamentary papers from 1904 recorded that:

> In the case of the Sierra Leone Railways the Native Insurrection that broke out in February 1898 had the effect of completely stopping the works and disorganizing the staff for some time. The rebels descended upon the railway, with the result that the entire staff had to be withdrawn to Freetown … Rotifunk, now situated upon the railways at 55 miles from Freetown, was at that time completely in the hands of the rebels.

In fact, Rotifunk was not on the planned railway line in 1894. The route was changed after the start of the rebellion, so that instead of going to the northeast, it went south, via Rotifunk and on to Bo, into Mendeland. The British wanted quick access to Mendeland, the heart of the rebellion, and to other potentially disruptive parts of the hinterland if other rebellions were to flare up.

When Sierra Leone became independent in 1961, the British handed power to Sir Milton Margai and his Sierra Leone People's Party (SLPP), which attracted support primarily in the south, particularly Mendeland, and the east. Sir Milton was followed as prime minister by his brother, Sir Albert Margai, in 1964. In 1967 the SLPP narrowly lost a hotly contested election to the opposition, the All People's Congress Party (APC), led by Siaka Stevens. Stevens was a Limba, from the north, and the APC got most of their support from northern ethnic groups, the Limba, the Temne, and the Loko.

Though the railway to the south was initially designed by the British to rule Sierra Leone, by 1967 its role was economic, transporting most of the country's exports: coffee, cocoa, and diamonds. The farmers who grew coffee and cocoa were Mende, and the railway was Mendeland's window to the world. Mendeland had voted hugely for Albert Margai in the 1967 election. Stevens was much more interested in holding on to power than promoting Mendeland's exports. His reasoning was simple: whatever was good for the Mende was good for the SLPP, and bad for Stevens. So he pulled up the railway line to

Mendeland. He then went ahead and sold off the track and rolling stock to make the change as irreversible as possible. Now, as you drive out of Freetown to the east, you pass the dilapidated railway stations of Hastings and Waterloo. There are no more trains to Bo. Of course, Stevens's drastic action fatally damaged some of the most vibrant sectors of Sierra Leone's economy. But like many of Africa's postindependence leaders, when the choice was between consolidating power and encouraging economic growth, Stevens chose consolidating his power, and he never looked back. Today you can't take the train to Bo anymore, because like Tsar Nicholas I, who feared that the railways would bring revolution to Russia, Stevens believed the railways would strengthen his opponents. Like so many other rulers in control of extractive institutions, he was afraid of challenges to his political power and was willing to sacrifice economic growth to thwart those challenges.

Stevens's strategy at first glance contrasts with that of the British. But in fact, there was a significant amount of continuity between British rule and Stevens's regime that illustrates the logic of vicious circles. Stevens ruled Sierra Leone by extracting resources from its people using similar methods. He was still in power in 1985 not because he had been popularly reelected, but because after 1967 he set up a violent dictatorship, killing and harassing his political opponents, particularly the members of the SLPP. He made himself president in 1971, and after 1978, Sierra Leone had only one political party, Stevens's APC. Stevens thus successfully consolidated his power, even if the cost was impoverishing much of the hinterland.

During the colonial period, the British used a system of indirect rule to govern Sierra Leone, as they did with most of their African colonies. At the base of this system were the paramount chiefs, who collected taxes, distributed justice, and kept order. The British dealt with the cocoa and coffee farmers not by isolating them, but by forcing them to sell all their produce to a marketing board developed by the colonial office purportedly to help the farmers. Prices for agricultural commodities fluctuated wildly over time. Cocoa prices might be high one year but low the next. The incomes of farmers fluctuated in tandem. The justification for marketing boards was that they, not the farmers, would absorb the price fluctuations. When world prices were high, the board would pay the farmers in Sierra Leone less than the world price, but when world prices were low, they would do the opposite. It seemed a good idea in principle. The reality was very different, however. The Sierra Leone Produce Marketing Board was set up in 1949. Of course the board needed a source of revenues to function. The natural way to attain these was by paying farmers just a little less than they should have received either in good or bad years. These funds could then be used for overhead expenditures and administration. Soon the little less became a lot less. The colonial state was using the marketing board as a way of heavily taxing farmers.

Many expected the worst practices of colonial rule in sub-Saharan Africa to stop after independence, and the use of marketing boards to excessively tax farmers to come to an end. But neither happened. In fact, the extraction of farmers using marketing boards got much worse. By the mid-1960s, the farmers of palm kernels were getting 56 percent of the world price from the marketing board; cocoa farmers, 48 percent; and coffee farmers, 49 percent. By the time Stevens left office in 1985, resigning to allow his handpicked successor, Joseph Momoh, to become president, these numbers were 37, 19, and 27 percent, respectively. As pitiful as this might sound, it was better than what the farmers were getting during Stevens's reign, which had often been as low as 10 percent—that is, 90 percent of the income of the farmers was extracted by Stevens's government, and not to provide public services, such as roads or education, but to enrich himself and his cronies and to buy political support.

As part of their indirect rule, the British had also stipulated that the office of the paramount chief would be held for life. To be eligible to be a chief, one had to be a member of a recognized "ruling house." The identity of the ruling houses in a chieftaincy developed over time, but it was essentially based on the lineage of the kings in a particular area and of the elite families who signed treaties with the British in the late nineteenth century. Chiefs were elected, but not democratically. A body called the Tribal Authority, whose members were lesser village chiefs or were appointed by paramount chiefs, village chiefs, or the British authorities, decided who would become the paramount chief. One might have imagined that this colonial institution would also have been abolished or at least reformed after independence. But just like the marketing board, it was not, and continued unchanged. Today paramount chiefs are still in charge of collecting taxes. It is no longer a hut tax, but its close descendant, a poll tax. In 2005 the Tribal Authority in Sandor elected a new paramount chief. Only candidates from the Fasuluku ruling house, which is the only ruling house, could stand. The victor was Sheku Fasuluku, King Suluku's great-great-grandson.

The behavior of the marketing boards and the traditional systems of land ownership go a long way to explain why agricultural productivity is so low in Sierra Leone and much of sub-Saharan Africa. The political scientist Robert Bates set out in the 1980s to understand why agriculture was so unproductive in Africa even though according to textbook economics this ought to have been the most dynamic economic sector. He realized that this had nothing to do with geography or the sorts of factors discussed in chapter 2 that have been claimed to make agricultural productivity intrinsically low. Rather, it was simply because the pricing policies of the marketing boards removed any incentives for the farmers to invest, use fertilizers, or preserve the soil.

The reason that the policies of the marketing boards were so unfavorable to rural interests was that these interests had no political power. These pricing policies interacted with other fundamental factors making tenure insecure, further undermining investment incentives. In Sierra Leone, paramount chiefs not only provide law and order and judicial services, and raise taxes, but they are also the "custodians of the land." Though families, clans, and dynasties have user rights and traditional rights to land; at the end of the day chiefs have the last say on who farms where. Your property rights to land are only secure if you are connected to the chief, perhaps from the same ruling family. Land cannot be bought or sold or used as collateral for a loan, and if you are born outside a chieftaincy, you cannot plant any perennial crop such as coffee, cocoa, or palm for fear that this will allow you to establish "de facto" property rights.

The contrast between the extractive institutions developed by the British in Sierra Leone and the inclusive institutions that developed in other colonies, such as Australia, is illustrated by the way mineral resources were managed. Diamonds were discovered in Kono in eastern Sierra Leone in January 1930. The diamonds were alluvial, that is, not in deep mines. So the primary method of mining them was by panning in rivers. Some social scientists call these "democratic diamonds," because they allow many people to become involved in mining, creating a potentially inclusive opportunity. Not so in Sierra Leone. Happily ignoring the intrinsically democratic nature of panning for diamonds, the British government set up a monopoly for the entire protectorate, called it the Sierra Leone Selection Trust, and granted it to De Beers, the giant South African diamond mining company. In 1936 De Beers was also given the right to create the Diamond Protection Force, a private army that would become larger than that of the colonial government in Sierra Leone. Even so, the widespread availability of the alluvial diamonds made the situation difficult to police. By the 1950s, the Diamond Protection Force was overwhelmed by thousands of illegal diamond miners, a massive source of conflict and chaos. In 1955 the British government opened up some of the diamond fields to licensed diggers outside the Sierra Leone Selection Trust, though the company still kept the richest areas in Yengema and Koidu and Tongo Fields. Things only got worse after independence. In 1970 Siaka Stevens effectively nationalized the Sierra Leone Selection Trust, creating the National Diamond Mining Company (Sierra Leone) Limited, in which the government, effectively meaning Stevens, had a 51 percent stake. This was the opening phase of Stevens's plan to take over diamond mining in the country.

In nineteenth-century Australia it was gold, discovered in 1851 in New South Wales and the newly created state of Victoria, not diamonds, that attracted everyone's attention. Like diamonds in Sierra Leone, the gold was alluvial, and a decision had to be made about how to exploit it. Some, such as James Macarthur, son of John Macarthur, the prominent leader of the Squatters we discussed earlier

(this page–this page), proposed that fences be placed around the mining areas and the monopoly rights auctioned off. They wanted an Australian version of the Sierra Leone Selection Trust. Yet many in Australia wanted free access to the gold mining areas. The inclusive model won, and instead of setting up a monopoly, Australian authorities allowed anyone who paid an annual mining license fee to search and dig for gold. Soon the diggers, as these adventurers came to be known, were a powerful force in Australian politics, particularly in Victoria. They played an important role in pushing forward the agenda of universal suffrage and the secret ballot.

We have already seen two pernicious effects of European expansion and colonial rule in Africa: the introduction of the transatlantic slave trade, which encouraged the development of African political and economic institutions in an extractive direction, and the use of colonial legislation and institutions to eliminate the development of African commercial agriculture that might have competed with Europeans. Slavery was certainly a force in Sierra Leone. At the time of colonization there was no strong centralized state in the interior, just many small, mutually antagonistic kingdoms continually raiding one another and capturing one another's men and women. Slavery was endemic, with possibly 50 percent of the population working as slaves. The disease environment meant that large-scale white settlement was not possible in Sierra Leone, as it was in South Africa. Hence there were no whites competing with the Africans. Moreover, the lack of a mining economy on the scale of Johannesburg meant that, in addition to the lack of demand for African labor from white farms, there was no incentive to create the extractive labor market institutions so characteristic of Apartheid South Africa.

But other mechanisms were also in play. Sierra Leone's cocoa and coffee farmers did not compete with whites, though their incomes were still expropriated via a government monopoly, the marketing board. Sierra Leone also suffered from indirect rule. In many parts of Africa where the British authorities wished to use indirect rule, they found peoples who did not have a system of centralized authority who could be taken over. For example, in eastern Nigeria the Igbo peoples had no chiefs when the British encountered them in the nineteenth century. The British then created chiefs, the warrant chiefs. In Sierra Leone, the British would base indirect rule on existing indigenous institutions and systems of authority.

Nevertheless, regardless of the historical basis for the individuals recognized as paramount chiefs in 1896, indirect rule, and the powers that it invested in paramount chiefs, completely changed the existing politics of Sierra Leone. For one, it introduced a system of social stratification—the ruling houses—where none had existed previously. A hereditary aristocracy replaced a situation that had been much more fluid and where chiefs had required popular support. Instead what emerged was a rigid system with chiefs holding office for life, beholden to their patrons in Freetown or Britain, and far less accountable to the people they ruled. The British were happy to subvert the institutions in other ways, too, for example, by replacing legitimate chiefs with people who were more cooperative. Indeed, the Margai family, which supplied the first two prime ministers of independent Sierra Leone, came to power in the Lower Banta chieftaincy by siding with the British in the Hut Tax Rebellion against the reigning chief, Nyama. Nyama was deposed, and the Margais became chiefs and held the position until 2010.

What is remarkable is the extent of continuity between colonial and independent Sierra Leone. The British created the marketing boards and used them to tax farmers. Postcolonial governments did the same extracting at even higher rates. The British created the system of indirect rule through paramount chiefs. Governments that followed independence didn't reject this colonial institution; rather, they used it to govern the countryside as well. The British set up a diamond monopoly and tried to keep out African miners. Postindependence governments did the same. It is true that the British thought that building railways was a good way to rule Mendeland, while Siaka Stevens thought the opposite. The British could trust their army and knew it could be sent to Mendeland if a rebellion arose. Stevens, on the other hand, could not do so. As in many other African nations, a strong army would have become a threat to Stevens's rule. It was for this reason that he emasculated the army, cutting it down and privatizing violence through specially created paramilitary units loyal only to him, and in the process, he accelerated the decline of the little state authority that existed in Sierra Leone. Instead of the army, first came the Internal Security Unit, the ISU, which Sierra Leone's long-suffering people knew as "I Shoot U." Then came the Special Security Division, the SSD, which the people knew as "Siaka Stevens's Dogs." In the end, the absence of an army supporting the regime would also be its undoing. It was a group of only thirty soldiers, led by Captain Valentine Strasser, that pitched the APC regime from power on April 29, 1992.

Sierra Leone's development, or lack thereof, could be best understood as the outcome of the vicious circle. British colonial authorities built extractive institutions in the first place, and the postindependence African politicians were only too happy to take up the baton for themselves. The pattern was eerily similar all over sub-Saharan Africa. There were similar hopes for postindependence Ghana, Kenya, Zambia, and many other African countries. Yet in all these cases, extractive institutions were re-created in a pattern predicted by the vicious circle—only they became more vicious as time went by. In all these countries, for example, the British creation of marketing boards and indirect rule were sustained.

There are natural reasons for this vicious circle. Extractive political institutions lead to extractive economic institutions, which enrich a few at the expense of many. Those who benefit from extractive institutions thus have the resources to build their (private) armies and mercenaries, to buy their judges, and to rig their elections in order to remain in power. They also have every interest in defending the system. Therefore, extractive economic institutions create the platform for extractive political institutions to persist. Power is valuable in regimes with extractive political institutions, because power is unchecked and brings economic riches.

Extractive political institutions also provide no checks against abuses of power. Whether power corrupts is debatable, but Lord Acton was certainly right when he argued that absolute power corrupts absolutely. We saw in the previous chapter that even when Franklin Roosevelt wished to use his presidential powers in a way that he thought would be beneficial for the society, unencumbered by constraints imposed by the Supreme Court, the inclusive U.S. political institutions prevented him from setting aside the constraints on his power. Under extractive political institutions, there is little check against the exercise of power, however distorted and sociopathic it may become. In 1980 Sam Bangura, then the governor of the central bank in Sierra Leone, criticized Siaka Stevens's policies for being profligate. He was soon murdered and thrown from the top floor of the central bank building onto the aptly named Siaka Stevens Street. Extractive political institutions thus also tend to create a vicious circle because they provide no line of defense against those who want to further usurp and misuse the powers of the state.

Yet another mechanism for the vicious circle is that extractive institutions, by creating unconstrained power and great income inequality, increase the potential stakes of the political game. Because whoever controls the state becomes the beneficiary of this excessive power and the wealth that it generates, extractive institutions create incentives for infighting in order to control power and its benefits, a dynamic that we saw played out in Maya city-states and in Ancient Rome. In this light, it is no surprise that the extractive institutions that many African countries inherited from the colonial powers sowed the seeds of power struggles and civil wars. These struggles would be very different conflicts from the English Civil War and the Glorious Revolution. They would not be fought to change political institutions, introduce constraints on the exercise of power, or create pluralism, but to capture power and enrich one group at the expense of the rest. In Angola, Burundi, Chad, Côte d'Ivoire, the Democratic Republic of Congo, Ethiopia, Liberia, Mozambique, Nigeria, Republic of Congo Brazzaville, Rwanda, Somalia, Sudan, and Uganda, and of course in Sierra Leone, as we will see in more detail in the next chapter, these conflicts would turn into bloody civil wars and would create economic ruin and unparalleled human suffering—as well as cause state failure.

FROM *ENCOMIENDA* TO LAND GRAB

On January 14, 1993, Ramiro De León Carpio was sworn in as the president of Guatemala. He named Richard Aitkenhead Castillo as his minister of finance, and Ricardo Castillo Sinibaldi as his minister of development. These three men all had something in common: all were direct descendants of Spanish conquistadors who had come to Guatemala in the early sixteenth century. De León's illustrious ancestor was Juan De León Cardona, while the Castillos were related to Bernal Díaz del Castillo, a man who wrote one of the most famous eyewitness accounts of the conquest of Mexico. In reward for his service to Hernán Cortés, Díaz del Castillo was appointed governor of Santiago de los Caballeros, which is today the city of Antigua in Guatemala. Both Castillo and De León founded dynasties along with other conquistadors, such as Pedro de Alvarado. The Guatemalan sociologist Marta Casaús Arzú identified a core group of twenty-two families in Guatemala that had ties through marriage to another twenty-six families just outside the core. Her genealogical and political study suggested that these families have controlled economic and political power in Guatemala since 1531. An even broader definition of which families were part of this elite suggested that they accounted for just over 1 percent of the population in the 1990s.

In Sierra Leone and in much of sub-Saharan Africa, the vicious circle took the form of the extractive institutions set up by colonial powers being taken over by postindependence leaders. In Guatemala, as in much of Central America, we see a simpler, more naked form of the vicious circle: those who have economic and political power structure institutions to ensure the continuity of their power, and succeed in doing so. This type of vicious circle leads to the persistence of extractive institutions and the persistence of the same elites in power together with the persistence of underdevelopment.

At the time of the conquest, Guatemala was densely settled, probably with a population of around two million Mayas. Disease and exploitation took a heavy toll as everywhere else in the Americas. It was not until the 1920s that its total population returned to this level. As elsewhere in the Spanish Empire, the indigenous people were allocated to conquistadors in grants of *encomienda*. As we saw in the context of the colonization of Mexico and Peru, the *encomienda* was a system of forced labor, which subsequently gave way to other similar coercive institutions, particularly to the *repartimiento*, also called the *mandamiento* in Guatemala. The elite, made up of the descendants of the conquistadors and some indigenous elements, not only benefited from the various forced labor systems but also controlled and monopolized trade through a merchant guild called the Consulado de Comercio. Most of the population in Guatemala was high in the mountains and far from the coast. The high transportation costs reduced the extent of the export economy, and initially land was not very valuable. Much of it was still in the hands of indigenous peoples, who had large communal landholdings called *ejidos*. The remainder was largely unoccupied and notionally owned by the government. There was more money in controlling and taxing trade, such as it was, than in controlling the land.

Just as in Mexico, the Guatemalan elite viewed the Cadiz Constitution (this page–this page) with hostility, which encouraged them to declare independence just as the Mexican elites did. Following a brief union with Mexico and the Central American Federation, the colonial elite ruled Guatemala under the dictatorship of Rafael Carrera from 1839 to 1871. During this period the descendants of the conquistadors and the indigenous elite maintained the extractive economic institutions of the colonial era largely unchanged. Even the organization of the Consulado did not alter with independence. Though this was a royal institution, it happily continued under a republican government.

Independence then was simply a coup by the preexisting local elite, just as in Mexico; they carried on as usual with the extractive economic institutions from which they had benefited so much. Ironically enough, during this period the Consulado remained in charge of the economic development of the country. But as had been the case pre-independence, the Consulado had its own interests at heart, not those of the country. Part of its responsibility was for the development of infrastructure, such as ports and roads, but as in Austria-Hungary, Russia, and Sierra Leone, this often threatened creative destruction and could have destabilized the system. Therefore, the development of infrastructure, rather than being implemented, was often resisted. For example, the development of a port on the Suchitepéquez coast, bordering the Pacific Ocean, was one of the proposed projects. At the time the only proper ports were on the Caribbean coast, and these were controlled by the Consulado. The Consulado did nothing on the Pacific side because a port in that region would have provided a much easier outlet for goods from the highland towns of Mazatenango and Quezaltenango, and access to a different market for these goods would have undermined the Consulado's monopoly on foreign trade. The same logic applied to roads, where, again, the Consulado had the responsibility for the entire country. Predictably it also refused to build roads that would have strengthened competing groups or would have potentially undone its monopoly. Pressure to do so again came from western Guatemala and Quezaltenango, in the Los Altos region. But if the road between Los Altos and the Suchitepéquez coast had been improved, this could have created a merchant class, which would have been a competitor to the Consulado merchants in the capital. The road did not get improved.

As a result of this elite dominance, Guatemala was caught in a time warp in the middle of the nineteenth century, as the rest of the world was changing rapidly. But these changes would ultimately affect Guatemala. Transportation costs were falling due to technological innovations such as the steam train, the railways, and new, much faster types of ships. Moreover, the rising incomes of people in Western Europe and North America were creating a mass demand for many products that a country such as Guatemala could potentially produce.

Early in the century, some indigo and then cochineal, both natural dyes, had been produced for export, but the more profitable opportunity would become coffee production. Guatemala had a lot of land suitable for coffee, and cultivation began to spread—without any assistance from the Consulado. As the world price of coffee rose and international trade expanded, there were huge profits to be made, and the Guatemalan elite became interested in coffee. In 1871 the long-lasting regime of the dictator Carrera was finally overthrown by a group of people calling themselves Liberals, after the worldwide movement of that name. What liberalism means has changed over time. But in the nineteenth century in the United States and Europe, it was similar to what is today called libertarianism, and it stood for freedom of individuals, limited government, and free trade. Things worked a little differently in Guatemala. Led initially by Miguel García Granados, and after 1873 by Justo Rufino Barrios, the Guatemalan Liberals were, for the most part, not new men with liberal ideals. By and large, the same families remained in charge. They maintained extractive political institutions and implemented a huge reorganization of the economy to exploit coffee. They did abolish the Consulado in 1871, but economic circumstances had changed. The focus of extractive economic institutions would now be the production and export of coffee.

Coffee production needed land and labor. To create land for coffee farms, the Liberals pushed through land privatization, in fact really a land grab in which they would be able to capture land previously held communally or by the government. Though their attempt was bitterly contested, given the highly extractive political institutions and the concentration of political power in Guatemala, the elite were ultimately victorious. Between 1871 and 1883 nearly one million acres of land, mostly indigenous communal land and frontier lands, passed into the hands of the elite, and it was only then that coffee developed rapidly. The aim was the formation of large estates. The privatized lands were auctioned off typically to members of the traditional elite or those connected with them. The coercive power of the Liberal state was then used to help large landowners gain access to labor by adapting and intensifying various systems of forced labor. In November 1876, President Barrios wrote to all the governors of Guatemala noting that

> because the country has extensive areas of land that it needs to exploit by cultivation using the multitude of workers who today remain outside the movement of development of the nation's productive elements, you are to give all help to export

agriculture:

> 1. From the Indian towns of your jurisdiction provide to the owners of fincas [farms] of that department who ask for labor the number of workers they need, be it fifty or one hundred.

The *repartimiento*, the forced labor draft, had never been abolished after independence, but now it was increased in scope and duration. It was institutionalized in 1877 by Decree 177, which specified that employers could request and receive from the government up to sixty workers for fifteen days of work if the property was in the same department, and for thirty days if it was outside it. The request could be renewed if the employer so desired. These workers could be forcibly recruited unless they could demonstrate from their personal workbook that such service had recently been performed satisfactorily. All rural workers were also forced to carry a workbook, called a *libreta*, which included details of whom they were working for and a record of any debts. Many rural workers were indebted to their employers, and an indebted worker could not leave his current employer without permission. Decree 177 further stipulated that the only way to avoid being drafted into the *repartimiento* was to show you were currently in debt to an employer. Workers were trapped. In addition to these laws, numerous vagrancy laws were passed so that anyone who could not prove he had a job would be immediately recruited for the *repartimiento* or other types of forced labor on the roads, or would be forced to accept employment on a farm. As in nineteenth- and twentieth-century South Africa, land policies after 1871 were also designed to undermine the subsistence economy of the indigenous peoples, to force them to work for low wages. The *repartimiento* lasted until the 1920s; the *libreta* system and the full gamut of vagrancy laws were in effect until 1945, when Guatemala experienced its first brief flowering of democracy.

Just as before 1871, the Guatemalan elite ruled via military strongmen. They continued to do so after the coffee boom took off. Jorge Ubico, president between 1931 and 1944, ruled longest. Ubico won the presidential election in 1931 unopposed, since nobody was foolish enough to run against him. Like the Consulado, he didn't approve of doing things that would have induced creative destruction and threatened both his political power and his and the elite's profits. He therefore opposed industry for the same reason that Francis I in Austria-Hungary and Nicholas I in Russia did: industrial workers would have caused trouble. In a legislation unparalleled in its paranoid repressiveness, Ubico banned the use of words such as *obreros* (workers), *sindicatos* (labor unions), and *huelgas* (strikes). You could be jailed for using any one of them. Even though Ubico was powerful, the elite pulled the strings. Opposition to his regime mounted in 1944, headed by disaffected university students who began to organize demonstrations. Popular discontent increased, and on June 24, 311 people, many of them from the elite, signed the Memorial de los 311, an open letter denouncing the regime. Ubico resigned on July 1. Though he was followed by a democratic regime in 1945, this was overthrown by a coup in 1954, leading to a murderous civil war. Guatemala democratized again after only 1986.

The Spanish conquistadors had no compunction about setting up an extractive political and economic system. That was why they had come all the way to the New World. But most of the institutions they set up were meant to be temporary. The *encomienda*, for example, was a temporary grant of rights over labor. They did not have a fully worked-out plan of how they would set up a system that would persist for another four hundred years. In fact, the institutions they set up changed significantly along the way, but one thing did not: the extractive nature of the institutions, the result of the vicious circle. The form of extraction changed, but neither the extractive nature of the institutions nor the identity of the elite did. In Guatemala the *encomienda*, the *repartimiento*, and the monopolization of trade gave way to the *libreta* and the land grab. But the majority of the indigenous Maya continued to work as low-wage laborers with little education, no rights, and no public services.

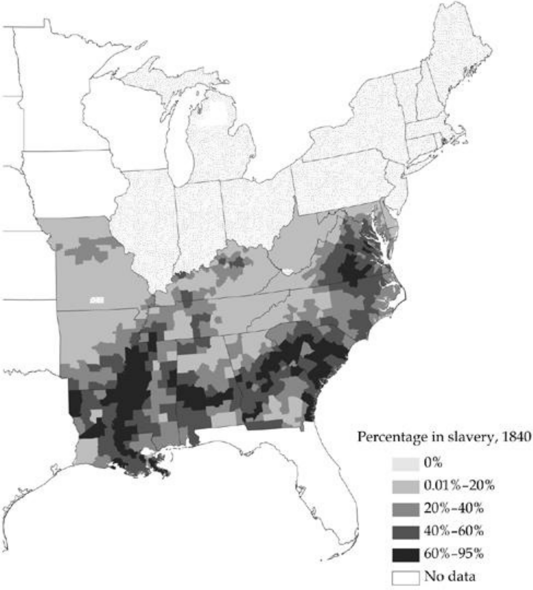
In Guatemala, as in much of Central America, in a typical pattern of the vicious circle, extractive political institutions supported extractive economic institutions, which in turn provided the basis for extractive political institutions and the continuation of the power of the same elite.

FROM SLAVERY TO JIM CROW

In Guatemala, extractive institutions persisted from colonial to modern times with the same elite firmly in control. Any change in institutions resulted from adaptations to changing environments, as was the case with the land grab by the elite motivated by the coffee boom. The institutions in the U.S. South were similarly extractive until the Civil War. Economics and politics were dominated by the southern elite, plantation owners with large land and slave holdings. Slaves had neither political nor economic rights; indeed, they had few rights of any kind.

The South's extractive economic and political institutions made it considerably poorer than the North by the middle of the nineteenth century. The South lacked industry and made relatively little investment in infrastructure. In 1860 its total manufacturing output was less than that of Pennsylvania, New York, or Massachusetts. Only 9 percent of the southern population lived in urban areas, compared with 35 percent in the Northeast. The density of railroads (i.e., miles of track divided by land area) was three times higher in the North than in southern states. The ratio of canal mileage was similar.

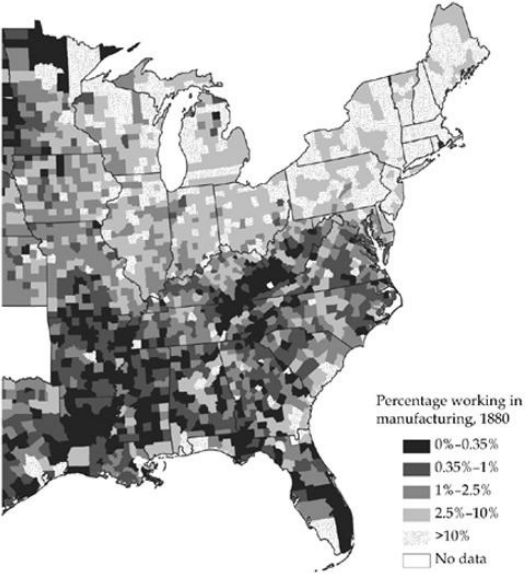
Map 18 (this page) shows the extent of slavery by plotting the percentage of the population that were slaves across U.S. counties in 1840. It is apparent that slavery was dominant in the South with some counties, for example, along the Mississippi River having as much as 95 percent of the population slaves. Map 19 (this page) then shows one of the consequences of this, the proportion of the labor force working in manufacturing in 1880. Though this was not high anywhere by twentieth-century standards, there are marked differences between the North and the South. In much of the Northeast, more than 10 percent of the labor force worked in manufacturing. In contrast in much of the South, particularly the areas with heavy concentrations of slaves, the proportion was basically zero.



Map 18: Slavery across U.S. counties in 1840

The South was not even innovative in the sectors in which it specialized: from 1837 to 1859, the numbers of patents issued per year for innovations related to corn and wheat were on average twelve and ten, respectively; there was just one per year for the most important crop of the South, cotton. There was no indication that industrialization and economic growth would commence anytime soon. But defeat in the Civil War was followed by fundamental economic and political reform at bayonet point. Slavery was abolished, and black men were allowed to vote.

These major changes should have opened the way for a radical transformation of southern extractive institutions into inclusive ones, and launched the South onto a path to economic prosperity. But in yet another manifestation of the vicious circle, nothing of the sort happened. A continuation of extractive institutions, this time of the Jim Crow kind rather than of slavery, emerged in the South. The phrase *Jim Crow*, which supposedly originated from "Jump Jim Crow," an early-nineteenth-century satire of black people performed by white performers in "blackface," came to refer to the whole gamut of segregationist legislation that was enacted in the South after 1865. These persisted for almost another century, until yet another major upheaval, the civil rights movement. In the meantime, blacks continued to be excluded from power and repressed. Plantation-type agriculture based on low-wage, poorly educated labor persisted, and southern incomes fell further relative to the U.S. average. The vicious circle of extractive institutions was stronger than many had expected at the time.



Map 19: Manufacturing employment across U.S. counties in 1880

The reason that the economic and political trajectory of the South never changed, even though slavery was abolished and black men were given the right to vote, was because blacks' political power and economic independence were tenuous. The southern planters lost the war, but would win the peace. They were still organized and they still owned the land. During the war, freed slaves had been offered the promise of forty acres and a mule when slavery was abolished, and some even got it during the famous campaigns of General William T. Sherman. But in 1865, President Andrew Johnson revoked Sherman's orders, and the hoped-for land redistribution never took place. In a debate on this issue in Congress, Congressman George Washington Julian presciently noted, “Of what avail would be an act of congress totally abolishing slavery … if the old agricultural basis of aristocratic power shall remain?” This was the beginning of the “redemption” of the old South and the persistence of the old southern landed elite.

The sociologist Jonathan Wiener studied the persistence of the planter elite in five counties of the Black Belt, prime cotton country, of southern Alabama. Tracking families from the U.S. census and considering those with at least $10,000 of real estate, he found that of the 236 members of the planter elite in 1850, 101 maintained their position in 1870. Interestingly, this rate of persistence was very similar to that experienced in the pre–Civil War period; of the 236 wealthiest planter families of 1850, only 110 remained so a decade later. Nevertheless, of the 25 planters with the largest landholdings in 1870, 18 (72 percent) had been in the elite families in 1860; 16 had been in the 1850 elite group. While more than 600,000 were killed in the Civil War, the planter elites suffered few casualties. The law, designed by the planters and for the planters, exempted one slaveholder from military service for every twenty slaves held. As hundreds of thousands of men died to preserve the southern plantation economy, many big slaveholders and their sons sat out the war on their porches and thus were able to ensure the persistence of the plantation economy.

After the end of the war, the elite planters controlling the land were able to reexert their control over the labor force. Though the economic institution of slavery was abolished, the evidence shows a clear line of persistence in the economic system of the South based on plantation-type agriculture with cheap labor. This economic system was maintained through a variety of channels, including both control of local politics and exercise of violence. As a consequence, in the words of the African American scholar W.E.B. Du Bois, the South became “simply an armed camp for intimidating black folk.”

In 1865 the state legislature of Alabama passed the Black Code, an important landmark toward the repression of black labor. Similar to Decree 177 in Guatemala, the Black Code of Alabama consisted of a vagrancy law and a law against the "enticement" of laborers. It was designed to impede labor mobility and reduce competition in the labor market, and it ensured that southern planters would still have a reliable low-cost labor pool.

Following the Civil War, the period called Reconstruction lasted from 1865 until 1877. Northern politicians, with the help of the Union Army, engineered some social changes in the South. But a systematic backlash from the southern elite in the guise of support for the so-called Redeemers, seeking the South's redemption, re-created the old system. In the 1877 presidential election, Rutherford Hayes needed southern support in the electoral college. This college, still used today, was at the heart of the indirect election for president created by the U.S. Constitution. Citizens' votes do not directly elect the president but instead elect electors who then choose the president in the electoral college. In exchange for their support in the electoral college, the southerners demanded that Union soldiers be withdrawn from the South and the region left to its own devices. Hayes agreed. With southern support, Hayes became president and pulled out the troops. The period after 1877 then marked the real reemergence of the pre–Civil War planter elite. The redemption of the South involved the introduction of new poll taxes and literacy tests for voting, which systematically disenfranchised blacks, and often also the poor white population. These attempts succeeded and created a one-party regime under the Democratic Party, with much of the political power vested in the hands of the planter elite.

The Jim Crow laws created separate, and predictably inferior, schools. Alabama, for example, rewrote its constitution in 1901 to achieve this. Shockingly, even today Section 256 of Alabama's constitution, though no longer enforced, still states:

> Duty of legislature to establish and maintain public school system; apportionment of public school fund; separate schools for white and colored children.
>
> The legislature shall establish, organize,

> and maintain a liberal system of public schools throughout the state for the benefit of the children thereof between the ages of seven and twenty-one years. The public school fund shall be apportioned to the several counties in proportion to the number of school children of school age therein, and shall be so apportioned to the schools in the districts or townships in the counties as to provide, as nearly as practicable, school terms of equal duration in such school districts or townships. Separate schools shall be provided for white and colored children, and no child of either race shall be permitted to attend a school of the other race.

An amendment to strike Section 256 from the constitution was narrowly defeated in the state legislature in 2004.

Disenfranchisement, the vagrancy laws such as the Black Code of Alabama, various Jim Crow laws, and the actions of the Ku Klux Klan, often financed and supported by the elite, turned the post–Civil War South into an effective apartheid society, where blacks and whites lived different lives. As in South Africa, these laws and practices were aimed at controlling the black population and its labor.

Southern politicians in Washington also worked to make sure that the extractive institutions of the South could persist. For instance, they ensured that no federal projects or public works that would have jeopardized southern elite control over the black workforce ever got approved. Consequently, the South entered the twentieth century as a largely rural society with low levels of education and backward technology, still employing hand labor and mule power virtually unassisted by mechanical implements. Though the proportion of people in urban areas increased, it was far less than in the North. In 1900, for example, 13.5 percent of the population of the South was urbanized, as compared with 60 percent in the Northeast.

All in all, the extractive institutions in the southern United States, based on the power of the landed elite, plantation agriculture, and low-wage, low-education labor, persisted well into the twentieth century. These institutions started to crumble only after the Second World War and then truly after the civil rights movement destroyed the political basis of the system. And it was only after the demise of these institutions in the 1950s and '60s that the South began its process of rapid convergence to the North.

The U.S. South shows another, more resilient side of the vicious circle: as in Guatemala, the southern planter elite remained in power and structured economic and political institutions in order to ensure the continuity of its power. But differently from Guatemala, it was faced with significant challenges after its defeat in the Civil War, which abolished slavery and reversed the total, constitutional exclusion of blacks from political participation. But there is more than one way of skinning a cat: as long as the planter elite was in control of its huge landholdings and remained organized, it could structure a new set of institutions, Jim Crow instead of slavery, to achieve the same objective. The vicious circle turned out to be stronger than many, including Abraham Lincoln, had thought. The vicious circle is based on extractive political institutions creating extractive economic institutions, which in turn support the extractive political institutions, because economic wealth and power buy political power. When forty acres and a mule was off the table, the southern planter elite's economic power remained untarnished. And, unsurprisingly and unfortunately, the implications for the black population of the South, and the South's economic development, were the same.

THE IRON LAW OF OLIGARCHY

The Solomonic dynasty in Ethiopia lasted until it was overthrown by a military coup in 1974. The coup was led by the Derg, a group of Marxist army officers. The regime that the Derg pitched from power looked like it was frozen in some earlier century, a historical anachronism. The emperor Haile Selassie would start his day by arriving in the courtyard at the Grand Palace, which had been built by Emperor Menelik II in the late nineteenth century. Outside the palace would be a crowd of dignitaries anticipating his arrival, bowing and desperately trying to get his attention.

The emperor would hold court in the Audience Hall, sitting on the imperial throne. (Selassie was a small man; so that his legs were not left swinging in the air, it was the job of a special pillow bearer to accompany him wherever he went to make sure there was a suitable pillow to put under his feet. The bearer kept a stock of fifty-two pillows to cope with any situation.) Selassie presided over an extreme set of extractive institutions and ran the country as his own private property, handing out favors and patronage and ruthlessly punishing lack of loyalty. There was no economic development to speak of in Ethiopia under the Solomonic dynasty.

The Derg initially formed out of 108 representatives of different military units from all over the country. The representative of the Third Division in Harar province was a major named Mengistu Haile Mariam. Though in their initial declaration of July 4, 1974, the Derg officers declared their loyalty to the emperor, they soon started to arrest members of the government, testing how much opposition it would create. As they became more confident that the support for Selassie's regime was hollow, they moved on the emperor himself, arresting him on September 12. Then the executions began. Many politicians at the core of the old regime were swiftly killed. By December, the Derg had declared that Ethiopia was a socialist state. Selassie died, probably murdered, on August 27, 1975. In 1975 the Derg started nationalizing property, including all urban and rural land and most kinds of private property. The increasingly authoritarian behavior of the regime sparked opposition around the country. Large parts of Ethiopia were put together during the European colonial expansion in the late nineteenth and early twentieth centuries by the policies of Emperor Menelik II, the victor of the battle of Adowa, which we encountered before (this page). These included Eritrea and Tigray in the north and the Ogaden in the east. Independence movements in response to the Derg's ruthless regime emerged in Eritrea and Tigray, while the Somali army invaded the Somali-speaking Ogaden. The Derg itself started to disintegrate and split into factions. Major Mengistu turned out to be the most ruthless and clever of them. By mid-1977 he had eliminated his major opponents and effectively taken charge of the regime, which was saved from collapse only by a huge influx of weapons and troops from the Soviet Union and Cuba later in November of that year.

In 1978 the regime organized a national celebration marking the fourth anniversary of the overthrow of Haile Selassie. By this time Mengistu was the unchallenged leader of the Derg. As his residence, the place from where he would rule Ethiopia, he had chosen Selassie's Grand Palace, left unoccupied since the monarchy was abolished. At the celebration, he sat on a gilded armchair, just like the emperors of old, watching the parade. Official functions were now held once again at the Grand Palace, with Mengistu sitting on Haile Selassie's old throne. Mengistu started to compare himself to Emperor Tewodros, who had refounded the Solomonic Dynasty in the mid-nineteenth century after a period of decline.

One of his ministers, Dawit Wolde Giorgis, recalled in his memoir:

> At the beginning of the Revolution all of us had utterly rejected anything to do with the past. We would no longer drive cars, or wear suits; neckties were considered criminal. Anything that made you look well-off or bourgeois, anything that smacked of affluence or sophistication, was scorned as part of the old order. Then, around 1978, all that began to change. Gradually materialism became accepted, then required. Designer clothes from the best European tailors were the uniform of all senior government officials and members of the Military Council. We had the best of everything: the best homes, the best cars, the best whiskey, champagne, food. It was a complete reversal of the ideals of the Revolution.

Giorgis also vividly recorded how Mengistu changed once he became sole ruler:

> The real Mengistu emerged: vengeful, cruel and authoritarian … Many of us who used to talk to him with hands in our pockets, as if he

> were one of us, found ourselves standing stiffly to attention, cautiously respectful in his presence. In addressing him we had always used the familiar form of “you,” *ante;* now we found ourselves switching to the more formal “you,” *ersiwo*. He moved into a bigger, more lavish office in the Palace of Menelik … He began using the Emperor’s cars … We were supposed to have a revolution of equality; now he had become the new Emperor.

The pattern of vicious circle depicted by the transition between Haile Selassie and Mengistu, or between the British colonial governors of Sierra Leone and Siaka Stevens, is so extreme and at some level so strange that it deserves a special name. As we already mentioned in chapter 4, the German sociologist Robert Michels called it the iron law of oligarchy. The internal logic of oligarchies, and in fact of all hierarchical organizations, is that, argued Michels, they will reproduce themselves not only when the same group is in power, but even when an entirely new group takes control. What Michels did not anticipate perhaps was an echo of Karl Marx’s remark that history repeats itself—the first time as tragedy, the second time as farce.

It is not only that many of the postindependence leaders of Africa moved into the same residences, made use of the same patronage networks, and employed the same ways of manipulating markets and extracting resources as had the colonial regimes and the emperors they replaced; but they also made things worse. It was indeed a farce that the staunchly anticolonial Stevens would be concerned with controlling the same people, the Mende, whom the British had sought to control; that he would rely on the same chiefs whom the British had empowered and then used to control the hinterland; that he would run the economy in the same way, expropriating the farmers with the same marketing boards and controlling the diamonds under a similar monopoly. It was indeed a farce, a very sad farce indeed, that Laurent Kabila, who mobilized an army against Mobutu’s dictatorship with the promise of freeing the people and ending the stifling and impoverishing corruption and repression of Mobutu's Zaire, would then set up a regime just as corrupt and perhaps even more disastrous. It was certainly farcical that he tried to start a Mobutuesque personality cult aided and abetted by Dominique Sakombi Inongo, previously Mobutu's minister of information, and that Mobutu's regime was itself fashioned on patterns of exploitation of the masses that had started more than a century previously with King Leopold's Congo Free State. It was indeed a farce that the Marxist officer Mengistu would start living in a palace, viewing himself as an emperor, and enriching himself and his entourage just like Haile Selassie and other emperors before him had done.

It was all a farce, but also more tragic than the original tragedy, and not only for the hopes that were dashed. Stevens and Kabila, like many other rulers in Africa, would start murdering their opponents and then innocent citizens. Mengistu and the Derg's policies would bring recurring famine to Ethiopia's fertile lands. History was repeating itself, but in a very distorted form. It was a famine in Wollo province in 1973 to which Haile Selassie was apparently indifferent that did so much finally to solidify opposition to his regime. Selassie had at least been only indifferent. Mengistu instead saw famine as a political tool to undermine the strength of his opponents. History was not only farcical and tragic, but also cruel to the citizens of Ethiopia and much of sub-Saharan Africa.

The essence of the iron law of oligarchy, this particular facet of the vicious circle, is that new leaders overthrowing old ones with promises of radical change bring nothing but more of the same. At some level, the iron law of oligarchy is harder to understand than other forms of the vicious circle. There is a clear logic to the persistence of the extractive institutions in the U.S. South and in Guatemala. The same groups continued to dominate the economy and the politics for centuries. Even when challenged, as the U.S. southern planters were after the Civil War, their power remained intact and they were able to keep and re-create a similar set of extractive institutions from which they would again benefit. But how can we understand those who come to power in the name of radical change re-creating the same system? The answer to this question reveals, once again, that the vicious circle is stronger than it first appears.

Not all radical changes are doomed to failure. The Glorious Revolution was a radical change, and it led to what perhaps turned out to be the most important political revolution of the past two millennia. The French Revolution was even more radical, with its chaos and excessive violence and the ascent of Napoleon Bonaparte, but it did not re-create the *ancien régime*.

Three factors greatly facilitated the emergence of more inclusive political institutions following the Glorious Revolution and the French Revolution. The first was new merchants and businessmen wishing to unleash the power of creative destruction from which they themselves would benefit; these new men were among the key members of the revolutionary coalitions and did not wish to see the development of yet another set of extractive institutions that would again prey on them.

The second was the nature of the broad coalition that had formed in both cases. For example, the Glorious Revolution wasn't a coup by a narrow group or a specific narrow interest, but a movement backed by merchants, industrialists, the gentry, and diverse political groupings. The same was largely true for the French Revolution.

The third factor relates to the history of English and French political institutions. They created a background against which new, more inclusive regimes could develop. In both countries there was a tradition of parliaments and power sharing going back to the Magna Carta in England and to the Assembly of Notables in France. Moreover, both revolutions happened in the midst of a process that had already weakened the grasp of the absolutist, or aspiring absolutist, regimes. In neither case would these political institutions make it easy for a new set of rulers or a narrow group to take control of the state and usurp existing economic wealth and build unchecked and durable political power. In the aftermath of the French Revolution, a narrow group under the leadership of Robespierre and Saint-Just did take control, with disastrous consequences, but this was temporary and did not derail the path toward more inclusive institutions. All this contrasts with the situation of societies with long histories of extreme extractive economic and political institutions, and no checks on the power of rulers. In these societies, there would be no new strong merchants or businessmen supporting and bankrolling the resistance against the existing regime in part to secure more inclusive economic institutions; no broad coalitions introducing constraints against the power of each of their members; no political institutions inhibiting new rulers intent on usurping and exploiting power.

In consequence, in Sierra Leone, Ethiopia, and the Congo, the vicious circle would be far harder to resist, and moves toward inclusive institutions far more unlikely to get under way. There were also no traditional or historical institutions that could check the power of those who would take control of the state. Such institutions had existed in some parts of Africa, and some, as in Botswana, even survived the colonial era. But they were much less prominent throughout Sierra Leone's history, and to the extent that they existed, they were warped by indirect rule. The same was true in other British colonies in Africa, such as Kenya and Nigeria. They never existed in the absolutist kingdom of Ethiopia. In the Congo, indigenous institutions were emasculated by Belgian colonial rule and the autocratic policies of Mobutu. In all these societies, there were also no new merchants, businessmen, or entrepreneurs supporting the new regimes and demanding secure property rights and an end to previous extractive institutions. In fact, the extractive economic institutions of the colonial period meant that there was not much entrepreneurship or business left at all.

The international community thought that postcolonial African independence would lead to economic growth through a process of state planning and cultivation of the private sector. But the private sector was not there—except in rural areas, which had no representation in the new governments and would thus be their first prey. Most important perhaps, in most of these cases there were enormous benefits from holding power. These benefits both attracted the most unscrupulous men, such as Stevens, who wished to monopolize this power, and brought the worst out of them once they were in power. There was nothing to break the vicious circle.

NEGATIVE FEEDBACK AND VICIOUS CIRCLES

Rich nations are rich largely because they managed to develop inclusive institutions at some point during the past three hundred years. These institutions have persisted through a process of virtuous circles. Even if inclusive only in a limited sense to begin with, and sometimes fragile, they generated dynamics that would create a process of positive feedback, gradually increasing their inclusiveness. England did not become a democracy after the Glorious Revolution of 1688. Far from it. Only a small fraction of the population had formal representation, but crucially, she was pluralistic. Once pluralism was enshrined, there was a tendency for the institutions to become more inclusive over time, even if this was a rocky and uncertain process.

In this, England was typical of virtuous circles: inclusive political institutions create constraints against the exercise and usurpation of power. They also tend to create inclusive economic institutions, which in turn make the continuation of inclusive political institutions more likely.

Under inclusive economic institutions, wealth is not concentrated in the hands of a small group that could then use its economic might to increase its political power disproportionately. Furthermore, under inclusive economic institutions there are more limited gains from holding political power, thus weaker incentives for every group and every ambitious, upstart individual to try to take control of the state. A confluence of factors at a critical juncture, including interplay between existing institutions and the opportunities and challenges brought by the critical juncture, is generally responsible for the onset of inclusive institutions, as the English case demonstrates. But once these inclusive institutions are in place, we do not need the same confluence of factors for them to survive. Virtuous circles, though still subject to significant contingency, enable the institutions' continuity and often even unleash dynamics taking society toward greater inclusiveness.

As virtuous circles make inclusive institutions persist, vicious circles create powerful forces toward the persistence of extractive institutions. History is not destiny, and vicious circles are not unbreakable, as we will see further in chapter 14. But they are resilient. They create a powerful process of negative feedback, with extractive political institutions forging extractive economic institutions, which in turn create the basis for the persistence of extractive political institutions. We saw this most clearly in the case of Guatemala, where the same elite held power, first under colonial rule, then in independent Guatemala, for more than four centuries; extractive institutions enrich the elite, and their wealth forms the basis for the continuation of their domination.

The same process of the vicious circle is also apparent in the persistence of the plantation economy in the U.S. South, except that it also showcases the vicious circle's great resilience in the face of challenges. U.S. southern planters lost their formal control of economic and political institutions after their defeat in the Civil War. Slavery, which was the basis of the plantation economy, was abolished, and blacks were given equal political and economic rights. Yet the Civil War did not destroy the political power of the planter elite or its economic basis, and they were able to restructure the system, under a different guise but still under their own local political control, and to achieve the same objective: abundance of low-cost labor for the plantations.

This form of the vicious circle, where extractive institutions persist because the elite controlling them and benefiting from them persists, is not its only form. At first a more puzzling, but no less real and no less vicious, form of negative feedback shaped the political and economic development of many nations, and is exemplified by the experiences of much of sub-Saharan Africa, in particular Sierra Leone and Ethiopia. In a form that the sociologist Robert Michels would recognize as the iron law of oligarchy, the overthrow of a regime presiding over extractive institutions heralds the arrival of a new set of masters to exploit the same set of pernicious extractive institutions.

The logic of this type of vicious circle is also simple to understand in hindsight: extractive political institutions create few constraints on the exercise of power, so there are essentially no institutions to restrain the use and abuse of power by those overthrowing previous dictators and assuming control of the state; and extractive economic institutions imply that there are great profits and wealth to be made merely by controlling power, expropriating the assets of others, and setting up monopolies.

Of course, the iron law of oligarchy is not a true law, in the sense that the laws of physics are. It does not chart an inevitable path, as the Glorious Revolution in England or the Meiji Restoration in Japan illustrate.

A key factor in these episodes, which saw a major turn toward inclusive institutions, was the empowerment of a broad coalition that could stand up against absolutism and would replace the absolutist institutions by more inclusive, pluralistic ones. A revolution by a broad coalition makes the emergence of pluralistic political institutions much more likely. In Sierra Leone and Ethiopia, the iron law of oligarchy was made more likely not only because existing institutions were highly extractive but also because neither the independence movement in the former nor the Derg coup in the latter were revolutions led by such broad coalitions, but rather by individuals and groups seeking power so that they could do the extracting.

There is yet another, even more destructive facet of the vicious circle, anticipated by our discussion of the Maya city-states in chapter 5. When extractive institutions create huge inequalities in society and great wealth and unchecked power for those in control, there will be many wishing to fight to take control of the state and institutions. Extractive institutions then not only pave the way for the next regime, which will be even more extractive, but they also engender continuous infighting and civil wars. These civil wars then cause more human suffering and also destroy even what little state centralization these societies have achieved. This also often starts a process of descent into lawlessness, state failure, and political chaos, crushing all hopes of economic prosperity, as the next chapter will illustrate.

13.

WHY NATIONS FAIL TODAY

HOW TO WIN THE LOTTERY IN ZIMBABWE

IT WAS JANUARY 2000 in Harare, Zimbabwe. Master of Ceremonies Fallot Chawawa was in charge of drawing the winning ticket for the national lottery organized by a partly state-owned bank, the Zimbabwe Banking Corporation (Zimbank). The lottery was open to all clients who had kept five thousand or more Zimbabwe dollars in their accounts during December 1999. When Chawawa drew the ticket, he was dumfounded. As the public statement of Zimbank put it, "Master of Ceremonies Fallot Chawawa could hardly believe his eyes when the ticket drawn for the Z$100,000 prize was handed to him and he saw His Excellency RG Mugabe written on it."

President Robert Mugabe, who had ruled Zimbabwe by hook or by crook, and usually with an iron fist, since 1980, had won the lottery, which was worth a hundred thousand Zimbabwe dollars, about five times the annual per capita income of the country. Zimbank claimed that Mr. Mugabe's name had been drawn from among thousands of eligible customers. What a lucky man! Needless to say he didn't really need the money. Mugabe had in fact only recently awarded himself and his cabinet salary hikes of up to 200 percent.

The lottery ticket was just one more indication of Zimbabwe's extractive institutions. One could call this corruption, but it is just a symptom of the institutional malaise in Zimbabwe. The fact that Mugabe could even win the lottery if he wanted showed how much control he had over matters in Zimbabwe, and gave the world a glimpse of the extent of the country's extractive institutions.

The most common reason why nations fail today is because they have extractive institutions. Zimbabwe under Mugabe's regime vividly illustrates the economic and social consequences. Though the national statistics in Zimbabwe are very unreliable, the best estimate is that by 2008, Zimbabwe's per capita income was about half of what it was when the country gained its independence in 1980. Dramatic as this sounds, it does not in fact begin to capture the deterioration in living standards in Zimbabwe. The state has collapsed and more or less stopped providing any basic public services. In 2008–2009 the deterioration in the health systems led to an outbreak of cholera across the country. As of January 10, 2010, there have been 98,741 reported cases and 4,293 deaths, making it the deadliest cholera outbreak in Africa over the previous fifteen years. In the meantime, mass unemployment has also reached unprecedented levels. In early 2009, the UN Office for the Coordination of Humanitarian Affairs claimed that the unemployment rate had hit an incredible 94 percent.

The roots of many economic and political institutions in Zimbabwe, as is the case for much of sub-Saharan Africa, can be traced back to the colonial period. In 1890 Cecil Rhodes's British South Africa Company sent a military expedition into the then-kingdom of the Ndebele, based in Matabeleland, and also into the neighboring Mashonaland. Their superior weaponry quickly suppressed African resistance, and by 1901 the colony of Southern Rhodesia, named after Rhodes, had been formed in the area that is currently Zimbabwe. Now that the area was a privately owned concession of the British South Africa Company, Rhodes anticipated making money there through prospecting and mining for precious minerals. The ventures never got off the ground, but the very rich farmlands began attracting white migration. These settlers soon annexed much of the land. By 1923 they had freed themselves from the rule of the British South Africa Company and persuaded the British government to grant them self-government. What then occurred is very similar to what had happened in South Africa a decade or so previously. The 1913 Natives Land Act (this page–this page) created a dual economy in South Africa. Rhodesia passed very similar laws, and inspired by the South African model, a white-only apartheid state was constructed soon after 1923.

As the European colonial empires collapsed in the late

1950s and early 1960s, the white elite in Rhodesia, led by Ian Smith, comprising possibly 5 percent of the population, declared independence from Britain in 1965. Few international governments recognized Rhodesia's independence, and the United Nations levied economic and political sanctions against it. The black citizens organized a guerrilla war from bases in the neighboring countries of Mozambique and Zambia. International pressure and the rebellion waged by the two main groups, Mugabe's ZANU (the Zimbabwe African National Union) and ZAPU (the Zimbabwe African People's Union), led by Joshua Nkomo, resulted in a negotiated end to white rule. The state of Zimbabwe was created in 1980.

After independence, Mugabe quickly established his personal control. He either violently eliminated his opponents or co-opted them. The most egregious acts of violence happened in Matabeleland, the heartland of support for ZAPU, where as many as twenty thousand people were killed in the early 1980s. By 1987 ZAPU had merged with ZANU to create ZANU-PF, and Joshua Nkomo was sidelined politically. Mugabe was able to rewrite the constitution he had inherited as a part of the independence negotiation, making himself president (he had started as prime minister), abolishing white voter rolls that were part of the independence agreement, and eventually, in 1990, getting rid of the Senate altogether and introducing positions in the legislature that he could nominate. A de facto one-party state headed by Mugabe was the result.

Upon independence, Mugabe took over a set of extractive economic institutions created by the white regime. These included a host of regulations on prices and international trade, state-run industries, and the obligatory agricultural marketing boards. State employment expanded rapidly, with jobs given to supporters of ZANU-PF. The tight government regulation of the economy suited the ZANU-PF elites because it made it difficult for an independent class of African businessmen, who might then have challenged the former's political monopoly, to emerge. This was very similar to the situation we saw in Ghana in the 1960s in chapter 2 (this page–this page). Ironically, of course, this left whites as the main business class. During this period the main strengths of the white economy, particularly the highly productive agricultural export sector, was left untouched. But this would last only until Mugabe became unpopular.

The model of regulation and market intervention gradually became unsustainable, and a process of institutional change, with the support of the World Bank and the International Monetary Fund, began in 1991 after a severe fiscal crisis. The deteriorating economic performance finally led to the emergence of a serious political opposition to ZANU-PF's one-party rule: the Movement for Democratic Change (MDC). The 1995 parliamentary elections were far from competitive. ZANU-PF won 81 percent of the vote and 118 out of the 120 seats. Fifty-five of these members of Parliament were elected unopposed. The presidential election the following year showed even more signs of irregularities and fraud. Mugabe won 93 percent of the vote, but his two opponents, Abel Muzorewa and Ndabaningi Sithole, had already withdrawn their candidacy prior to the election, accusing the government of coercion and fraud.

After 2000, despite all the corruption, ZANU-PF's grip was weakening. It took only 49 percent of the popular vote, and only 63 seats. All were contested by the MDC, who took every seat in the capital, Harare. In the presidential election of 2002, Mugabe scraped home with only 56 percent of the vote. Both sets of elections went ZANU-PF's way only because of violence and intimidation, coupled with electoral fraud.

The response of Mugabe to the breakdown of his political control was to intensify both the repression and the use of government policies to buy support. He unleashed a full-scale assault on white landowners. Starting in 2000, he encouraged and supported an extensive series of land occupations and expropriations. They were often led by war veterans' associations, groups supposedly comprised of former combatants in the war of independence. Some of the expropriated land was given to these groups, but much of it also went to the ZANU-PF elites. The insecurity of property rights wrought by Mugabe and ZANU-PF led to a collapse of agricultural output and productivity. As the economy crumbled, the only thing left was to print money to buy support, which led to enormous hyperinflation. In January 2009, it became legal to use other currencies, such as the South African rand, and the Zimbabwean dollar vanished from circulation, a worthless piece of paper.

What happened in Zimbabwe after 1980 was commonplace in sub-Saharan Africa since independence. Zimbabwe inherited a set of highly extractive political and economic institutions in 1980. For the first decade and a half, these were maintained relatively untouched. While elections took place, political institutions were anything but inclusive. Economic institutions changed somewhat; for example, there was no longer explicit discrimination against blacks. But on the whole the institutions remained extractive, with the only difference being that instead of Ian Smith and the whites doing the extracting, it was Robert Mugabe and the ZANU-PF elites filling their pockets. Over time the institutions became even more extractive, and incomes in Zimbabwe collapsed. The economic and political failure in Zimbabwe is yet another manifestation of the iron law of oligarchy—in this instance, with the extractive and repressive regime of Ian Smith being replaced by the extractive, corrupt, and repressive regime of Robert Mugabe. Mugabe's fake lottery win in 2000 was then simply the tip of a very corrupt and historically shaped iceberg.

NATIONS FAIL TODAY because their extractive economic institutions do not create the incentives needed for people to save, invest, and innovate. Extractive political institutions support these economic institutions by cementing the power of those who benefit from the extraction. Extractive economic and political institutions, though their details vary under different circumstances, are always at the root of this failure. In many cases, for example, as we will see in Argentina, Colombia, and Egypt, this failure takes the form of lack of sufficient economic activity, because the politicians are just too happy to extract resources or quash any type of independent economic activity that threatens themselves and the economic elites. In some extreme cases, as in Zimbabwe and Sierra Leone, which we discuss next, extractive institutions pave the way for complete state failure, destroying not only law and order but also even the most basic economic incentives. The result is economic stagnation and—as the recent history of Angola, Cameroon, Chad, the Democratic Republic of Congo, Haiti, Liberia, Nepal, Sierra Leone, Sudan, and Zimbabwe illustrates—civil wars, mass displacements, famines, and epidemics, making many of these countries poorer today than they were in the 1960s.

A CHILDREN'S CRUSADE?

On March 23, 1991, a group of armed men under the leadership of Foday Sankoh crossed the border from Liberia into Sierra Leone and attacked the southern frontier town of Kailahun. Sankoh, formerly a corporal in the Sierra Leonean army, had been imprisoned after taking part in an abortive coup against Siaka Stevens's government in 1971. After being released, he eventually ended up in Libya, where he entered a training camp that the Libyan dictator Colonel Qaddafi ran for African revolutionaries. There he met Charles Taylor, who was plotting to overthrow the government in Liberia. When Taylor invaded Liberia on Christmas Eve 1989, Sankoh was with him, and it was with a group of Taylor's men, mostly Liberians and Burkinabes (citizens of Burkina Faso), that Sankoh invaded Sierra Leone. They called themselves the RUF, the Revolutionary United Front, and they announced that they were there to overthrow the corrupt and tyrannical government of the APC.

As we saw in the previous chapter, Siaka Stevens and his All People's Congress, the APC, took over and intensified the extractive institutions of colonial rule in Sierra Leone, just as Mugabe and ZANU-PF did in Zimbabwe. By 1985, when Stevens, ill with cancer, brought in Joseph Momoh to replace him, the economy was collapsing. Stevens, apparently without irony, used to enjoy quoting the aphorism "The cow eats where it is tethered." And where Stevens had once eaten, Momoh now gorged. The roads fell to pieces, and schools disintegrated. National television broadcasts stopped in 1987, when the transmitter was sold by the minister of information, and in 1989 a radio tower that relayed radio signals outside

Freetown fell down, ending transmissions outside the capital. An analysis published in a newspaper in the capital city of Freetown in 1995 rings very true:

> by the end of Momoh's rule he had stopped paying civil servants, teachers and even Paramount Chiefs. Central government had collapsed, and then of course we had border incursions, "rebels" and all the automatic weapons pouring over the border from Liberia. The NPRC, the "rebels" and the "sobels" [soldiers turned rebels] all amount to the chaos one expects when government disappears. None of them are the causes of our problems, but they are symptoms.

The collapse of the state under Momoh, once again a consequence of the vicious circle unleashed by the extreme extractive institutions under Stevens, meant that there was nothing to stop the RUF from coming across the border in 1991. The state had no capacity to oppose it. Stevens had already emasculated the military, because he worried they might overthrow him. It was then easy for a relatively small number of armed men to create chaos in most of the country. They even had a manifesto called "Footpaths to Democracy," which started with a quote from the black intellectual Frantz Fanon: "Each generation must, out of relative obscurity, discover its mission, fulfill it or betray it." The section "What Are We Fighting For?" begins:

> We continue to fight because we are tired of being perpetual victims of state sponsored poverty and human degradation visited on us by years of autocratic rule and militarism. But, we shall exercise restraint and continue to wait patiently at the rendezvous of peace—where we shall all be winners. We are committed to peace, by any means necessary, but what we are not committed to is becoming victims of peace. We know our cause to be just and God/Allah will never abandon us in our struggle to reconstruct a new Sierra Leone.

Though Sankoh and other RUF leaders may have started with political grievances, and the grievances of the people suffering under the APC's extractive institutions may have encouraged them to join the movement early on, the situation quickly changed and spun out of control. The "mission" of the RUF plunged the country into agony, as in the testimony of a teenager from Geoma, in the south of Sierra Leone:

> They gathered some of us … They chose some of our friends and killed them, two of them. These were people whose fathers were the chiefs, and they had soldiers' boots and property in their houses. They were shot, for no other reason than that they were accused of harbouring soldiers. The chiefs were also killed—as part of the government. They chose someone to be the new chief. They were still saying they had come to free us from the APC. After a point, they were not choosing people to kill, just shooting people.

In the first year of the invasion, any intellectual roots that the RUF may have had were completely extinguished. Sankoh executed those who criticized the mounting stream of atrocities. Soon, few voluntarily joined the RUF. Instead they turned to forcible recruitment, particularly of children. Indeed, all sides did this, including the army. If the Sierra Leonean civil war was a crusade to build a better society, in the end it was a children's crusade. The conflict intensified with massacres and massive human rights abuses, including mass rapes and the amputation of hands and ears. When the RUF took over areas, they also engaged in economic exploitation. It was most obvious in the diamond mining areas, where they press-ganged people into diamond mining, but was widespread elsewhere as well.

The RUF wasn't alone in committing atrocities, massacres, and organized forced labor. The government did so as well. Such was the collapse of law and order that it became difficult for people to tell who was a soldier and who was a rebel. Military discipline completely vanished. By the time the war ended in 2001, probably eighty thousand people had died and the whole country had been devastated. Roads, houses, and buildings were entirely destroyed. Today, if you go to Koidu, a major diamond-producing area in the east, you'll still see rows of burned-out houses scarred with bullet holes.

By 1991 the state in Sierra Leone had totally failed. Think of what King Shyaam started with the Bushong (this page–this page): he set up extractive institutions to cement his power and extract the output the rest of society would produce. But even extractive institutions with central authority concentrated in his hands were an improvement over the situation without any law and order, central authority, or property rights that characterized the Lele society on the other side of the river Kasai. Such lack of order and central authority has been the fate of many African nations in recent decades, partly because the process of political centralization was historically delayed in much of sub-Saharan Africa, but also because the vicious circle of extractive institutions reversed any state centralization that existed, paving the way for state failure.

Sierra Leone during her bloody civil war of ten years, from 1991 to 2001, was a typical case of a failed state. It started out as just another country marred by extractive institutions, albeit of a particularly vicious and inefficient type. Countries become failed states not because of their geography or their culture, but because of the legacy of extractive institutions, which concentrate power and wealth in the hands of those controlling the state, opening the way for unrest, strife, and civil war. Extractive institutions also directly contribute to the gradual failing of the state by neglecting investment in the most basic public services, exactly what happened in Sierra Leone.

Extractive institutions that expropriate and impoverish the people and block economic development are quite 18
common in Africa, Asia, and South America. Charles Taylor helped to start the civil war in Sierra Leone while at the same time initiating a savage conflict in Liberia, which led to state failure there, too. The pattern of extractive institutions collapsing into civil war and state failure has happened elsewhere in Africa; for example, in Angola,

Côte d'Ivoire, the Democratic Republic of Congo, Mozambique, Republic of Congo, Somalia, Sudan, and Uganda. Extraction paves the way for conflict, not unlike the conflict that the highly extractive institutions of the Maya city-states generated almost a thousand years ago. Conflict precipitates state failure. So another reason why nations fail today is that their states fail. This, in turn, is a consequence of decades of rule under extractive economic and political institutions.

WHO IS THE STATE?

The cases of Zimbabwe, Somalia, and Sierra Leone, even if typical of poor countries in Africa, and perhaps even some in Asia, seem rather extreme. Surely Latin American countries do not have failed states? Surely their presidents are not brazen enough to win the lottery?

In Colombia, the Andean Mountains gradually merge to the north with a large coastal plain that borders the Caribbean Ocean. Colombians call this the *tierra caliente*, the "hot country," as distinct from the Andean world of the *tierra fria*, the "cold country." For the last fifty years, Colombia has been regarded by most political scientists and governments as a democracy. The United States feels happy to negotiate a potential free trade agreement with the country and pours all kinds of aid into it, particularly military aid. After a short-lived military government, which ended in 1958, elections have been regularly held, even though until 1974 a pact rotated political power and the presidency between the two traditional political parties, the Conservatives and the Liberals. Still, this pact, the National Front, was itself ratified by the Colombian people via a plebiscite, and this all seems democratic enough.

Yet while Colombia has a long history of democratic elections, it does not have inclusive institutions. Instead, its history has been marred by violations of civil liberties, extrajudicial executions, violence against civilians, and civil war. Not the sort of outcomes we expect from a democracy. The civil war in Colombia is different from that in Sierra Leone, where the state and society collapsed and chaos reigned. But it is a civil war nonetheless and one that has caused far more casualties. The military rule of the 1950s was itself partially in response to a civil war known in Spanish simply as La Violencia, or "The Violence." Since that time quite a range of insurgent groups, mostly communist revolutionaries, have plagued the countryside, kidnapping and murdering. To avoid either of these unpleasant options in rural Colombia, you have to pay the *vacuna*, literally "the vaccination," meaning that you have to vaccinate yourself against being murdered or kidnapped by paying off some group of armed thugs each month.

Not all armed groups in Colombia are communists. In 1981 members of the main communist guerrilla group in Colombia, the Fuerzas Armadas Revolucionarias de Colombia (the FARC—the Revolutionary Armed Forces of Colombia) kidnapped a dairy farmer, Jesus Castaño, who lived in a small town called Amalfi in the hot country in the northeastern part of the department of Antioquia. The FARC demanded a ransom amounting to $7,500, a small fortune in rural Colombia. The family raised it by mortgaging the farm, but their father's corpse was found anyway, chained to a tree. Enough was enough for three of Castaño's sons, Carlos, Fidel, and Vicente. They founded a paramilitary group, Los Tangueros, to hunt down members of the FARC and avenge this act. The brothers were good at organizing, and soon their group grew and began to find a common interest with other similar paramilitary groups that had developed from similar causes. Colombians in many areas were suffering at the hands of left-wing guerrillas, and right-wing paramilitaries formed in opposition. Paramilitaries were being used by landowners to defend themselves against the guerrillas, but they were also involved in drug trafficking, extortion, and the kidnapping and murder of citizens.

By 1997 the paramilitaries, under the leadership of the Castaño brothers, had managed to form a national organization for paramilitaries called the Autodefensas Unidas de Colombia (the AUC—United Self-Defense Forces of Colombia). The AUC expanded into large parts of the country, particularly into the hot country, in the departments of Córdoba, Sucre, Magdalena, and César. By 2001 the AUC may have had as many as thirty thousand armed men at its disposal and was organized into different blocks. In Córdoba, the paramilitary Bloque Catatumbo was led by Salvatore Mancuso. As its power continued to grow, the AUC made a strategic decision to get involved in politics. Paramilitaries and politicians courted each other. Several of the leaders of the AUC organized a meeting with prominent politicians in the town of Santa Fé de Ralito in Córdoba. A joint document, a pact, calling for the "refounding of the country" was issued and signed by leading members of the AUC, such as "Jorge 40" (the nickname for Rodrigo Tovar Pupo), Adolfo Paz (a nom de guerre for Diego Fernando "Don Berna" Murillo), and Diego Vecino (real name: Edwar Cobo Téllez), along with politicians, including national senators William Montes and Miguel de la Espriella. By this point the AUC was running large tracts of Colombia, and it was easy for them to fix who got elected in the 2002 elections for the Congress and Senate. For example, in the municipality of San Onofre, in Sucre, the election was arranged by the paramilitary leader Cadena ("chain"). One eyewitness described what happened as follows:

> The trucks sent by Cadena went around the neighborhoods, corregimientos and rural areas of San Onofre picking people up. According to some inhabitants … for the 2002 elections hundreds of peasants were taken to the corregimiento Plan Parejo so they could see the faces of the candidates they had to vote for in the parliamentarian elections: Jairo Merlano for Senate and Muriel Benito Rebollo for Congress.
>
> Cadena put in a bag the names of the members of the municipal council, took out two and said that he would kill them and other people chosen randomly if Muriel did not win.

The threat seems to have worked: each candidate obtained forty thousand votes in the whole of Sucre. It is no surprise that the mayor of San Onofre signed the pact of Santa Fé de Ralito. Probably one-third of the congressmen and senators owed their election in 2002 to paramilitary support, and Map 20, which depicts the areas of Colombia under paramilitary control, shows how widespread their hold was. Salvatore Mancuso himself put it in an interview in the following way:

> 35 percent of the Congress was elected in areas where there were states of the Self-Defense groups, in those states we were the ones collecting taxes, we delivered justice, and we had the military and territorial control of the region and all the people who wanted to go into politics had to come and deal with the political representatives we had there.

It is not difficult to imagine the effect of this extent of paramilitary control of politics and society on economic institutions and public policy. The expansion of the AUC was not a peaceful affair. The group not only fought against the FARC, but also murdered innocent civilians and terrorized and displaced hundreds of thousands of people from their homes. According to the Internal Displacement Monitoring Centre (IDMC) of the Norwegian Refugee Council, in early 2010 around 10 percent of Colombia's population, nearly 4.5 million people, was internally displaced. The paramilitaries also, as Mancuso suggested, took over the government and all its functions, except that the taxes they collected were just expropriation for their own pockets. An extraordinary pact between the paramilitary leader Martín Llanos (real name: Héctor Germán Buitrago) and the mayors of the municipalities of Tauramena, Aguazul, Maní, Villanueva, Monterrey, and Sabanalarga, in the department of Casanare in eastern Colombia, lists the following rules to which the mayors had to adhere by order of the "Paramilitary Peasants of Casanare":

Map 20: Paramilitary presence across Colombia, 1997–2005

9) Give 50 percent of the municipality budget to be managed by the Paramilitary Peasants of Casanare.

10) 10 percent of each and every contract of the

municipality [to be given to the Paramilitary Peasants of Casanare].

11) Mandatory assistance to all the meetings called by the Paramilitary Peasants of Casanare.

12) Inclusion of the Paramilitary Peasants of Casanare in every infrastructure project.

13) Affiliation to the new political party formed by the Paramilitary Peasants of Casanare.

14) Accomplishment of his/hers governance program.

Casanare is not a poor department. On the contrary, it has the highest level of per capita income of any Colombian department, because it has significant oil deposits, just the kind of resources that attract paramilitaries. In fact, once they gained power, the paramilitaries intensified their systematic expropriation of property. Mancuso himself reputedly accumulated $25 million worth of urban and rural property. Estimates of land expropriated in Colombia by paramilitaries are as high as 10 percent of all rural land.

Colombia is not a case of a failed state about to collapse. But it is a state without sufficient centralization and with far-from-complete authority over all its territory. Though the state is able to provide security and public services in large urban areas such as Bogotá and Barranquilla, there are significant parts of the country where it provides few public services and almost no law and order. Instead, alternative groups and people, such as Mancuso, control politics and resources. In parts of the country, economic institutions function quite well, and there are high levels of human capital and entrepreneurial skill; in other parts the institutions are highly extractive, even failing to provide a minimal degree of state authority.

It might be hard to understand how a situation like this can sustain itself for decades, even centuries. But in fact, the situation has a logic of its own, as a type of vicious circle. Violence and the absence of centralized state institutions of this type enter into a symbiotic relationship with politicians running the functional parts of the society. The symbiotic relationship arises because national politicians exploit the lawlessness in peripheral parts of the country, while paramilitary groups are left to their own devices by the national government.

This pattern became particularly apparent in the 2000s. In 2002 the presidential election was won by Álvaro Uribe. Uribe had something in common with the Castaño brothers: his father had been killed by the FARC. Uribe ran a campaign repudiating the attempts of the previous administration to try to make peace with the FARC. In 2002 his vote share was 3 percentage points higher in areas with paramilitaries than without them. In 2006, when he was reelected, his vote share was 11 percentage points higher in such areas. If Mancuso and his partners could deliver the vote for Congress and the Senate, they could do so in presidential elections as well, particularly for a president strongly aligned with their worldview and likely to be lenient on them. As Jairo Angarita, Salvatore Mancuso's deputy and the former leader of the AUC's Sinú and San Jorge blocs, declared in September 2005, he was proud to work for the "reelection of the best president we have ever had."

Once elected, the paramilitary senators and congressmen voted for what Uribe wanted, in particular changing the constitution so that he could be reelected in 2006, which had not been allowed at the time of his first election, in 2002. In exchange, President Uribe delivered a highly lenient law that allowed the paramilitaries to demobilize. Demobilization did not mean the end of paramilitarism, simply its institutionalization in large parts of Colombia and the Colombian state, which the paramilitaries had taken over and were allowed to keep.

In Colombia many aspects of economic and political institutions have become more inclusive over time. But certain major extractive elements remain. Lawlessness and insecure property rights are endemic in large swaths of the country, and this is a consequence of the lack of control by the national state in many parts of the country, and the particular form of lack of state centralization in Colombia. But this state of affairs is not an inevitable outcome. It is itself a consequence of dynamics mirroring the vicious circle: political institutions in Colombia do not generate incentives for politicians to provide public services and law and order in much of the country and do not put enough constraints on them to prevent them from entering into implicit or explicit deals with paramilitaries and thugs.

El Corralito

Argentina was in the grip of an economic crisis in late 2001. For three years, income had been falling, unemployment had been rising, and the country had accumulated a massive international debt. The policies leading to this situation were adopted after 1989 by the government of Carlos Menem, to stop hyperinflation and stabilize the economy. For a time they were successful.

In 1991 Menem tied the Argentine peso to the U.S. dollar. One peso was equal to one dollar by law. There was to be no change in the exchange rate. End of story. Well, almost. To convince people that the government really meant to stick to the law, it persuaded people to open bank accounts in U.S. dollars. Dollars could be used in the shops of the capital city of Buenos Aires and withdrawn from cash machines all over the city. This policy may have helped stabilize the economy, but it had one big drawback. It made Argentine exports very expensive and foreign imports very cheap. Exports dribbled to a halt; imports gushed in. The only way to pay for them was to borrow. It was an unsustainable situation. As more people began worrying about the sustainability of the peso, they put more of their wealth into dollar accounts at banks. After all, if the government ripped up the law and devalued the peso, they would be safe with dollar accounts, right? They were right to be worried about the peso. But they were too optimistic about their dollars.

On December 1, 2001, the government froze all bank accounts, initially for ninety days. Only a small amount of cash was allowed for withdrawal on a weekly basis. First it was 250 pesos, still worth $250; then 300 pesos. But this was allowed to be withdrawn only from peso accounts. Nobody was allowed to withdraw money from their dollar accounts, unless they agreed to convert the dollars into pesos. Nobody wanted to do so. Argentines dubbed this situation El Corralito, “the Little Corral”: depositors were hemmed into a corral like cows, with nowhere to go. In January the devaluation was finally enacted, and instead of there being one peso for one dollar, there were soon four pesos for one dollar. This should have been a vindication of those who thought that they should put their savings in dollars. But it wasn't, because the government then forcibly converted all the dollar bank accounts into pesos, but at the old one-for-one exchange rate. Someone who had had \$1,000 saved suddenly found himself with only \$250. The government had expropriated three-quarters of people's savings.

For economists, Argentina is a perplexing country. To illustrate how difficult it was to understand Argentina, the Nobel Prize–winning economist Simon Kuznets once famously remarked that there were four sorts of countries: developed, underdeveloped, Japan, and Argentina. Kuznets thought so because, around the time of the First World War, Argentina was one of the richest countries in the world. It then began a steady decline relative to the other rich countries in Western Europe and North America, which turned, in the 1970s and '80s, into an absolute decline. On the surface of it, Argentina's economic performance is puzzling, but the reasons for its decline become clearer when looked at through the lens of inclusive and extractive institutions.

It is true that before 1914, Argentina experienced around fifty years of economic growth, but this was a classic case of growth under extractive institutions. Argentina was then ruled by a narrow elite heavily invested in the agricultural export economy. The economy grew by exporting beef, hides, and grain in the middle of a boom in the world prices of these commodities. Like all such experiences of growth under extractive institutions, it involved no creative destruction and no innovation. And it was not sustainable. Around the time of the First World War, mounting political instability and armed revolts induced the Argentine elites to try to broaden the political system, but this led to the mobilization of forces they could not control, and in 1930 came the first military coup. Between then and 1983, Argentina oscillated backward and forward between dictatorship and democracy and between various extractive institutions. There was mass repression under military rule, which peaked in the 1970s with at least nine thousand people and probably far more being illegally executed. Hundreds of thousands were imprisoned and tortured.

During the periods of civilian rule there were elections—a democracy of sorts. But the political system was far from inclusive. Since the rise of Perón in the 1940s, democratic Argentina has been dominated by the political party he created, the Partido Justicialista, usually just called the Perónist Party. The Perónists won elections thanks to a huge political machine, which succeeded by buying votes, dispensing patronage, and engaging in corruption, including government contracts and jobs in exchange for political support. In a sense this was a democracy, but it was not pluralistic. Power was highly concentrated in the Perónist Party, which faced few constraints on what it could do, at least in the period when the military restrained from throwing it from power. As we saw earlier (this page–this page), if the Supreme Court challenged a policy, so much the worse for the Supreme Court.

In the 1940s, Perón had cultivated the labor movement as a political base. When it was weakened by military repression in the 1970s and '80s, his party simply switched to buying votes from others instead. Economic policies and institutions were designed to deliver income to their supporters, not to create a level playing field. When President Menem faced a term limit that kept him from being reelected in the 1990s, it was just more of the same; he could simply rewrite the constitution and get rid of the term limit. As El Corralito shows, even if Argentina has elections and popularly elected governments, the government is quite able to override property rights and expropriate its own citizens with impunity. There is little check on Argentine presidents and political elites, and certainly no pluralism.

What puzzled Kuznets, and no doubt many others who visit Buenos Aires, is that the city seems so different from Lima, Guatemala City, or even Mexico City. You do not see indigenous people, and you do not see the descendants of former slaves. Mostly you see the glorious architecture and buildings put up during the Belle Epoch, the years of growth under extractive institutions. But in Buenos Aires you see only part of Argentina. Menem, for example, was not from Buenos Aires. He was born in Anillaco, in the province of La Rioja, in the mountains far to the northwest of Buenos

Aires, and he served three terms as governor of the province. At the time of the conquest of the Americas by the Spanish, this area of Argentina was an outlying part of the Inca Empire and had a dense population of indigenous people (see Map 1 on this page). The Spanish created *encomiendas* here, and a highly extractive economy developed growing food and breeding mules for the miners in Potosí to the north. In fact, La Rioja was much more like the area of Potosí in Peru and Bolivia than it was like Buenos Aires. In the nineteenth century, La Rioja produced the famous warlord Facundo Quiroga, who ruled the area lawlessly and marched his army on Buenos Aires. The story about the development of Argentine political institutions is a story about how the interior provinces, such as La Rioja, reached agreements with Buenos Aires. These agreements were a truce: the warlords of La Rioja agreed to leave Buenos Aires alone so that it could make money. In return, the Buenos Aires elites gave up on reforming the institutions of "the interior." So Argentina at first appears a world apart from Peru or Bolivia, but it is really not so different once you leave the elegant boulevards of Buenos Aires. That the preferences and the politics of the interior got embedded into Argentine institutions is the reason why the country has experienced a very similar institutional path to those of other extractive Latin American countries.

That elections have not brought either inclusive political or economic institutions is the typical case in Latin America. In Colombia, paramilitaries can fix one-third of national elections. In Venezuela today, as in Argentina, the democratically elected government of Hugo Chávez attacks its opponents, fires them from public-sector jobs, closes down newspapers whose editorials it doesn't like, and expropriates property. In whatever he does, Chávez is much more powerful and less constrained than Sir Robert Walpole was in Britain in the 1720s, when he was unable to condemn John Huntridge under the Black Act (this page–this page). Huntridge would have fared much less well in present-day Venezuela or Argentina.

While the democracy emerging in Latin America is in principle diametrically opposed to elite rule, and in rhetoric and action it tries to redistribute rights and opportunities away from at least a segment of the elite, its roots are firmly based in extractive regimes in two senses. First, inequities persisting for centuries under extractive regimes make voters in newly emerging democracies vote in favor of politicians with extreme policies. It is not that Argentinians are just naïve and think that Juan Perón or the more recent Perónist politicians such as Menem or the Kirchners are selfless and looking out for their interests, or that Venezuelans see their salvation in Chávez. Instead, many Argentinians and Venezuelans recognize that all other politicians and parties have for so long failed to give them voice, to provide them with the most basic public services, such as roads and education, and to protect them from exploitation by local elites. So many Venezuelans today support the policies that Chávez is adopting even if these come with corruption and waste in the same way that many Argentinians supported Perón's policies in the 1940s and 1970s. Second, it is again the underlying extractive institutions that make politics so attractive to, and so biased in favor of, strongmen such as Perón and Chávez, rather than an effective party system producing socially desirable alternatives. Perón, Chávez, and dozens of other strongmen in Latin America are just another facet of the iron law of oligarchy, and as the name suggests, the roots of this iron law lies in the underlying elite-controlled regimes.

THE NEW ABSOLUTISM

In November 2009, the government of North Korea implemented what economists call a currency reform. Severe bouts of inflation are often the reasons for such reforms. In France in January 1960, a currency reform introduced a new franc that was equal to 100 of the existing francs. Old francs continued in circulation and people even quoted prices in them as the change to the new francs was gradually made. Finally, old francs ceased to be legal tender in January 2002, when France introduced the euro. The North Korean reform looked similar on the face of it. Like the French in 1960, the North Korean government decided to take two zeros off the currency. One hundred old wons, the currency of North Korea, were to be worth one new won. Individuals were allowed to come forward to exchange their old currency for the newly printed currency, though this had to be done in one week, rather than forty-two years, as in the French case. Then came the catch: the government announced that no one could convert more than 100,000 won, though it later relaxed this to 500,000. One hundred thousand won was about $40 at the black market exchange rate. In one stroke, the government had wiped out a huge fraction of North Korean citizens' private wealth; we do not know exactly how much, but it is probably greater than that expropriated by the Argentine government in 2002.

The government in North Korea is a communist dictatorship opposed to private property and markets. But it is difficult to control black markets, and black markets make transactions in cash. Of course quite a bit of foreign exchange is involved, particularly Chinese currency, but many transactions use won. The currency reform was designed to punish people who used these markets and, more specifically, to make sure that they did not become too wealthy or powerful enough to threaten the regime. Keeping them poor was safer. Black markets are not the whole story. People in North Korea also keep their savings in wons because there are few banks in Korea, and they are all owned by the government. In effect, the government used the currency reform to expropriate much of people's savings.

Though the government says it regards markets as bad, the North Korean elite rather like what markets can produce for them. The leader, Kim Jong-Il, has a seven-story pleasure palace equipped with a bar, a karaoke machine, and a mini movie theater. The ground floor has an enormous swimming pool with a wave machine, where Kim likes to use a body board fitted with a small motor. When in 2006 the United States placed sanctions on North Korea, it knew how to really hit the regime where it hurt. It made it illegal to export more than sixty luxury items to North Korea, including yachts, water scooters, racing cars, motorcycles, DVD players, and televisions larger than twenty-nine inches. There would be no more silk scarves, designer fountain pens, furs, or leather luggage. These were exactly the items collected by Kim and his Communist Party elites. One scholar used sales figures from the French company

Hennessy to estimate that Kim's annual cognac budget before the sanctions could have been as high as $800,000 a year.

It is impossible to understand many of the poorest regions of the world at the end of the twentieth century without understanding the new absolutism of the twentieth century: communism. Marx's vision was a system that would generate prosperity under more humane conditions and without inequality. Lenin and his Communist Party were inspired by Marx, but the practice could not have been more different from the theory. The Bolshevik Revolution of 1917 was a bloody affair, and there was no humane aspect to it. Equality was not part of the equation, either, since the first thing Lenin and his entourage did was to create a new elite, themselves, at the head of the Bolshevik Party. In doing so, they purged and murdered not only non-communist elements, but also other communists who could have threatened their power. But the real tragedies were yet to come: first with the Civil War, and then under Stalin's collectivization and his all-too-frequent purges, which may have killed as many as forty million people. Russian communism was brutal, repressive, and bloody, but not unique. The economic consequences and the human suffering were quite typical of what happened elsewhere—for example, in Cambodia in the 1970s under the Khmer Rouge, in China, and in North Korea. In all cases communism brought vicious dictatorships and widespread human rights abuses. Beyond the human suffering and carnage, the communist regimes all set up various types of extractive institutions. The economic institutions, with or without markets, were designed to extract resources from the people, and by entirely abhorring property rights, they often created poverty instead of prosperity. In the Soviet case, as we saw in chapter 5, the Communist system at first generated rapid growth, but then faltered and led to stagnation. The consequences were much more devastating in China under Mao, in Cambodia under the Khmer Rouge, and in North Korea, where the Communist economic institutions led to economic collapse and famine.

The Communist economic institutions were in turn supported by extractive political institutions, concentrating all power in the hands of Communist parties and introducing no constraints on the exercise of this power. Though these were different extractive institutions in form, they had similar effects on the livelihoods of the people as the extractive institutions in Zimbabwe and Sierra Leone.

KING COTTON

Cotton accounts for about 45 percent of the exports of Uzbekistan, making it the most important crop since the country established independence at the breakup of the Soviet Union in 1991. Under Soviet communism all farmland in Uzbekistan was under the control of 2,048 state-owned farms. These were broken up and the land distributed after 1991. But that didn't mean farmers could act independently. Cotton was too valuable to the new government of Uzbekistan's first, and so far only, president, Ismail Karimov. Instead, regulations were introduced that determined what farmers could plant and exactly how much they could sell it for. Cotton was a valuable export, and farmers were paid a small fraction of world market prices for their crop, with the government taking the rest. Nobody would have grown cotton at the prices paid, so the government forced them. Every farmer now has to allocate 35 percent of his land to cotton. This caused many problems, difficulties with machinery being one. At the time of independence, about 40 percent of the harvest was picked by combine harvesters. After 1991, not surprisingly, given the incentives that President Karimov's regime created for farmers, they were not willing to buy these or maintain them. Recognizing the problem, Karimov came up with a solution, in fact, a cheaper option than combine harvesters: schoolchildren.

The cotton bolls start to ripen and are ready to be picked in early September, at about the same time that children return to school. Karimov issued orders to local governors to send cotton delivery quotas to schools. In early September the schools are emptied of 2.7 million children (2006 figures). Teachers, instead of being instructors, became labor recruiters. Gulnaz, a mother of two of these children, explained what happens:

> At the beginning of each school year,

> approximately at the beginning of September, the classes in school are suspended, and instead of classes children are sent to the cotton harvest. Nobody asks for the consent of parents. They don't have weekend holidays [during the harvesting season]. If a child is for any reason left at home, his teacher or class curator comes over and denounces the parents. They assign a plan to each child, from 20 to 60 kg per day depending on the child's age. If a child fails to fulfil this plan then next morning he is lambasted in front of the whole class.

The harvest lasts for two months. Rural children lucky enough to be assigned to farms close to home can walk or are bused to work. Children farther away or from urban areas have to sleep in the sheds or storehouses with the machinery and animals. There are no toilets or kitchens. Children have to bring their own food for lunch.

The main beneficiaries from all this forced labor are the political elites, led by President Karimov, the de facto king of all Uzbeki cotton. The schoolchildren are supposedly paid for their labor, but only supposedly. In 2006, when the world price of cotton was around \$1.40 (U.S.) per kilo, the children were paid about \$0.03 for their daily quota of twenty to sixty kilos. Probably 75 percent of the cotton harvest is now picked by children. In the spring, school is closed for compulsory hoeing, weeding, and transplanting.

How did it all come to this? Uzbekistan, like the other Soviet Socialist Republics, was supposed to gain its independence after the collapse of the Soviet Union and develop a market economy and democracy. As in many other Soviet Republics, this is not what happened, however. President Karimov, who began his political career in the Communist Party of the old Soviet Union, rising to the post of first secretary for Uzbekistan at the opportune moment of 1989, just as the Berlin Wall was collapsing, managed to reinvent himself as a nationalist. With the crucial support of the security forces, in December 1991 he won Uzbekistan's first-ever presidential election. After taking power, he cracked down on the independent political opposition. Opponents are now in prison or exile. There is no free media in Uzbekistan, and no nongovernmental organizations are allowed. The apogee of the intensifying repression came in 2005, when possibly 750, maybe more, demonstrators were murdered by the police and army in Andijon.

Using this command of the security forces and total control of the media, Karimov first extended his presidential term for five years, through a referendum, and then won reelection for a new seven-year term in 2000, with 91.2 percent of the vote. His only opponent declared that he had voted for Karimov! In his 2007 reelection, widely regarded as fraudulent, he won 88 percent of the vote. Elections in Uzbekistan are similar to those that Joseph Stalin used to organize in the heyday of the Soviet Union. One in 1937 was famously covered by *New York Times* correspondent Harold Denny, who reproduced a translation from *Pravda*, the newspaper of the Communist Party, which was meant to convey the tension and excitement of Soviet elections:

> Midnight has struck. The twelfth of December, the day of the first general, equal and direct elections to the Supreme Soviet, has ended. The result of the voting is about to be announced.
>
> The commission remains alone in its room. It is quiet, and the lamps are shining solemnly. Amid the general attentive and intense expectation the chairman performs all the necessary formalities before counting of the ballots—checking up by list how many voters there were and how many have voted—and the result is 100 per cent. 100 per cent! What election in what country for what candidate has given a 100 per cent response?
>
> The main business starts now. Excitedly the chairman inspects the seals on the boxes. Then the members of the commission inspect them. The seals are intact and are cut off. The boxes are opened.
>
> It is quiet. They sit attentively and seriously,

> these election inspectors and executives.
>
> Now it is time to open the envelopes. Three members of the commission take scissors. The chairman rises. The tellers have their copybooks ready. The first envelope is slit. All eyes are directed to it. The chairman takes out two slips—white [for a candidate for the Soviet of the Union] and blue [for a candidate for the Soviet of Nationalities]—and reads loudly and distinctly, "Comrade Stalin."
>
> Instantly the solemnity is broken. Everybody in the room jumps up and applauds joyously and stormily for the first ballot of the first general secret election under the Stalinist Constitution—a ballot with the name of the Constitution's creator.

This mood would have captured the suspense surrounding the reelections of Karimov, who appears an apt pupil of Stalin when it comes to repression and political control and seems to organize elections that compete with those of Stalin in their surrealism.

Under Karimov, Uzbekistan is a country with very extractive political and economic institutions. And it is poor. Probably one-third of the people live in poverty, and the average annual income is around $1,000. Not all the development indicators are bad. According to World Bank data, school enrollment is 100 percent … well, except possibly during the cotton picking season. Literacy is also very high, though apart from controlling all the media, the regime also bans books and censors the Internet. While most people are paid only a few cents a day to pick cotton, the Karimov family and former communist cadres who reinvented themselves after 1989 as the new economic and political elites of Uzbekistan have become fabulously wealthy.

The family economic interests are run by Karimov's daughter Gulnora, who is expected to succeed her father as president. In a country so untransparent and secretive, nobody knows exactly what the Karimov family controls or how much money they earn, but the experience of the U.S.

company Interspan is indicative of what has happened in the Uzbek economy in the last two decades. Cotton is not the only agricultural crop; parts of the country are ideal for growing tea, and Interspan decided to invest. By 2005 it had taken over 30 percent of the local market, but then it ran into trouble. Gulnora decided that the tea industry looked economically promising. Soon Interspan's local personnel started to be arrested, beaten up, and tortured. It became impossible to operate, and by August 2006 the company had pulled out. Its assets were taken over by the Karimov families' rapidly expanding tea interests, at the time representing 67 percent of the market, up from 2 percent a couple of years earlier.

Uzbekistan in many ways looks like a relic from the past, a forgotten age. It is a country languishing under the absolutism of a single family and the cronies surrounding them, with an economy based on forced labor—in fact, the forced labor of children. Except that it isn't. It's part of the current mosaic of societies failing under extractive institutions, and unfortunately it has many commonalities with other former Soviet Socialist Republics, ranging from Armenia and Azerbaijan to Kyrgyzstan, Tajikistan, and Turkmenistan, and reminds us that even in the twenty-first century, extractive economic and political institutions can take an unashamed atrociously extractive form.

KEEPING THE PLAYING FIELD AT AN ANGLE

The 1990s were a period of reform in Egypt. Since the military coup that removed the monarchy in 1954, Egypt had been run as a quasi-socialist society in which the government played a central role in the economy. Many sectors of the economy were dominated by state-owned enterprises. Over the years, the rhetoric of socialism lapsed, markets opened, and the private sector developed. Yet these were not inclusive markets, but markets controlled by the state and by a handful of businessmen allied with the National Democratic Party (NDP), the political party founded by President Anwar Sadat in 1978. Businessmen became more and more involved with the party, and the party became more and more involved with them under the government of Hosni Mubarak. Mubarak, who became president in 1981 following Anwar Sadat's assassination, ruled with the NDP until being forced from power by popular protests and the military in February 2011, as we discussed in the Preface (this page).

Major businesspeople were appointed to key government posts in areas closely related to their economic interests. Rasheed Mohamed Rasheed, former president of Unilever AMET (Africa, Middle East, and Turkey), became minister of foreign trade and industry; Mohamed Zoheir Wahid Garana, the owner and managing director of Garana Travel Company, one of the largest in Egypt, became minister of tourism; Amin Ahmed Mohamed Osman Abaza, founder of the Nile Cotton Trade Company, the largest cotton-exporting company in Egypt, became minister of agriculture.

In many sectors of the economy, businessmen persuaded the government to restrict entry through state regulation. These sectors included the media, iron and steel, the automotive industry, alcoholic beverages, and cement. Each sector was very concentrated with high entry barriers protecting the politically connected businessmen and firms. Big businessmen close to the regime, such as Ahmed Ezz (iron and steel), the Sawiris family (multimedia, beverages, and telecommunications), and Mohamed Nosseir (beverages and telecommunications) received not only protection from the state but also government contracts and large bank loans without needing to put up collateral. Ahmed Ezz was both the chairman of Ezz Steel, the largest company in the country's steel industry, producing 70 percent of Egypt's steel, and also a high-ranking member of the NDP, the chairman of the People's Assembly Budget and Planning Committee, and a close associate of Gamal Mubarak, one of President Mubarak's sons.

The economic reforms of the 1990s promoted by international financial institutions and economists were aimed at freeing up markets and reducing the role of the state in the economy. A key pillar of such reforms everywhere was the privatization of state-owned assets. Mexican privatization (this page–this page), instead of increasing competition, simply turned state-owned monopolies into privately owned monopolies, in the process enriching politically connected businessmen such as Carlos Slim. Exactly the same thing took place in Egypt. The businesspeople connected to the regime were able to heavily influence implementation of Egypt's privatization program so that it favored the wealthy business elite—or the "whales," as they are known locally. At the time that privatization began, the economy was dominated by thirty-two of these whales.

One was Ahmed Zayat, at the helm of the Luxor Group. In 1996 the government decided to privatize Al Ahram beverages (ABC), which was the monopoly maker of beer in Egypt. A bid came in from a consortium of the Egyptian Finance Company, led by real estate developer Farid Saad, along with the first venture capital company formed in Egypt in 1995. The consortium included Fouad Sultan, former minister of tourism, Mohamed Nosseir, and Mohamed Ragab, another elite businessman. The group was well connected, but not well connected enough. Its bid of 400 million Egyptian pounds was turned down as too low. Zayat was better connected. He didn't have the money to purchase ABC, so he came up with a scheme of Carlos Slim–type ingenuity. ABC shares were floated for the first time on the London Stock Exchange, and the Luxor Group acquired 74.9 percent of those shares at 68.5 Egyptian pounds per share. Three months later the shares were then split in two, and the Luxor Group was able to sell all of them at 52.5 pounds each, netting a 36 percent profit, with which Zayat was able to fund the purchase of ABC for 231 million pounds the next month. At the time, ABC was making an annual profit of around 41.3 million Egyptian pounds and had cash reserves of 93 million Egyptian pounds. It was quite a bargain. In 1999 the newly privatized ABC extended its monopoly from beer into wine by buying the privatized national wine monopoly Gianaclis. Gianaclis was a very profitable company, nestling behind a 3,000 percent tariff imposed on imported wines, and it had a 70 percent profit margin on what it sold. In 2002 the monopoly changed hands again when Zayat sold ABC to Heineken for 1.3 billion Egyptian pounds. A 563 percent profit in five years.

Mohamed Nosseir hadn't always been on the losing side. In 1993 he purchased the privatized El Nasr Bottling Company, which had the monopoly rights to bottle and sell Coca-Cola in Egypt. Nosseir's relations with the then-minister of the public business sector, Atef Ebeid, allowed him to make the purchase with little competition. Nosseir then sold the company after two years for more than three times the acquisition price. Another example was the move in the late 1990s to involve the private sector in the state cinema industry. Again political connections implied that only two families were allowed to bid for and operate the cinemas—one of whom was the Sawiris family.

Egypt today is a poor nation—not as poor as most countries to the south, in sub-Saharan Africa, but still one where around 40 percent of the population is very poor and lives on less than two dollars a day. Ironically, as we saw earlier (this page–this page), in the nineteenth century Egypt was the site of an initially successful attempt at institutional change and economic modernization under Muhammad Ali, who did generate a period of extractive economic growth before it was effectively annexed to the British Empire. From the British colonial period a set of extractive institutions emerged, and were continued by the military after 1954. There was some economic growth and investment in education, but the majority of the population had few economic opportunities, while the new elite could benefit from their connections to the government.

These extractive economic institutions were again supported by extractive political institutions. President Mubarak planned to begin a political dynasty, grooming his son Gamal to replace him. His plan was cut short only by the collapse of his extractive regime in early 2011 in the face of widespread unrest and demonstrations during the so-called Arab Spring. During the period when Nasser was president, there were some inclusive aspects of economic institutions, and the state did open up the education system and provide some opportunities that the previous regime of King Farouk had not. But this was an example of an unstable combination of extractive political institutions with some inclusivity of economic institutions.

The inevitable outcome, which came during Mubarak's reign, was that economic institutions became more extractive, reflecting the distribution of political power in society. In some sense the Arab Spring was a reaction to this. This was true not just in Egypt but also in Tunisia. Three decades of Tunisian growth under extractive political institutions started to go into reverse as President Ben Ali and his family began to prey more and more on the economy.

WHY NATIONS FAIL

Nations fail economically because of extractive institutions. These institutions keep poor countries poor and prevent them from embarking on a path to economic growth. This is true today in Africa, in places such as Zimbabwe and Sierra Leone; in South America, in countries such as Colombia and Argentina; in Asia, in countries such as North Korea and Uzbekistan; and in the Middle East, in nations such as Egypt. There are notable differences among these countries. Some are tropical, some are in temperate latitudes. Some were colonies of Britain; others, of Japan, Spain, and Russia. They have very different histories, languages, and cultures. What they all share is extractive institutions. In all these cases the basis of these institutions is an elite who design economic institutions in order to enrich themselves and perpetuate their power at the expense of the vast majority of people in society. The different histories and social structures of the countries lead to the differences in the nature of the elites and in the details of the extractive institutions. But the reason why these extractive institutions persist is always related to the vicious circle, and the implications of these institutions in terms of impoverishing their citizens are similar—even if their intensity differs.

In Zimbabwe, for example, the elite comprise Robert Mugabe and the core of ZANU-PF, who spearheaded the anticolonial fight in the 1970s. In North Korea, they are the clique around Kim Jong-Il and the Communist Party. In Uzbekistan it is President Islam Karimov, his family, and his reinvented Soviet Union–era cronies. These groups are obviously very different, and these differences, along with the variegated polities and economies they govern, mean that the specific form the extractive institutions take differs. For instance, because North Korea was created by a communist revolution, it takes as its political model the one-party rule of the Communist Party. Though Mugabe did invite the North Korean military into Zimbabwe in the 1980s to massacre his opponents in Matabeleland, such a model for extractive political institutions is not applicable in Zimbabwe. Instead, because of the way he came to power in the anticolonial struggle, Mugabe had to cloak his rule with elections, even if for a while he managed actually to engineer a constitutionally sanctified one-party state.

In contrast, Colombia has had a long history of elections, which emerged historically as a method for sharing power between the Liberal and Conservative parties in the wake of independence from Spain. Not only is the nature of elites different, but their numbers are. In Uzbekistan, Karimov could hijack the remnants of the Soviet state, which gave him a strong apparatus to suppress and murder alternative elites. In Colombia, the lack of authority of the central state in parts of the country has naturally led to much more fragmented elites—in fact, so much so that they sometimes murder one another. Nevertheless, despite these variegated elites and political institutions, these institutions often manage to cement and reproduce the power of the elite that created them. But sometimes the infighting they induce leads to the collapse of the state, as in Sierra Leone.

Just as different histories and structures mean that the identity of elites and the details of extractive political institutions differ, so do the details of the extractive economic institutions that the elites set up. In North Korea, the tools of extraction were again inherited from the communist toolkit: the abolition of private property, state-run farms, and industry.

In Egypt, the situation was quite similar under the avowedly socialist military regime created by Colonel Nasser after 1952. Nasser sided with the Soviet Union in the cold war, expropriating foreign investments, such as the British-owned Suez Canal, and took into public ownership much of the economy. However, the situation in Egypt in the 1950s and '60s was very different from that in North Korea in the 1940s. It was much easier for the North Koreans to create a more radically communist-style economy, since they could expropriate former Japanese assets and build on the economic model of the Chinese Revolution.

In contrast, the Egyptian Revolution was more a coup by a group of military officers. When Egypt changed sides in the cold war and became pro-Western, it was therefore relatively easy, as well as expedient, for the Egyptian military to change from central command to crony capitalism as a method of extraction. Even so, the better economic performance of Egypt compared with North Korea was a consequence of the more limited extractive nature of Egyptian institutions. For one thing, lacking the stifling control of the North Korean Communist Party, the Egyptian regime had to placate its population in a way that the North Korean regime does not. For another, even crony capitalism generates some incentives for investment, at least among those favored by the regime, that are totally absent in North Korea.

Though these details are all important and interesting, the more critical lessons are in the big picture, which reveals that in each of these cases, extractive political institutions have created extractive economic institutions, transferring wealth and power toward the elite.

The intensity of extraction in these different countries obviously varies and has important consequences for prosperity. In Argentina, for example, the constitution and democratic elections do not work well to promote pluralism, but they do function much better than in Colombia. At least the state can claim the monopoly of violence in Argentina. Partly as a consequence, income per capita in Argentina is double that of Colombia. The political institutions of both countries do a much better job of restraining elites than those in Zimbabwe and Sierra Leone, and as a result, Zimbabwe and Sierra Leone are much poorer than Argentina and Colombia.

The vicious circle also implies that even when extractive institutions lead to the collapse of the state, as in Sierra Leone and Zimbabwe, this doesn't put a conclusive end to the rule of these institutions. We have already seen that civil wars and revolutions, while they may occur during critical junctures, do not necessarily lead to institutional change. The events in Sierra Leone since the civil war ended in 2002 vividly illustrate this possibility.

In 2007 in a democratic election, the old party of Siaka Stevens, the APC, returned to power. Though the man who won the presidential election, Ernest Bai Koroma, had no association with the old APC governments, many of his cabinet did. Two of Stevens's sons, Bockarie and Jengo, were even made ambassadors to the United States and Germany. In a sense this is a more volatile version of what we saw happen in Colombia. There the lack of state authority in many parts of the country persists over time because it is in the interests of part of the national political elite to allow it to do so, but the core state institutions are also strong enough to prevent this disorder from turning into complete chaos. In Sierra Leone, partly because of the more extractive nature of economic institutions and partly because of the country's history of highly extractive political institutions, the society has not only suffered economically but has also tipped between complete disorder and some sort of order. Still, the long-run effect is the same: the state all but remains absent, and institutions are extractive.

In all these cases there has been a long history of extractive institutions since at least the nineteenth century. Each country is trapped in a vicious circle. In Colombia and Argentina, they are rooted in the institutions of Spanish colonial rule (this page–this page). Zimbabwe and Sierra Leone originated in British colonial regimes set up in the late nineteenth century. In Sierra Leone, in the absence of white settlers, these regimes built extensively on precolonial extractive structures of political power and intensified them. These structures themselves were the outcome of a long vicious circle that featured lack of political centralization and the disastrous effects of the slave trade. In Zimbabwe, there was much more of a construction of a new form of extractive institutions, because the British South Africa Company created a dual economy. Uzbekistan could take over the extractive institutions of the Soviet Union and, like Egypt, modify them into crony capitalism. The Soviet Union's extractive institutions themselves were in many ways a continuation of those of the tsarist regime, again in a pattern predicated on the iron law of oligarchy. As these various vicious circles played out in different parts of the world over the past 250 years, world inequality emerged, and persists.

The solution to the economic and political failure of nations today is to transform their extractive institutions toward inclusive ones. The vicious circle means that this is not easy. But it is not impossible, and the iron law of oligarchy is not inevitable. Either some preexisting inclusive elements in institutions, or the presence of broad coalitions leading the fight against the existing regime, or just the contingent nature of history, can break vicious circles. Just like the civil war in Sierra Leone, the Glorious Revolution in 1688 was a struggle for power. But it was a struggle of a very different nature than the civil war in Sierra Leone. Conceivably some in Parliament fighting to remove James II in the wake of the Glorious Revolution imagined themselves playing the role of the new absolutist, as Oliver Cromwell did after the English Civil War. But the fact that Parliament was already powerful and made up of a broad coalition consisting of different economic interests and different points of view made the iron law of oligarchy less likely to apply in 1688. And it was helped by the fact that luck was on the side of Parliament against James II. In the next chapter, we will see other examples of countries that have managed to break the mold and transform their institutions for the better, even after a long history of extractive institutions.

"Truly awesome . . . brilliant in its simplicity and power."—Steven Levitt, coauthor of *Freakonomics*

THE ORIGINS OF
POWER, PROSPERITY, AND POVERTY

WHY NATIONS FAIL

DARON ACEMOGLU JAMES A. ROBINSON