activity, they do not provide the full story. Firms benefit from being near each other. So do workers. Economies of scale and the network effects associated with close proximity (known in economics as agglomeration effects), rather than geographic fundamentals, often explain why certain cities outperform others. Most importantly, geography on its own cannot account for the *timing* of the Industrial Revolution, the onset of modern, sustained economic growth in the 19th century, or the various reversals of fortune that we observe in the historical record.

Where does this leave us? Is there a role for geography in explaining why the world became rich? Hopefully, this chapter has convinced you that geography has played an important role in determining certain outcomes that differ between societies, but that it cannot explain everything. If it could explain everything, our fate would have been written thousands of years ago with little room for human agency. In the remaining chapters, we will show that human actions have played a significant role in determining the economic trajectories of societies. These decisions range from the most intimate (how many babies to have) to the type of legal and political systems societies have. Yet, even though human actions have played a key role in determining the world's economic distribution, geography likely played some role in these decisions. To some degree, geography has helped shape societies' institutions (the subject of Chapter 3), culture (the subject of Chapter 4), demography (the subject of Chapter 5), and colonization (the subject of Chapter 6). We will keep these interactions in mind as we proceed through the first half of the book.

3

Is It All Just Institutions?

In Afghanistan, justice is mostly provided at the tribal level. Whether you win your case or not depends in part on its merit. But it also depends on who you are, which tribe you belong to, and who is overseeing your case (Murtazashvili, 2016). In the Ottoman Empire, a well-functioning court system existed. But it was biased: it favored men, Muslims, and elites, regardless of the merits of their cases (Kuran and Rubin, 2018). These systems of justice differ considerably from those found in the wealthier parts of the contemporary world. True, the rich can buy better lawyers, and justice is hardly color-blind, especially in the US. But the fact is that the disadvantaged can, and frequently do, win cases when the facts are in their favor. The relative impartiality of courts, in turn, encourages economic activity. When people know they have legal recourse should their partner cheat them, they are more likely to engage in exchange. Might these differences in legal systems have played some role in determining which parts of the world have become rich?

Differences across societies are hardly relegated to the legal sphere. At a higher level, differences in political systems can play an important role in economic decision-making. Where autocrats rule, violence often follows. In North Korea, those on the wrong side of the regime tend not to last long. In Stalin's Soviet Union, anyone remotely expected of having anti-regime sentiments ended up in a gulag or executed (as well as many who had no such sentiments).

Political systems affect more than just violence. They affect whether you have to pay bribes to do business, whether you have the right to sell and use your property as you please, and whether you have the freedom to move when the economic opportunity presents itself. The degree to which a society enables – or restricts – people in this fashion is of first-order importance in determining its economic potential.

These political, legal, religious, and economic organizations are a society's *institutions*. In this chapter, we examine the literature on institutions. We begin by explaining what institutions are and why they impact economic development. We then show why different institutions in different parts of the world have placed societies on different economic trajectories.

What Are Institutions?

Until recently, textbook accounts of economic growth focused on investment in physical capital and technological change as the key determinants of growth. Such an approach is natural for economists interested in creating mathematical models of economic growth. However, it is of limited value in understanding the historical origins of economic growth. There are simply too many examples of societies *not* investing in capital or technology to boil it down to investment decisions. This is as true of the past as it is today. Investment is relatively sparse in places like Afghanistan, Haiti, and Niger. The marginal return to capital in these countries is likely very high. Why, then, do individuals in these societies forgo what is more or less a free lunch? What are the constraints people face in these societies? And how do these constraints evolve over time?

These questions were posed by North and Thomas (1973). They argued that the factors on which economists focused – investment in capital machinery, factories, and schooling – were not independent causes of economic growth. They *were* economic growth. To understand the causes of growth, one has to study the incentives that led individuals in some societies to build factories and invest, to go to school, and to acquire new skills. One must also study why individuals in other societies were not incentivized to do these things. North and his co-authors called the aspects of society that formed these incentives *institutions* and they proposed reorienting the study of economic growth around the study of them.

For North, institutions are the rules of the game. For example, in a game of football (soccer), the rules determine the nature of the game and thus structure the incentives facing the players. If we want to explain the different behavior of players in football compared to a game of rugby, the best explanation may reside not in differences in the individuals playing the two games but in the different rules, and hence incentives, they face.

North's thinking about institutions evolved over the course of his career. Initially, North and Thomas (1973) supposed that institutions had a tendency to evolve towards efficiency. Over time, inefficient institutions would be weeded out and more efficient institutions would be promoted in the same manner as market competition weeds out less efficient firms in favor of their more efficient competitors.

Later, North (1981) came to the view that there is no process analogous to the competitive market process to "select" the most efficient institutions. The incentives in the political sphere differ from those in the marketplace. Hence, inefficient institutions can persist for decades or even centuries. Such inefficient institutions can be a source of lasting poverty. They may even be a leading cause of why some countries are rich and others are poor.

Building on North's insights, Greif (2006) developed an alternative definition of institutions. His framework incorporates the critical role cultural beliefs play in enforcing, and indeed constituting, institutions. For Greif, institutions are

not only the "rules of the game," but also comprise the beliefs and social norms that uphold these rules. Beliefs and social norms, like institutions, can be hard to change. This is true even when there are clear economic benefits from doing so. One of Greif's key insights is that cultural beliefs and institutions can reinforce each other. When one strengthens the other, both are all the harder to change. Greif proposes this as a key reason why some economies fail to grow while others prosper. We discuss these insights further in Chapter 4.

One key component of institutions is the degree to which they permit *economic freedom*. The more economic freedom a society has, the more individuals are free to allocate their resources as they see fit. Economic freedom is closely associated with the rule of law. When a society follows the rule of law, laws are applied equally and all types of rights are protected. This of course includes economic rights. Economic freedom is strongly correlated with per capita income (Gwartney, Lawson, and Holcombe, 1999; Gwartney, Lawson, Hall, and Murphy, 2019). Rodrik, Subramanian, and Trebbi (2004) show that raising the degree to which a society follows the rule of law one standard deviation – roughly corresponding to the difference in institutions between Bolivia and South Korea – is associated with a 6.4-fold difference in per capita income (see Figure 3.1). Incidentally, this is about the income gap between Bolivia and South Korea. This finding should not be taken as gospel, however. There are major challenges

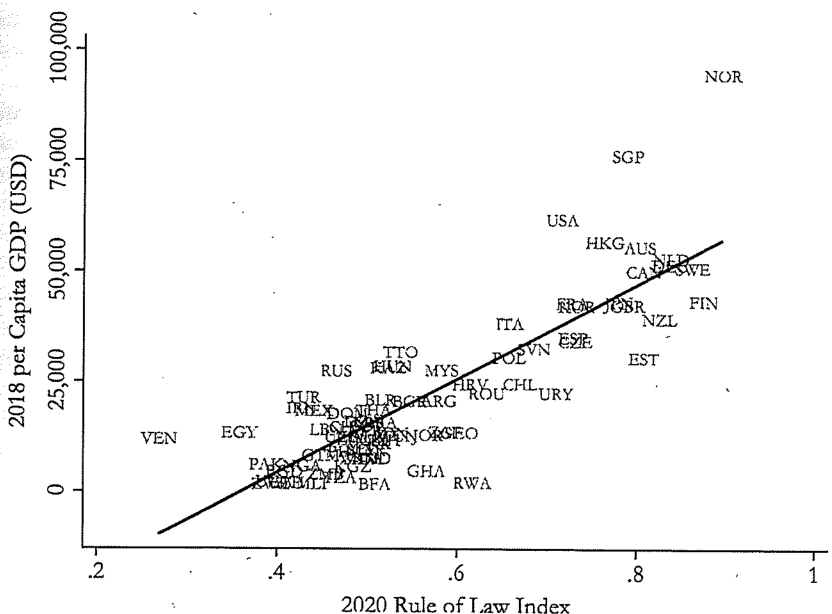

Figure 3.1 Rule of law vs. per capita GDP

Data sources: World Justice Project (2020) for rule of law, Bolt and van Zanden (2020) for per capita GDP.

in estimating the importance of institutions when relying solely on variation across countries. This is because countries differ across so many dimensions it is hard to isolate the specific effects of institutional quality.

For this reason, one of the most persuasive examples illustrating the importance of institutions is the comparison between North and South Korea (Acemoglu, Johnson, and Robinson, 2005*a*). For centuries, North and South Korea were part of a single country, with the same language, culture, and religious traditions. To the extent that there were major regional differences, the north of the country was more industrialized and developed. Then, in 1948, the Communists took over the North. Ever since the war that followed, the economic story between the two nations has been one of divergence. The contrast between the prosperity of the South and the poverty of the North reveals the importance of their different institutions: market-based in the South versus Communist in the North. These differences are immediately visible in Figure 3.2, which is a photograph taken from space of the two countries at night. Night lights are a measure of economic prosperity because they reveal economic activity and electrification. One can clearly see the South Korean border and its many economic hubs. North Korea is almost entirely pitch black.

We rarely have examples as clean-cut as North and South Korea, however. In their absence, institutional arguments can be difficult to test. One critique is that "good institutions" is simply a label for all things a particular author approves of (Clark, 2007, pp. 145–65). The problem is that a label of approbation has little explanatory value. It cannot be operationalized or used to discriminate between points of view.

There are two ways around this problem. One is that pioneered by Greif and his co-authors. Greif developed carefully specified theoretical models of how specific economic institutions functioned. These models have two virtues:

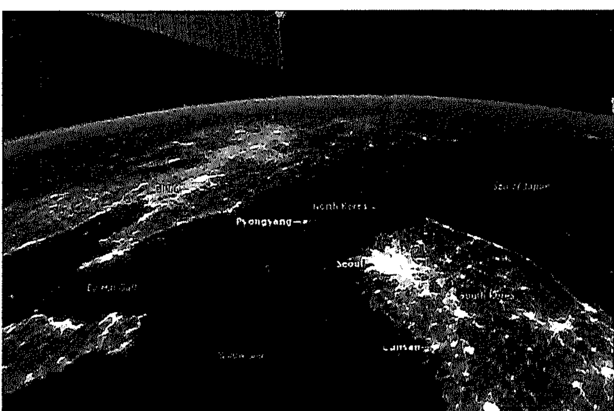

Figure 3.2 Night lights on the Korean Peninsula

Source: https://visibleearth.nasa.gov/view.php?id=83182.

the assumptions are transparent and they are capable of generating novel predictions.

Another approach relies on studying so-called "natural experiments." The comparison of North and South Korea made by Acemoglu, Johnson, and Robinson (2005*a*) is one such natural experiment. In comparing the two Koreas, factors like culture and history are held constant. Borders often provide a nice way to test the effects of institutions. When borders change, people are immediately subjected to a new institutional environment, although nothing changes about their culture, history, or geography. Yet, finding clean natural experiments is difficult, and there are many questions we wish to answer for which no natural experiment exists.

Property Rights

One of the lessons from England's political and economic rise is that well-functioning property rights matter for economic success. Property rights are perhaps the most basic institution studied by economists. Besley and Ghatak (2010, p. 4526) define a property right as an owner's right to use a good or asset (for either consumption or production), to transfer it, or to contract on the basis of it (for example, by mortgaging it). But why do property rights matter? Why have they intrigued economists for so long? What do they have to do with institutions?

Secure property rights ensure that individuals earn a return on their investments. One widely used measure of property rights is the World Bank Governance Indicators (WBGI). These measure a range of factors relating to rule of law, government quality, and regulation. We briefly discussed the rule of law earlier as being key to economic freedom. According to the WBGI, the rule of law index reflects the "perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence" (Kaufmann, Kraay, and Mastruzzi, 2011). Figure 3.3 shows the variations in rule of law throughout the world in 2017. Another widely used measure for security of property is expropriation risk. This is the risk that the state will deprive, expropriate, nationalize, or confiscate the assets of private business, whether domestic or foreign. Such confiscation can be disastrous for investment. Why would one invest if doing so makes them a target for expropriation?

Restrictions on property rights were a dominant feature of pre-modern European property law. Feudal property rights, in particular, were often impediments to developing markets for land or encouraging investment in infrastructure. In England, for example, the law of entails meant that land had to be passed undivided to one male heir. The purpose of this law was to preserve the land of the nobility, who were expected to possess enough land to supply a

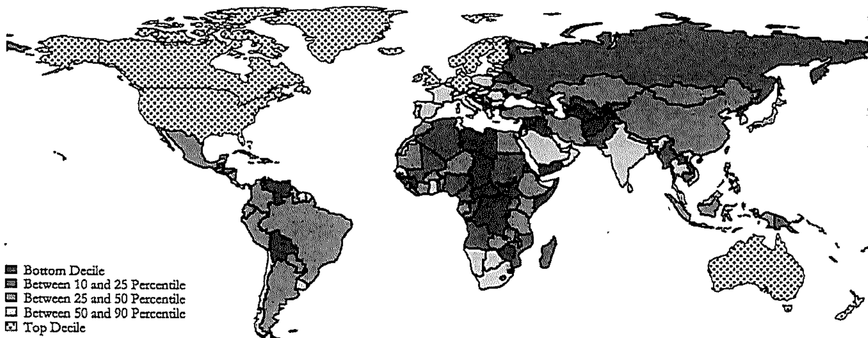

Figure 3.3 Rule of law index, 2017

Data source: Kaufmann, Kraay, and Mastruzzi (2011). Data available at https://info.worldbank.org/governance/wgi/#home.

mounted warrior to serve the king in times of war. The law traded off military security for economic productivity. Land is more valuable when users have the right to sell or divide it. Over the centuries, the military justification for the law of entails diminished, as professional armies replaced knightly hosts. Yet, the law remained the writ of land.

For property to be put to its most efficient use, there must also be a way of rearranging property rights when it is productive to do so. Different systems of property rights have different consequences for economic development. In 17th- and 18th-century England, property rights had to be reordered in order to make the best use of new investment opportunities in areas such as mining (Bogart and Richardson, 2011). For this to happen, there needs to be a way that property rights can be renegotiated and reallocated if it is beneficial to do so. This was made possible only when political institutions also changed.

As Cox (2016) points out, the crucial advantage that England had by the 18th century was that in Parliament it had a centralized forum where these claims could be adjudicated. Local veto players could be, and often were, overruled if the gains to reallocating property rights were large enough. The problem facing rulers who were not constrained by institutions like Parliament is that they could not credibly commit to not using these powers to reallocate property rights in their favor. In France, the monarchy could not be trusted to refrain from abusing this power. Landlords therefore insisted on retaining local veto power on any adjustments in land use. All of this suggests that property rights cannot be viewed in isolation from the legal system as a whole.

Van Bavel, Buringh, and Dijkman (2018) present a fascinating case of weak property rights discouraging investment. They study investments in the major labor-saving capital goods of the late medieval period: water mills, windmills, and cranes. These were expensive to build, but offered an immense return on investment. Because of their cost, they were precisely the type of capital goods

that were only likely to be invested in when property rights were secure. Van Bavel, Buringh, and Dijkman find that, over the period 900–1600, their use increased over time in Western Europe but diminished over time in the Middle East. The decline in Middle Eastern capital investments coincided precisely with a decline in the security of property rights.

It is not just the security of property rights that matters. How rights are assigned is also crucial. In particular, overlapping property rights, in which multiple parties are able to claim rights to the same good, generate holdup problems and impede investment (Lamoreaux, 2011). A holdup problem occurs when one party has an incentive to strategically delay or impede the investment in order to extract a larger share of the resulting profits (Williamson, 1985). For example, in France prior to the Revolution, complex and rigid overlapping property rights prevented landowners from investing in irrigation or in drainage projects (Rosenthal, 1992). If anything, these property rights were "too secure."

It is also possible for property rights to sow the seeds of an economy's demise. Van Bavel (2016) shows that the rise and fall of Abbasid Iraq (750–1258), medieval northern Italy, and the early modern Dutch Republic all followed a similar pattern. In each case, there was a feedback loop in which secure property rights in factor markets allowed underutilized resources to become more productively used. This led to specialization and division of labor, which led to economic growth, which resulted in greater use of factor markets, and so on. However, with factor market growth came political and economic inequality. Those who owned the factors of production gained more political power. They used it to dominate the markets for land, labor, and capital, as well as financial markets, making these markets less free in the process. These vested interests squeezed the little remaining productive power out of the economy, leaving little for the rest of society.

The Legal System

The legal system is the meta-institution that spells out the formal rules of the game and the manner in which they are enforced. Legal systems vary tremendously across societies. Many small-scale societies rely on informal and decentralized legal systems based around ostracism and feuding. These legal systems differ considerably from those of large-scale agrarian states, let alone highly commercialized, market-oriented societies.

Small-scale societies rely on repeated interactions and social sanctions to enforce order. In small-scale, closely knit societies, these mechanisms are powerful in enforcing cooperation. Individuals in these societies expect to interact with each other again in the future. They thus have an incentive to develop a reputation for fair dealing. It is the threat of missing out on future rewards that induces them to refrain from violence or theft in the present (Kandori, 1992; Dixit, 2004).

However, mechanisms dependent on the logic of repeated dealings may be unable to support cooperation in large-scale societies, where people are unlikely to interact consistently. As societies become more complex, therefore, formal legal systems tend to emerge. A formal legal system prescribes rules that outline licit and illicit behavior and specify punishments for those who disobey the law. The first such legal system that we have a record of is the Code of Hammurabi, dated to 1754 BCE (see Figure 3.4).

Legal systems structure the incentives that individuals face. For example, the Code of Hammurabi prescribed numerous harsh punishments. A son who hit his father was to have his hands chopped off, while lovers who conspired to murder a husband were to be impaled. The Code of Hammurabi specified harsher punishment for lower-status individuals and for crimes directed against one's social betters. For instance, "if a man knocks out the teeth of his equal, his teeth shall be knocked out," while knocking out the teeth of a social inferior only resulted in a fine. The Code of Hammurabi thus reinforced existing power structures and a hierarchical social system.

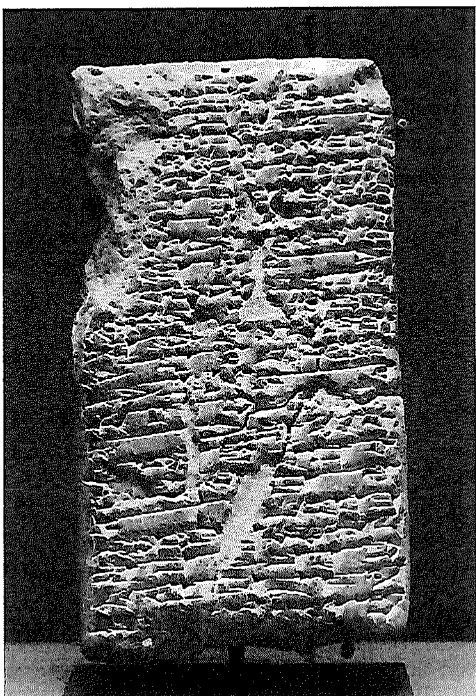

Figure 3.4 The first(?) formal legal system: the Code of Hammurabi

Source: https://upload.wikimedia.org/wikipedia/commons/0/06/Prologue_code_Lipit-Ishtar_AO_5473.jpg.

In other words, the Code of Hammurabi was a set of laws based on *identity rules* rather than impersonal rules (Johnson and Koyama, 2019). The kind of punishment varied according to the social identity of the parties involved. Such identity rules tend to entrench the power of existing elites. There are many historical examples of identity rules. Medieval European Jews faced their own set of laws, including what they could wear and where they could live. Many medieval European cities had sumptuary laws, specifying which clothes individuals could wear based on their social class (Desierto and Koyama, 2020).

In contrast, a society structured around *impersonal rules* is committed to treating individuals the same regardless of their social identity. Impersonal rules are often more consistent with the protection of property than are identity rules. They enable impersonal rather than personal exchange and hence are more likely to promote investment and economic growth.

The distinction between identity rules and impersonal rules is one of the key differences between *rule by law* and the *rule of law*. In a "rule by law" society, laws exist but are not necessarily applied to everyone. The ruling elites tend to be above the law. In these societies, the dictum "rules for thee but not for me" applies. As we noted earlier in this chapter, moving away from such a society to one in which rulers are constrained by the law was key to economic growth (Greif and Rubin, 2021).

Societies that have a strong rule of law tend to be richer (see Figure 3.1). What is it about the rule of law that leads to positive economic outcomes? Fuller (1969) argued that the rule of law requires (1) a concept of legal equality – that is, all individuals from the ruler downwards are equally subject to the law; (2) that laws should be prospective, open, and clear; (3) that laws should be stable over time; (4) that the making of laws should be open and guided by impersonal rules; (5) that the judiciary should be politically independent; (6) that legal institutions such as courts should be accessible to all; and (7) that rules should be general and apply uniformly. These are all conditions that decrease uncertainty, and hence increase investment and incentivize exchange.

Importantly, the rule of law depends on a nexus of different institutions. It depends on a political system that is bound by rules that constrain the power of the executive. It also depends on a standardized legal system in which individuals know what to expect and in which there are clearly defined standards that cannot be subverted through corruption or patronage.

Many scholars have argued that the rule of law played a critical role in the rise of Western Europe. It provided a check on the power of government, guaranteed each individual their own private sphere of non-interference, and offered a platform of institutional stability conducive to long-run economic growth (Hayek, 1960; Cooter, 1997; Weingast, 1997). The idea is that rule of law, rather than rule by men, provides the stability and certainty that enables individuals to truck, barter, and exchange their way to prosperity (Dicey, 1908, pp. 198–9). The question is: how do societies develop impersonal rules? Where does the rule of law come from?

In *Law and Revolution*, Harold Berman (1983) attributed particular significance to the emergence of the Western legal tradition during the Middle Ages. This legal tradition has its origins in ancient Rome. As the Roman economy was commercialized and market-oriented, Roman legal scholars developed a theory of contracts and property (Arruñada, 2016). Roman law was largely lost in Western Europe following the collapse of the Western Roman Empire. During this period, Germanic or customary legal codes dominated. These laws were based on identity rules and better suited to governing small-scale societies rather than providing a framework for impersonal trade. Beginning in the 11th and 12th centuries, however, Roman law was rediscovered and the principles underlying the Roman law of contracts began to be studied by lawyers of canon law in universities like that of Bologna (Fernández-Villaverde, 2022). While customary laws remained in place for centuries, the principles of Roman law would eventually form the basis for both the French and Germanic legal systems.

At roughly the same time, a different legal system was emerging when Germanic and local legal systems were unified into common law during the 12th and 13th centuries. English common law was influenced by Roman law, but only indirectly via canon law. Numerous scholars, including the great legal scholars Pollack and Maitland (1895) and Dicey (1908), saw English common law as it evolved in the Middle Ages as playing a critical role in enshrining and protecting property rights. They reasoned that common law provided a set of stable but adaptive principles that in later centuries enabled more complex organizational forms to emerge.

Building on this tradition, La Porta, de Silanes, Shleifer, and Vishny (1998) argued that the English common law tradition is associated with better protection of property rights, less onerous regulations, and a more favorable environment for markets than systems based on Roman Law (such as the French or German legal systems). In particular, La Porta et al. argued that protection for investors was systematically stronger in common law countries. Common law countries tend to give both shareholders and creditors the strongest rights, regardless of GDP. This is important for financial development because it limits the extent to which corporate insiders can expropriate investors. Other research has found that within a country such as France, which had both civil and common law traditions prior to the Revolution, there is little evidence that civil law was worse for economic development (Le Bris, 2019). Particularly problematic for the legal origins argument is the finding that financial development varies greatly over time, and that in 1913 France's stock market capitalization, as a proportion of the economy, was almost twice that of the US (Rajan and Zingales, 2003).

Nevertheless, the type of law that countries have is one of the important relics of the colonial era. Countries that were colonized by Europeans – which, as we will see in Chapter 6, was much of the world – tended to get their legal institutions (see Figure 3.5). Some speculate that this is why English colonies, steeped as they were in English common law, tended to do better than other

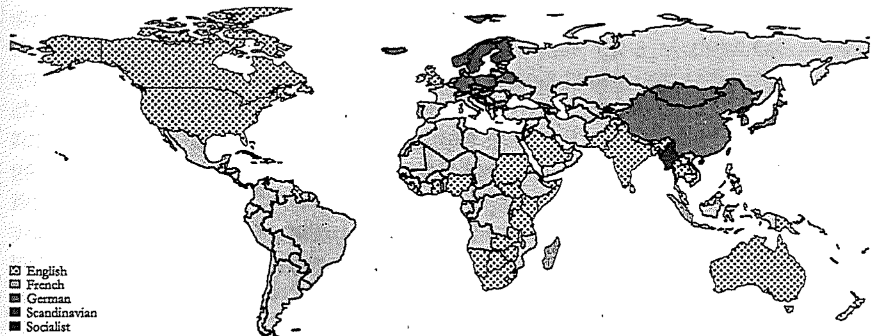

Figure 3.5 Legal origins throughout the world

Data source: Shleifer, de Silanes, and Porta (2008).

European colonies. Other scholarship suggests that what matters is how different legal traditions were transplanted to the colonies (Berkowitz, Pistor, and Richard, 2003). In particular, Oto-Peralías and Romero-Ávila (2014) argue that whether or not the British introduced common law to a colony hinged on initial levels of population density, which determined how costly it would be to enforce a new legal code. We leave the discussion of the long-run effects of colonization to Chapter 6.

What about other legal systems? Non-European systems of law dominated much of the rest of the world in the pre-colonial period. How did they affect economic growth? While there has been much less research undertaken in the social sciences on this issue, Kuran (2011) has highlighted the role that the rigidity of Islamic law played in Middle Eastern economies falling behind those of Europe. We delve into this in more detail in Chapter 4, since it is as much an issue of culture as it is of institutions.

Political Institutions

If institutions form the "rules of the game" that people play, one reason that people around the world play by different rules is the *political institutions* of their society. Political institutions shape how we act. For one, the state generally has something close to a monopoly on violence. It can therefore use force – or the threat of it – to incentivize people to act in certain ways.

The distribution of a society's resources is also highly dependent on political institutions. Resource distribution often reflects the desires of those with political power as well as the desires of those who can keep them in power. In a democracy, this might be "the people." Elsewhere, this may be a select group of warlords. Political institutions are also crucial for encouraging economic activity. Institutions designed to protect rights, limit confiscation, and encourage

investment in public goods are more likely to incentivize investment and economic exchange.

Political institutions vary widely. This variation is one of the leading causes – and consequences – of differences in economic development. Social scientists have come up with numerous ways to measure these institutions, allowing us to compare them over space and time. One of the most widely used metrics, the polity score index (Marshall and Elzinga-Marshall, 2017), measures countries on a range from absolute autocracy to absolute democracy (see Figure 3.6). These data track well with economic development. Europe and North America have high scores and Africa and the Middle East have low scores. This is also true of the democracy index (see Figure 3.7). Of course, there are major outliers: India

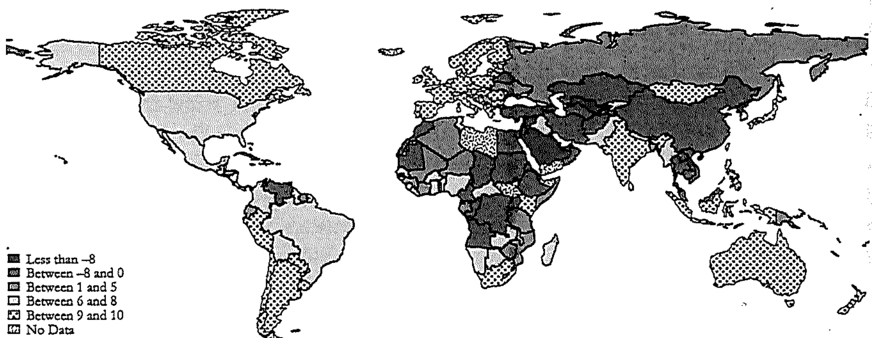

Figure 3.6 Polity score, 2017

Data source: Marshall and Elzinga-Marshall (2017). Note: polity scores range from -10 (strongly autocratic) to 10 (strongly democratic).

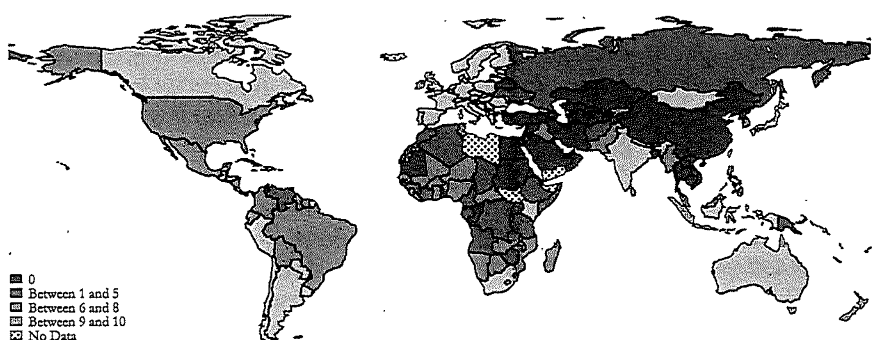

Figure 3.7 Democracy score, 2017

Data source: Marshall and Elzinga-Marshall (2017). Note: democracy scores range from 0 to 10, with higher scores indicating more competitiveness in executive recruitment, more openness of executive recruitment, more constraint on the chief executive, and more competitiveness of political participation.

and Mongolia score relatively well on these indices, despite not having close to the per capita income of OECD countries. Moreover, these indices capture the autocratic and democratic leanings of a society's institutions. This is just one aspect of how "good" or "bad" political institutions are.

But does the type of government one lives under matter for economic growth? If so, why does it matter? What can governments do – or not do – to facilitate growth?

Institutional economists argue that political institutions matter for growth because they ultimately shape the incentives that exist in society. In an unconstrained autocracy, the rule of law cannot function because it does not apply to the ruler, who can use the legal system to favor his cronies and harm his opponents. Autocratic regimes are often more corrupt than non-autocratic regimes. Liberal democracies – those which secure basic rights and liberties for their citizens – tend to have less corruption. Yet, the evidence for whether (liberal) democracy promotes economic growth, or is itself a product of economic growth, is mixed. There are examples of democracies outperforming autocracies and vice versa. As we discuss in more depth in Chapter 10, China has outperformed India since the 1980s despite similar initial levels of development. The Chinese government is autocratic while India is the world's largest democracy.

Whether a country is democratic or not is not all that matters for economic growth. Whether or not state power is limited is likely more important, as we shall discuss below. Moreover, among nondemocratic states it matters how broad the ruling coalition is and whether it represents the interests of merchants and owners of capital. In what follows, we discuss some of the additional factors that connect political institutions to economic development.

More Equal Rights for All

Which are more important for a society's long-run growth: economic institutions or political institutions? Political institutions determine who has power in society. However, political institutions tend to depend on economic resources and hence past economic institutions. Those who control political institutions can use them to increase their control of economic resources, thereby increasing their political power (Acemoglu and Robinson, 2006). This is one reason why institutions can be a source of lasting poverty. In societies where the political system is in the hands of a narrow elite, this elite can use their power to construct economic institutions that benefit themselves but impoverish the rest of society.

There is therefore a feedback between economic institutions and political institutions. The economic institutions of a society are a function of its political institutions, while political institutions help dictate the economic incentives people face. But in what ways do they do so? Are some political institutions "good" for growth and others "bad"?

Acemoglu and Robinson (2012*b*) argue that whether political institutions are good or bad for growth hinges on whether they are inclusive or extractive. Inclusive institutions create broad-based economic incentives and opportunities. They tend to be characterized by a broad distribution of political power. They bring people from varying socio-economic classes, ethnicities, and cultures to the political bargaining table. This makes it tough for one group to transgress the rights of another. Democratic institutions are often seen as relatively inclusive. In a well-functioning democracy, office holders can be challenged, anyone that meets certain qualifications (age, citizenship) can run, and unpopular officials can be voted out.

On the other end of the spectrum are extractive economic and political institutions. These institutions do not create broad-based incentives. Instead, power is held by a small group of powerful elites who are subject to few constraints from the rest of society. When power is held narrowly or the state is too weak to be effective, the conditions are ripe for some individuals or groups to extract from each other. Long-lasting economic development is rarely achieved in such an environment. One need look no further than modern-day autocracies for examples of extractive political institutions. In North Korea, Eritrea, and Turkmenistan, the political elite live in luxury while much of the population barely gets by. Funds that could be used for the public good line the pockets of the elite. There is little incentive for either domestic or foreign investment, since the proceeds of such investment are insecure. These economies tend to stagnate over time.

But what institutions count as "inclusive" as opposed to "extractive"? One problem with this taxonomy is that institutions that are inclusive for some can also be extractive for others. The Athenian democracy was a very inclusive system for adult male Athenians (Ober, 2018). Yet it behaved in a highly extractive manner to its large slave population – many of whom labored in the silver mines of Laurion in horrific conditions – and to the members of its alliance or empire who had to pay tribute in order to maintain the fleet. Was the Athenian democracy more inclusive and less extractive than later societies that did not practice democracy but also did not have slaves or an empire? A similar insight applies to Republican Florence, which granted its citizens civil and political rights but imposed extractive institutions on neighboring towns and cities (Epstein, 2000). To take a more recent example, since the 17th century, political institutions in North America have been more open and equal – for white males – than in almost any other part of the world. At the same time, even after the abolition of the slavery, the US suppressed the voting (and many other) rights of African-Americans until the 1960s.

Acemoglu and Robinson (2019) introduced the concept of a "narrow corridor to liberty" to resolve this apparent paradox. They acknowledge that political development is a process. Inclusive institutions are made possible by a functioning state – one able to repress powerful interest groups or corrupt elites. But inclusive institutions will also be upended by an overly powerful state. Civil society needs to develop alongside the growth of state power because it is civil society that is capable of limiting state power. When the state becomes too powerful relative to society, a society can be knocked out of the corridor, resulting in despotic rule and extractive institutions. But the state cannot be too weak, either. This is when we get anarchism, tribalism, and weak public good provision. For Acemoglu and Robinson, the former situation describes societies like China, while the latter situation explains the poverty and stagnation of much of sub-Saharan Africa.

A related concept is the distinction between what North, Wallis, and Weingast (2009) call the "natural state" and the "open-access order." In the natural state, which was the only form of political organization more or less anywhere prior to the 18th century, personal relationships form the basis of social organization. People are treated differently based on whom they know and where they stand in the social and economic hierarchy. The law, where it is applied at all, applies differently depending on one's standing. Although growth can happen in the natural state, it is limited. The elite can extract from the non-elite, and there is little investment in any goods that do not directly benefit the elite. Nearly every society in history was structured this way, as are many contemporary societies. Feudal Europe, Shogunate Japan, and practically every European colony had natural state political institutions. Today, many states in Africa, the Middle East, and Central Asia are in the "natural state."

Key to this argument is the concept of *rents*. Economic rents are returns above opportunity cost. Rents can be generated by government policies and restrictions on economic activity that limit entry. The insight of Tullock (1967) was that the cost of rents includes all the resources economic actors expend in pursuit of them. These costs can be very large. The concept of rent-seeking can be used to explain many institutions that were prevalent in the past, including restrictions on interest rates (Ekelund, Hebert, Tollison, Anderson, and Davidson, 1996; Koyama, 2010*a*), the sale of monopolies (Ekelund and Tollison, 1997), and government offices. What North, Wallis, and Weingast (2009) mean when they argue that rents hold a natural state together is that they bind together individuals with the capacity to use violence. Rents make it in their interest to adhere to the existing political order rather than challenge it by resorting to force.

According to North, Wallis, and Weingast (2009), long-run economic development requires a society to transition to the open-access order. This is a type of political system similar to what Acemoglu and Robinson (2012*b*) call "inclusive." But the emphasis for North, Wallis, and Weingast is on the ability of individuals and groups to form long-lasting organizations that are independent of the state. What matters for sustained economic growth is less whether or not a country is a formal democracy, but rather how easy it is to set up a rival political party or establish a large business without getting favors from political insiders. By this measure, North, Wallis, and Weingast argue that only a small number of countries today are truly open-access orders.

The question is: how do societies transition to an open-access order or inclusive institutions? How do they enter what Acemoglu and Robinson (2019) call the "narrow corridor"? This is the key question from a development standpoint. Noting which institutions are necessary is one thing. Figuring out how we get there is another. For Acemoglu and Robinson (2012*b*), there are "critical junctures" in a society's past that allow it to make the leap (or not). For North, Wallis, and Weingast (2009), there are a set of "doorstep conditions" a natural state society needs to achieve in order to be able to make the leap. But even then the leap is not assured.

In short, there is no general answer to how a society achieves "good" political institutions. Some societies have done it and others have not. Yet, this does not mean this literature teaches us nothing. Understanding how the societies that took off did so is important for teaching us what *can* work.

Institutions and the Commercial Revolution

Between 1000 and 1300, Europe experienced a period of sustained economic growth. The population grew, per capita income increased, urbanization rates increased dramatically, and the volume of trade expanded many times over. This was not modern economic growth, as we described in Chapter 1, because it eventually petered out. But it was growth nonetheless. For some economic historians, it was this growth that set the stage for what was to come. Southern (1970, p. 34) went so far as to suggest that "that moment of self-generating expansion, for which economists now look so anxiously in underdeveloped countries, came to Western Europe in the late eleventh century." Cipolla (1976, p. 139) described the rise of cities in Europe in the 10th through 12th centuries as one of the "turning point[s] in the history of the West – and, for that matter, of the whole world." These statements may be exaggerations. But there is no doubt that this was a period of remarkable economic growth.

The Commercial Revolution was characterized by the revival of long-distance trade. This trade spanned political borders and jurisdictions. Yet, there were no political authorities capable of enforcing contracts between merchants operating in different cities. So how did trade occur? How could a merchant punish someone who cheated him?

Medieval merchants selling in different jurisdictions faced the following problem. Say an English merchant has wool valued at £5 in England. He decides to take the wool to Flanders, where it is valued more. He meets a Flemish merchant who can sell it for £8. If goods and money can be exchanged there and then, the trade can take place at a price between £5 and £8 and both merchants will gain. Trade is mutually beneficial.

But now suppose that there is some delay between either the goods or payment arriving. Let's say the Flemish merchant has to pay on credit since hard specie is in short supply. This delay between purchase and payment gives the Flemish merchant an opportunity to renege on his promise. The English merchant cannot therefore know whether the Flemish merchant will actually deliver the payment. Once the Flemish merchant has the wool, he has the choice of transferring the money to England or not. But why should he? The shadow cast by the future will not be sufficiently strong to enforce cooperation if the two merchants are not expected to regularly trade with each other in the future. Knowing this, the English merchant will never agree to trade unless he can be paid up front and immediately. The potential gains from trade are never realized.

Greif (2000) called this the fundamental problem of exchange. Trade is limited to cases where either a spot exchange can take place or merchants trust each other enough to trade on credit. This insight is generalizable: in the absence of this trust, trade is extremely costly. A lack of information about the quality of goods could generate similar problems. Manufactured goods might have suffered from defects that were hard to detect on purchase but only became apparent later on. Purchasers alone bore the consequences of defective purchases, because people who purchased defective merchandise had few remedies (Richardson, 2008). Merchants had an incentive to pass on defective goods to traders they met anonymously and to retain quality products for personal exchange.

Greif used game theory to explain how merchant communities in the Middle Ages attempted to resolve this fundamental problem of exchange. First, he considered the Maghribi traders – Jewish traders in Muslim North Africa – who were prominent in Mediterranean trade during the 10th and 11th centuries (Greif, 1989, 1993, 2006). Conducting long-distance trade in the absence of a reliable and centralized legal system, these traders faced a version of the problem of exchange. To conduct their business in different parts of the Mediterranean, they needed to hire agents. But what guarantee did a merchant have that the agent would not cheat him?

The key to the Maghribi traders' success was their ability to utilize a multilateral punishment strategy. This is a situation in which multiple parties agree to punish anyone who cheats any individual in the group. The effectiveness of this punishment hinged upon the nature of their close-knit community. Greif provides evidence that the Maghribi traders were able to disseminate information about the behavior of commercial agents among a coalition of merchants.

In the 11th and 12th centuries, the Maghribi traders were displaced from Mediterranean trade by Italian merchants. Greif (2002, 2006) documents that community responsibility enabled these merchants to overcome the problem of exchange. Community responsibility meant that individual merchants were made responsible for the behavior of their peers. The goods of a merchant could be seized simply because another merchant from his town had refused to repay a loan or had cheated another merchant. Under this system, the cost of an individual merchant cheating was borne by the community. This meant that internal community enforcement could be used to deter domestic merchants from

cheating foreigners. The court of the Flemish merchant had an incentive to uphold complaints from the court of the English merchant, which in turn only championed genuine complaints. This allowed borrowers to credibly commit to repaying lenders even though they might never trade again in future.

Greif's models help to explain how trade worked without widespread rule of law. Although these institutions differed over time and place, they had one common element: they incentivized people to follow through with their promises. They thus allowed trade to flourish long before states were capable of providing the rule of law over the entirety of their domains.

Between the State and the Market: Guilds

Not all institutions are formal ones emanating from the state. The rules of the game come from many facets of life. One important set of institutions that played a role in the Commercial Revolution were *guilds*. Various craft and merchant guilds regulated large swaths of the European economy from the late Middle Ages to the Industrial Revolution. Indeed, the story of the medieval economic efflorescence is impossible to tell without some mention of guild activity.

Craft guilds dominated urban life for much of the period between the 12th and 16th centuries. To participate in a trade, one generally had to be a member of a guild. This was true for smiths, tanners, butchers, bakers, brewers, clerks, and entertainers, among many others. Guild membership was restricted, and one had to serve a lengthy apprenticeship before becoming a master. Craft guilds performed a range of different functions, including regulating goods and labor markets. They also aided local political authorities in tax collection. They kept their privileges by allying with local political authorities (Ogilvie, 2019).

Craft guilds have been the subject of considerable controversy. Were guilds rent-seeking institutions that imposed costs on the wider economy? Or did they serve an important economic function? In theory, both are possible. Guilds did restrict entry and hence generate higher profits for members than they would have otherwise obtained. At the same time, it is possible that guilds could have helped to resolve coordination problems, protect merchants from state predation, and resolve market failures.

Richardson (2008) argued that craft guilds helped resolve problems of asymmetric information such as verifying product quality. Guilds had an incentive to keep product quality high since they had a "name brand" which allowed them to sell goods well outside of the city. This is why some goods are still associated with cities. London's guild of pewterers made a particularly high-quality pewter (which we still call London pewter). The wine-makers of Burgundy made a particularly rich and sought-after red wine. The cheese-makers of Parma made a very nice cheese.

Craft guilds served other purposes as well. Epstein (1998) argued that the craft guilds and the apprentice system played a critical role in incentivizing investment in industry-specific human capital. The work done by master guild members required extremely high skill, taking years to learn. Epstein argues that restricting labor via the apprenticeship system was the only way for masters to recoup the costs of transferring skills. Teaching these skills entailed upfront costs with little immediate payoff. Masters needed to recoup costs later to make it worth it. The rents available via the guild system provided these incentives. They could also have incentivized innovation. Monopoly rents largely went to inventors, since not just anyone could pick up a trade. As a result, there were numerous minor improvements made to techniques and capital in the late medieval period. De la Croix, Doepke, and Mokyr (2017) show that these novel techniques slowly spread throughout Europe via journeymen. These master guildsmen would go from town to town and employ the tacit knowledge of "how to do things" gained in their previous employment. In this light, guilds may have encouraged innovation and the diffusion of best-practice techniques.

In contrast, Ogilvie (2019) argues that craft guilds were rent-seeking organizations that did best by their members but often imposed costs on the rest of society. Drawing on two databases of guild activities, she provides evidence that guilds routinely restricted the economic activity of non-guild members, lobbied governments for privileges, and used a combination of fines, fees, confiscations, and occasionally violence to enforce these privileges. They opposed innovations when they threatened guild profits. In this view, the late medieval economy was held back by guild activity. It is of course possible that the truth lies somewhere in between these opposing views.

Merchant guilds, on the other hand, dominated trade throughout most of the late Middle Ages, especially in the free towns of Northern Europe (Ogilvie, 2011). The most famous of these guilds was the German Hansa. Members of the Hansa traded with other members in cities throughout modern-day Germany, Poland, Denmark, Netherlands, Belgium, England, and beyond.

Travel to foreign lands can be dangerous, especially for merchants carrying valuable cargo. Beyond the physical dangers imposed by outlaws, local rulers have an incentive to confiscate the wares of foreign merchants. What's to stop them? To whom could merchants appeal? According to Greif, Milgrom, and Weingast (1994), one purpose of the merchant guild was to serve as a collective voice against rulers who wished to transgress upon their rights. If a merchant had his rights transgressed, the guild would boycott the city. This dried up the tax revenue that rulers received from trade, providing a huge disincentive to transgress merchants' rights. The guild could do this while an individual could not because boycotts require collective action. Any merchant caught trading with a city under boycott – a "scab" in modern parlance – could be kicked out of the guild, forgoing a lifetime of healthy earnings.

Once states became larger and able to protect the rights of more of their citizens, merchant guilds became unnecessary. They imposed a cost on society

through their monopoly privileges (Ogilvie, 2011). In the medieval context, where state and legal capacity was limited, this cost was worth it, since it allowed some degree of trade to occur. But how did states gain the legal capacity to make institutions such as merchant guilds obsolete? It is to this issue that we turn next.

Parliaments and Limited Government

Among the most important changes in European political institutions in the build-up to the Industrial Revolution was the rise of *limited government*. As the name implies, in a limited government, the powers of the ruling elite are constrained. There are checks on authority in each part of government. Such checks help economies grow. They limit the ability of any one group to extract too much. There are no inclusive political institutions or open-access orders without some limits on government.

But how do societies get limited government? Throughout history, most governments were not limited. Small groups of ruling elites held most power, and the rest of society had relatively little capacity to push back. This began to change in Europe during the Commercial Revolution. In the 11th and 12th centuries, self-governing or independent cities emerged in northern Italy, the Low Countries, and then later in the Rhineland and across Germany. The absence of a single strong ruler enabled merchants to rise to political prominence. They implemented reforms that favored their own interests but were also largely favorable to the overall expansion of trade. For this reason, independent city-states were pioneers of the revival of trade during the Commercial Revolution.

The late Middle Ages also saw the rise of *parliaments*. The term "parliament" refers to political assemblies that had the ability to limit the power of the sovereign. They were generally comprised of three types of elites: churchmen, landed nobles, and the urban bourgeoisie. The first medieval parliaments, or *cortes*, emerged in medieval Spain, in León and Castile, before spreading across Europe in the 12th and 13th centuries (van Zanden, Buringh, and Bosker, 2012). Parliaments were important because, at least in theory, they allowed the interests of the broader population to be *represented* in a single body. In the ancient world, democracy always meant direct democracy. In classical Athens, the entire citizen body meeting in the assembly was sovereign. The limitation to this mode of government was that it scaled poorly. Individuals who did not reside in Athens could not attend the assembly. The development of *representative* institutions like parliaments made it possible for individuals living in much more extensive polities to participate, albeit indirectly, in the governing process.

Parliaments emerged out of the councils medieval kings called among their nobility. In England, the Magna Carta, signed by King John (r. 1199–1216) and his barons in 1215, played a crucial role in the formation of medieval parliaments.

The Magna Carta arose in response to the abuses of King John following his defeat by (and the loss of his Norman territories to) the King of France (Koyama, 2016). To understand why this led to the rise of representative institutions, one has to understand the feudal system that John presided over. One way to do this is through the lens of the natural state framework introduced by North, Wallis, and Weingast (2009).

In their framework, discussed earlier in this chapter, political elites restrict entry in areas of the economy in order to generate monopoly rents. These rents are distributed to those with enough coercive power to disrupt the existing power structure. In medieval England, these were the great barons, who had purview over significant territory. In such a setting, states cannot provide rule of law to all precisely because they rely on treating individuals differently in order to maintain political order. Instead of impersonal rules, natural states rely on personal enforcement and on identity rules to govern. This reliance on identity rules meant that individuals received different treatment depending on their religion, social class, place of birth, or residency.

In England circa 1200, the authority of the king stemmed from two sources. First, he was the largest landowner. This was crucial, as most royal revenue came from land. Second, the king was at the top of the feudal system. All other lords in the country swore him fealty. He could apportion estates in the absence of a direct male heir and was endowed with a range of prerogatives, which included the right to hold court and enforce justice. But these prerogatives did not include the right to freely impose taxes on his subjects during times of peace. In theory, he was expected to "live on his own" from the proceeds of the royal demesne.

To meet the costs of war with France, John exploited all of the revenue sources available to a feudal ruler. He sold wardships – the right to take control of the land of minors until they came of age – and sold the right to marry rich heiresses. Feudal rights such as these were traditionally employed to cement alliances among the nobility. But John exploited them for immediate profit. His position as feudal overlord enabled him to exploit his control over the royal courts to sell justice. And he mercilessly taxed England's Jewish community (Koyama, 2010*b*).

Through these abuses, John alienated the powerful lords who made up the ruling coalition of 13th-century England. These lords had the military power to defeat the king if they could coordinate their resistance. In forcing him to agree to the Magna Carta, they got him to acknowledge certain limits to royal authority – particularly the right to levy taxes without consent. They succeeded: the Magna Carta forcefully articulated the position that the king was subject to the law.

During the reign of John's successor, Henry III (r. 1216–72), the term "parliament" emerged to describe meetings between the king and the leading barons. The initial role of Parliament was judicial. It was employed to hear a variety of legal cases, and it soon became important in fiscal matters as well. In the reign

of Edward I (r. 1272–1307), Parliament became crucial in granting the king the right to collect taxes. Importantly, Edward I regularly called Parliaments in which representatives of the urban bourgeoisie attended, in addition to the nobility and members of the clergy. For the rest of the Middle Ages, English kings relied on Parliament to grant them the taxes they needed, particularly in wartime.

During the Middle Ages, representative institutions flourished across Europe. However, they did not all become national institutions like the English Parliament. For example, the *parlement* of Paris was in origin a similar body to the English Parliament and it covered the entirety of the kingdom of France. However, during and after the Hundred Years' War, the French king devolved authority to regional *parlements* such as the those of Toulouse, Rouen, Guyenne and Gascony, Burgundy, and Provence. These *parlements* came to represent local rather than national interests. They therefore precluded the development of an institution strong enough to constrain the French monarch. Similarly, the *cortes* of León and Aragon, although they remained active, were only able to provide local resistance against the Spanish monarchies after the unification of the crowns of Aragon and Castile in 1469.

Unlike today, parliaments were only in session when they were called by the monarch. And monarchs were only likely to call parliaments when they needed something – mainly taxes. In return, parliaments could expect to receive some favorable laws, policies, or rights. Hence, the frequency with which parliaments were called reflects their power to constrain the crown, and it can thus be used as an indicator of the degree of limited governance (van Zanden, Buringh, and Bosker, 2012). Figure 3.8 depicts the frequency of parliamentary activity in several Western European states. One feature is immediately clear: parliaments in Southern Europe became less active over time while parliamentary activity in northwestern Europe rose rapidly in the 16th and 17th centuries. This was just prior to the economic rise of northwestern Europe.

Is this just a coincidence? The timing suggests not. The parliaments of northwestern Europe, especially in England and the Netherlands, became strong around the time when economic growth accelerated. This is precisely what we would expect to see if limited government played a role in forging the modern economy. Although it could be a coincidence, the literature described above suggests that the connection is causal. For this reason, the rise of limited governance is among the most important issues we will return to in Chapter 7, when we discuss the preconditions for the rise of the modern economy.

War and State Finances

How does a society get institutions that provide equal rights for individuals, enforce property rights, and allow for flourishing factor markets? One important factor dictating the path of institutional development has been warfare.

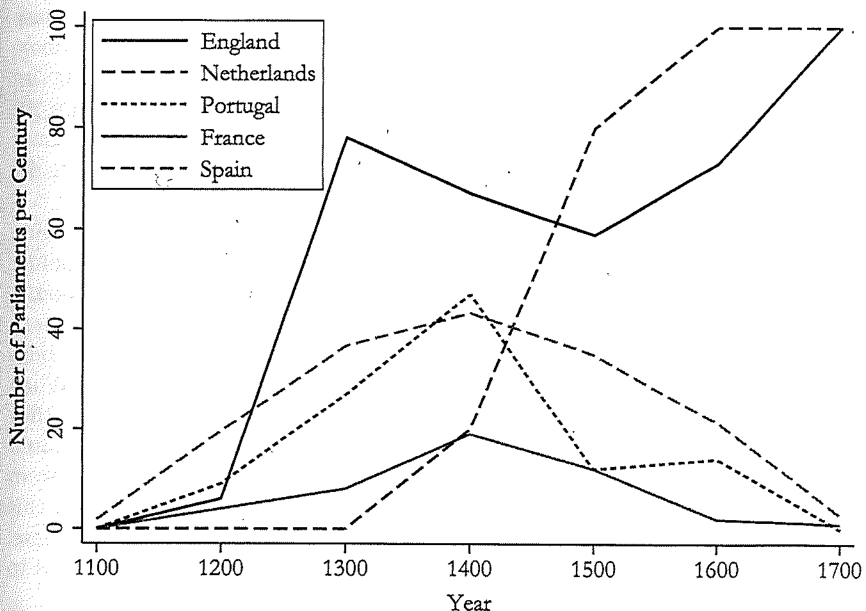

Figure 3.8 Parliamentary activity in Europe, 1100–1800

Data source: van Zanden, Buringh, and Bosker (2012). The values for Spain are the average of the number of parliaments in León and Castile, Catalonia, Aragon, Valencia, and Navarre.

In ancient Athens, the granting of the franchise to all adult male citizens was partly a response to the crucial role ordinary citizens played in maintaining the Athenian navy (Stasavage, 2020). But this was not the norm in European history. Why not? After all, prior to the 19th century, Europeans were almost constantly at war with each other. How did this shape institutional development in Europe? One set of arguments focuses on the simple fact that wars require money. And acquiring money generally requires the capacity to tax. Economists and political scientists call this a state's "fiscal capacity."

How do states acquire fiscal capacity? Raising enough revenue to support a state is a difficult problem to solve. People generally don't like paying taxes unless they feel like they are getting something in return. Even then, a major free-rider problem exists if it is difficult to enforce punishment for tax evasion. Solving this problem requires institutions that align the incentives of the relevant players: rulers, those who collect the taxes, and those who pay the taxes. Once this problem is solved, however, the state generally has to provide something besides just war in order to convince people to pay taxes. These include public goods such as transport networks, education, and infrastructure.

Perhaps more importantly, fiscal capacity does not arise in a vacuum. It is generally associated with other aspects of what scholars call state capacity. This

is the state's ability to implement policy, render justice, protect its citizens, and so on. These are some of the key things a state can provide to promote economic growth. States that solve these problems tend to have better long-run economic prospects. State capacity is different from the size of the state. An expansive state with a large but inefficient public bureaucracy may attempt but fail to achieve many goals. On the other hand, a smaller state may actually achieve its more modest ends.

Over the long run, one way European states built fiscal capacity was by solving the credible commitment problem they faced. Rulers who aspire to absolute power or are unconstrained by either the law or other representative bodies like Parliament cannot make credible promises. Medieval and early modern kings often broke promises to their subjects. For instance, they might raise money for a crusade but use the revenues for their own private purposes. When they broke their promise, it was difficult to hold them to account. After all, they were the king! King Philip II of Spain (r. 1556–98) was so renowned for defaulting on his debts that Drelichman and Voth (2016) call him "the borrower from hell." As we discuss in more detail in Chapter 7, in England it was only once this commitment problem was solved that tax revenues began to seriously increase. This was done by making Parliament a permanent institution (as opposed to something that could be called and dismissed by the monarch).

Another problem for pre-industrial states is that they tended to be fiscally fragmented. As a result, they suffered from local tax free-riding. Why should one region send money to the central government when it believed the other regions would send in enough to cover expenses? Fiscal centralization helped states overcome this free-rider problem. In doing so, it enabled states to increase revenues and often complemented the expansion of markets and the division of labor. This made it easier for responsible governments to follow sound fiscal policies and thus lowered credit risk. However, there was always the chance that rulers would waste the new funds on reckless wars. Thus, it was only the conjunction of limitations on the discretionary authority of government with centralized fiscal systems that enabled states to borrow at low cost (Dincecco, 2009). This allowed a massive take-off in European government revenues between the 16th and 19th centuries (see Figure 3.9).

Why and how did some states in some places and at some times invest in fiscal capacity? In early modern Europe, parliaments played a key role in authorizing and legitimating tax collection. Because states were fiscally fragmented, rulers needed buy-in from other elites in order to collect revenue. Parliaments were the means through which they received this buy-in. Thus far, we have only focused on how parliaments limited the arbitrary power of rulers. But by constraining the arbitrary power of the monarch, parliaments also strengthened rulers by giving them the ability to make credible commitments. This enabled rulers to collect more money and fight longer and more expensive wars. For instance, after throwing off Habsburg rule in 1568, the Dutch Republic embarked upon an eighty-year war with the Spanish Empire followed by a conflict with France

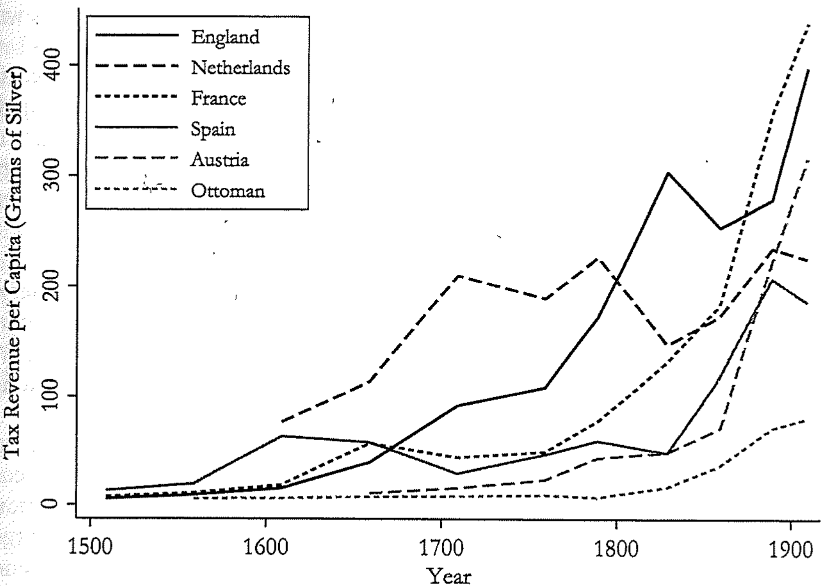

Figure 3.9 Tax revenues per capita for six European powers, 1500–1900
Data source: Karaman and Pamuk (2013).

that would last, on and off, for over forty years. For England, the Glorious Revolution of 1688 marked the start of a "Second Hundred Years' War" with France. But how and why did states form in a manner that they could conduct such long and expensive wars?

This question was asked by Tilly (1975, 1990). His statement "war made the state, and the state made war" succinctly summarizes the matter. War was the main preoccupation of medieval and early modern European states. Before 1800, states spent most of their revenue on warfare or on paying interest on debts occurred in previous wars. Nonetheless, though almost all pre-modern states were preoccupied with war, Tilly (1990) discerned three paths of state formation: a "coercive-intensive" path, a "capital-intensive" path, and a mix between the two.

The coercive-intensive path was taken by Prussia and Russia. These states were built from and for war. The Prussians and Russians did not just tax their populations. They directly mobilized labor and other resources through conscription. In the 17th century, Russia was involved in a war on its western frontier one year out of every two. On its southern and eastern frontiers, war was essentially continuous (Pipes, 1974). As war machines became stronger, they were able to exert force over a greater expanse of people and demand tax revenue from them. This, however, did not lead to positive long-run economic outcomes. There was little incentive for individuals to invest in capital in such a regime.

On the other extreme were states on the "capital-intensive" path. This was taken by many of the city-states of northern Italy, Central Europe, and the Low Countries. These states expanded as their wealth expanded. Capital-intensive states were also often involved in warfare. They generally protected the rights of their citizens, especially those engaged in commerce. They were thus able to collect tax revenue from the economic elite in return for protection. These city-states were able to borrow at relatively low rates precisely because they were run *by* merchants *for* merchants (Stasavage, 2011). Merchants had an interest in maintaining the credit-worthiness of the state, and thus the risk of default was low. This is important: access to credit is a necessary feature of the modern state. Absent natural resource windfalls, it is impossible to think of a modern state functioning well without the capacity to borrow. Yet, as Stasavage (2014) points out, when certain groups dominate the political scene for too long, vested interests eventually undermine further growth.

The late medieval Italian city-states provided a particularly important example of the institutional innovations associated with warfare. Almost continuously at war, Venice and Florence pioneered institutional innovations which enabled them to compete with much larger states and raise ever larger revenues to pay for their mercenary armies. The annual revenue of Florence in the 14th century varied between 250,000 and 350,000 florins, but historians estimate the direct cost of the three-year war between Florence and the Papacy in 1375–8 exceeded 2.5 million florins (Caferro, 2008, p. 177). To meet this shortfall, Italian city-states developed impersonal systems of public debt and permanent systems of taxation (Epstein, 2000). In these cities, the holders of capital were represented in government, ensuring that the promise to repay was credible. Thus, their political institutions allowed city-states to pay lower interest on their debts than territorial monarchies did (see Figure 3.10).

Yet, Tilly (1990) argues that the path to the modern economy did not lie in the purely capital-intensive political institutions of the city-states. Instead, it lay in those states that employed some hybrid model that combined both coercion and capital. England is a prototypical example. These states were able to encourage capital accumulation while also acquiring the capacity to tax large swaths of their population. And those with capital were willing to provide taxes in return for protection of their property rights.

Hoffman (2015) presents a modified version of this thesis. Like Tilly, he focuses on incessant European warfare as a prime mover of European economic success. But instead of emphasizing its role in state formation, he stresses the role that warfare played in encouraging innovation in military technology. These technological advances, especially when combined with gunpowder, gave Europe the upper hand in colonizing the rest of the world. But why were European patterns of warfare different? Hoffman argues that a cultural proclivity to warfare made European military competition more intense than elsewhere.

Yet another impact of warfare was its effect on the location of economic activity. Rosenthal and Wong (2011) argue that European warfare disproportionately

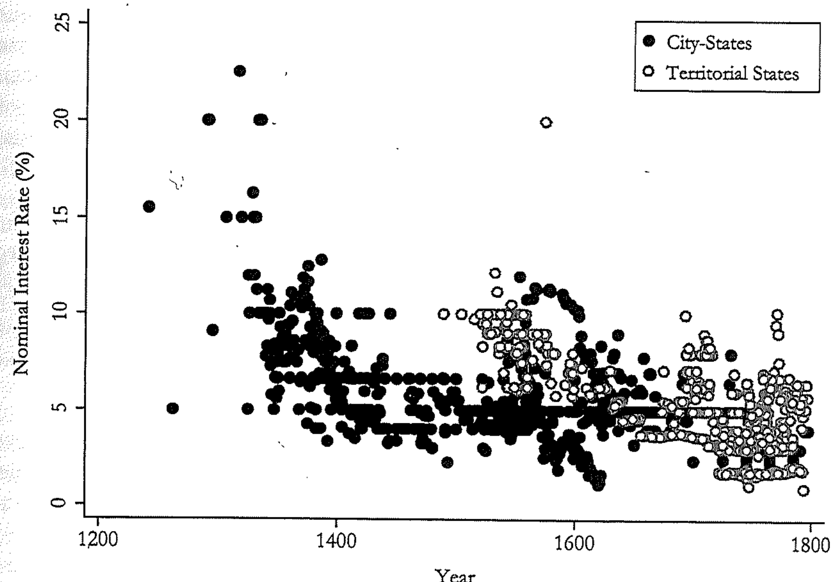

Figure 3.10 Interest rates in city-states and territorial states, 1200–1800
Data source: Stasavage (2016).

affected the countryside because cities tended to be better fortified. This contrasted with China, where less frequent warfare meant that urban and rural locations were equally affected by war. The result was that trade and manufacturing were pushed to urban areas in Europe but not in China. This in turn meant higher urban wages in Europe, which encouraged investment in labor-saving technologies. As we will see in Chapter 8, these were precisely the types of technologies that were key to Europe's industrialization.

In highlighting the role played by warfare in Europe's institutional development, we do not mean to suggest that war was "good" for the economy. Naïve historical accounts that sometimes suggest this are usually mistaken. Warfare is destructive. The direct costs include the destruction caused by war and the opportunity costs of investment in defense, walls, weapons, and armies. There are also harder to ascertain indirect costs that include the effect of warfare on institutional quality or culture.

In any case, warfare was hardly unique to Europe. For instance, conflict was endemic to the Indian sub-continent. Nath (2018, p. 245, quoted in Dincecco, Fenske, Menon, and Mukherjee, 2019) notes that "the Mughals fought their enemies ceaselessly ... war was a constant preoccupation of the Mughal Empire." The Mughal state was also almost entirely devoted to military activity. In Southeast Asia, the Toungoo Empire in Burma warred more or less continuously with Pagan, Ava, and Siam. As we discussed in Chapter 2, there was

also frequent warfare on China's northern frontier between China and various nomadic groups and confederacies. The critical point, however, is that these types of warfare were fundamentally different from interstate warfare between European states.

Warfare in Europe from the late medieval period onwards took place between similarly sized states. As Scheidel (2019, pp. 338–9) observes:

> The fall of Rome ultimately gave rise to multiple states that did not dramatically differ in terms of capabilities (smaller but more cohesive polities balanced less-well-organized larger ones), mobilization intensity (Roman-style levels of conscription did not return until the French Revolution), mode of production (most Europeans were farmers and lived far from the steppe frontier), and religion (Christianity steadily spread into the northern and eastern reaches of the continent while Islam failed to make much headway).

The symmetric nature of European interstate conflict accounts for why warfare was ongoing, with no permanent hegemonic empire emerging. This symmetry was rooted in European geography and in political economy factors that stemmed from this geography, as discussed in Chapter 2.

Other authors push back against this bellicose hypothesis. One counter-argument is that frequent warfare did not promote urbanization, economic development, or the rise of more inclusive institutions in other parts of the world (Centeno, 1997). Another criticism is that when population density is low, warfare may promote slave raiding rather than state-building (Herbst, 2000). We discuss this latter point in the African context in Chapter 6.

Even in Europe, not all states invested as much in fiscal capacity as England did after 1688. To account for why some European states built much more fiscal capacity than others, Gennaioli and Voth (2015) laid out a model in which the incentive to build capacity depended on the chances a state had of defeating its rivals. In this account, interstate competition provided incentives for some states – those for whom the relative payoff to such investments wass high – to standardize their fiscal systems and invest in more capital-intensive forms of war. But other states, which faced higher initial costs of centralizing because they were more heterogeneous, might not find this worthwhile. Nations that did not have a robust military tended to be gobbled up by their rivals, as was Poland-Lithuania.

Chapter Summary

In this chapter, we assessed the role of *institutions* in economic development. Institutions form the "rules of the game" that people play in their day-to-day lives. They come in many forms: political, economic, legal, social, and religious. They form the incentives which shape how people act. Institutions differ across societies and throughout history. They can therefore help explain why different societies have had varying degrees of economic success.

Among the most important institutions a society can have to facilitate growth are those that uphold the rule of law and protect property rights. Whether a society has such institutions is a result of some nexus of their political and legal institutions. Understanding why some societies have had institutions at (or near) this nexus and others have not is an inherently historical question. This chapter has provided insight into why some societies have achieved this and others have not.

Yet, unresolved questions remain. Most importantly, why do institutions work differently in different parts of the world? Democracy is a prime example. Democratic institutions have broadly succeeded in some countries, especially in Europe, in facilitating open exchanges of ideas, empowering a broader set of citizens, and addressing broad-based economic needs. But democratic institutions have failed in other settings, such as many post-Soviet transition states or the post-Arab Spring Middle East. Why do similar institutions work in some settings but not in others? Does culture play some role in the efficacy of institutions? More generally, is a society's culture an important independent determinant of a society's economic success? We turn to these issues in the next chapter.