Chapter 1

▪ ▪ ▪

Economics and Ancient History

> Ignorance is the first requisite of the historian—ignorance, which simplifies and clarifies, which selects and omits, with a placid perfection unattainable by the highest art.
>
> —Lytton Strachey, *Eminent Victorians*

The reputation of the Roman Empire lives on long after the empire itself vanished. Roman literature, Roman archaeological remains, and Roman analogies—particularly now in our time of troubles—confront us at every term. Books like *Are We Rome?* trumpet the analogy, and less extensive allusions are frequent (Murphy 2007; Smil 2010). It often seems as if we are as familiar with the history of ancient Rome as much as of the recent history of the Western world.

While this was true in the late eighteenth century, as witnessed by the writings of our founding fathers, it is no longer so. Most of us do not study Greek and Latin in school, and we do not read the Classics in the original. Most of us know them only by allusion and summary. Classicists and ancient historians by contrast know the ancient languages and read ancient texts, but even they are subject to Strachey's critique. In particular, many accounts of ancient affairs neglect their economic aspects since most ancient historians have only limited training in the dismal science. The application of economic reasoning to ancient history is growing, but more ancient historians than economists are interested in ancient economies.

This book is a contribution to the economic analysis of ancient history from an economic historian who spent most of his academic career writing about modern and early-modern economies. Sometime before the end of the twentieth century, my interest in ancient economies turned from casual to serious.

This book is based on the papers that have resulted from the decade or so of research I have conducted into the economy of ancient Rome, updated and altered to fit into a coherent account. I hope to convince you of five points in this narrative.

First, economics provides useful insights into ancient history. Much of modern economics is devoted to the analysis of modern industrial economies and is not very useful to ancient historians. But the basic elements of economics, still taught in introductory economics classes, provide valuable tools. Supply and demand and comparative advantage allow historians to ask and occasionally answer a variety of questions that have plagued scholars for many years.

Second, ancient Rome had a market economy. There are many references to markets in ancient history, and it does not take much reading to see that they were ubiquitous. Focusing on markets allows us to ask how these markets worked, whether they were helped or hurt by the structure of Roman society, and how far they extended. I argue that markets knit the Roman economy together enough to call it a market economy.

Third, the Pax Romana stimulated Mediterranean trade. Shipping costs over sea were far less than over land before the Industrial Revolution and the advent of the railroad. Extensive Mediterranean trade promoted regional specialization, and comparative advantage worked to raise incomes across the Roman Empire.

Fourth, ordinary Romans lived well, probably better than any other large group—consisting of many millions of people—before the Industrial Revolution. They lived well as a result of extensive markets, comparative advantage, and technological change. True, the Industrial Revolution did not occur in Roman times, and conditions there were not propitious for this momentous change, at least in the form that it took in eighteenth-century Britain. But living conditions were better in the earlier Roman Empire than anywhere else and anytime else before the Industrial Revolution.

Fifth, we are learning more about the Roman economy all the time. Economics helps us ask new questions, and new information is coming to light all the time. Archaeology constantly provides new evidence of economic activity, and new questions suggest reinterpretations of previously known information. This book is a progress report on one part of an ongoing reinterpretation of the Roman economy being undertaken by many historians.

Consider two well-known Romans: Cicero and Trimalchio. They are quite different. One was a historical figure; the other, a fictional one. One lived through the start of the Roman Empire; the other was created a century later. Yet they are together in appearing regularly in the pages of modern ancient historians. It may be interesting to note how they are similar.

Neither of them was a member of a royal household. Often in the study of ancient empires we know only of royal families and their immediate helpers. Even for Rome, they are the most familiar figures, and there is a lot of Roman history that looks only at the emperors and their frequently dysfunctional families. It is uncommon to have abundant evidence of ordinary people long ago, and Roman times are unusual in having records of many ordinary people that have survived for two millennia. This gives us hope that we are not discussing just a tiny royal minority when we analyze records from the Roman Empire. True, most people were farmers and farm laborers who left few records, but even they have left records that have survived.

Both Cicero and Trimalchio were urban residents, in fact residents of Rome itself. Rome was a large city, perhaps a million strong. We need to be careful about that number as with all ancient numbers, but it seems clear that Rome was one of the largest cities that existed before the Industrial Revolution. The existence of this large city, as well as its smaller cousins, tells us that Roman agriculture was efficient enough to feed a lot of nonfarmers. I argue here that this accomplishment was achieved more by long-distance trade than through new technology; I will explain later in this chapter how trade improves incomes. In addition, large cities have their own ecology with lots of urban activities, from crafts to finance. The existence of these people raises questions related to their varied occupations, from how they were paid to whether they had contracts for their work. These questions will engage our attention in several chapters of this book.

Cicero and Trimalchio were both free men and Roman citizens (to the extent that a fictional character can be a citizen). Trimalchio was a freedman, and the *Satyricon* in which he appears satirizes the pretensions of freedmen in the early Roman Empire. Trimalchio was a member of the nouveau riche and subject to the time-honored ritual of being ridiculed for his inability to act like the scion of a respectable, that is, rich household. The ridicule comes from the fear of established people that newcomers will displace them in society, and a freedman contained that threat in ancient Rome. This implies that Roman slavery was far different from slavery in the antebellum United States with which it often is compared. The nature of Roman slavery will be explained further in chapter 6 on the Roman labor force.

Both men had urban occupations. Cicero was a lawyer who pleaded cases in Roman courts. For him to practice this profession there must have been laws and courts in which the laws were applied and tested. The existence of such a legal structure often is used today as a marker for modern societies and for economic growth in less-developed countries. Their existence in ancient Rome indicates that Rome had an important prerequisite for economic growth. The branch of economics that considers these prerequisites is known as the New

Institutional Economics (NIE), as I will explain in this chapter. Trimalchio was a merchant, and he recounts that he had to send out several ships that did not return until he sent our one that did return to his great profit. We learn from this story that there were private merchants, and they were common enough among the literate population for Petronius to assume that his readers knew what he was talking about. More evidence has survived about literary figures who did not like trade than about merchants themselves, so we have to infer their activities from a variety of sources. Chapters 2 and 5 approach this task from different directions. We also learn that being a merchant was a risky occupation, very different from the practice of agriculture. Some ways in which Romans dealt with risk are explained in chapters 5 and 6.

It should clear by now that we need some kind of framework in which to organize all these observations and questions about them. I propose that simple economic tools will help us in this attempt to organize our thoughts, and this chapter will explain a few important economic concepts and their applications to ancient times.

The economy of the early Roman Empire has been an object of study for at least the last century. The discussion has been marked by continuing debate, known sometimes as the primitivist/modern debate and at other times as the Finley debate, following M. I. Finley's famous Sather lectures, *The Ancient Economy*. Finley (1973, 22–23) declared that "ancient society did not have an economic system which was an enormous conglomeration of interdependent markets." He drew implicitly on research by Polanyi (1944, 1977) to oppose the views of Rostovtzeff (1957) within the field of ancient history and those of Fogel and Engerman (1974) in economic history, but he did not explicitly join their conceptual apparatuses. Morris (1999) summarized the debate fueled by Finley's dramatic lectures in his foreword to the twenty-fifth anniversary edition and argued that the controversy is still vigorous today. I hope to clarify the issues in this debate and even resolve the debate for the period of the early Roman Empire.

I argue that the economy of the early Roman Empire was primarily a market economy. The parts of this economy located far from each other were not tied together as tightly as markets often are today, but they still functioned as part of a comprehensive Mediterranean market. There are two reasons why this conclusion is important. First, it brings the description of the Roman economy as a whole into accord with the fragmentary evidence we have about individual market transactions. Second, this synthetic view provides a platform on which to investigate further questions about the origins and eventual demise of the Roman economy and about conditions for the formation and preservation of markets in general.

In his lectures and his subsequent "Further Thoughts," Finley (1999, 27, 182)

called for models of the ancient economy. This is a good approach. But what does it mean to use a model of the ancient economy? A model is an abstract representation of reality. It is simpler than reality because it is created by social and natural scientists who can only conceptualize a few dimensions of reality at a time. Models typically are manipulated in order to reach conclusions, and they have to be simple enough for their formulators to manipulate. With the advent of computers, we can deal with much more complex models than before, but the most useful models often are the simplest.

Most economic models assume the existence of a market economy. The models show how institutions or other economic forces affect prices, quantities, and related variables in one or more industries or, sometimes, in the economy as a whole. The model provides a simplified description of events that can be repeated and discussed, and it allows economists to test *counterfactual* propositions. That is, the economist can ask what would have happened if the institutions or other economic forces had been different than they actually were. The resulting counterfactual history is not an account of events as they happened; it is a conjecture about what would have happened had history been different. The conjecture is conditional on the model. If the model is a poor one, the conjecture will be poor as well. And the conjecture is limited by the model; it can only track the variables in the model in the counterfactual world.

How can we tell whether a model is poor? This is a question that has energized generations of philosophers of science, and I will attempt only the most concrete answer here. A good model fits the observed facts more closely than a poor one. This apparently simple statement has several important components. First, any model depends on the facts behind it. If new data are discovered, models may need to be changed. Stated differently, good models are not made up out of whole cloth; they are distillations of the available data. One advantage of using a model is that it often suggests the need for more data to settle open questions and sets in motion data searches that have proven successful in many fields of economic history. Second, there must be a ranking by which one can tell which model fits the facts more closely than another. When there is an abundance of numerical data, modern statistics and econometrics provide tests that economic historians use. When the data are qualitative, as they generally are for the early Roman Empire, less formal tests have to be used. Third, no model is good in the abstract; it is better or worse than an alternative.

This last point is critical. Economics is a comparative science. The story is told of an economist who meets a colleague while walking across campus. The colleague hails the economist and asks, "How are your children?" The economist responds, "Compared to what?" This response, only slightly exaggerated here, is typical of economists. Economic models are supported by showing that they are superior to another, often called the "null hypothesis." The null

hypothesis of most economics is that there is a well-functioning market, that prices are determined by supply and demand. This is a problem for the study of the Roman economy, because it is precisely this typical null hypothesis that needs to be tested.

I propose to test the hypothesis that there was a market economy in the early Roman Empire in two stages. I argue first that many individual actions and interactions are seen best as market transactions. I then argue that there were enough market transactions to constitute a market economy, that is, an economy where many resources are allocated by prices that are free to move in response to changes in underlying conditions. More technically I argue that markets in the early Roman Empire typically were equilibrated by means of prices.

I begin by presenting the alternatives to which market transactions are to be compared. The logical starting point, as for so much of this literature, is Polanyi. He provided a taxonomy of interactions that has been used widely. He asserted that "the main forms of integration in the human economy are, as we find them, reciprocity, redistribution, and exchange" (Polanyi 1977, 35–36). These forms describe different ways to organize the economic functions of any society. Reciprocity, as the term suggests, is a system in which people aim toward a rough balance between the goods and services they receive and that they give to others. The reciprocal obligations are determined by social obligations and tradition, and they change only slowly. This organization can be formalized, as in Malinowski's Trobriand Islanders, or simply followed with informal or implicit rules. Redistribution is a system in which goods are collected in one hand and distributed by virtue of custom, law, or ad hoc central decision. This system is present in units as small as households, where it is known as householding, as well as in the taxation levied by modern states. The essential characteristic is that a central authority collects and distributes goods and services. Exchange is the familiar economic transaction where people voluntarily exchange one or more goods for other goods or for money. Polanyi's categories appear frequently in books about various aspects of classical antiquity, from Peacock and Williams (1986) on amphorae to Jongman (1988) on Pompeii and Garnsey (1999) on food.

Polanyi's definitions of these different forms of integration are appealing, but imprecise. They suggest three models of interaction; we need to make them precise enough that we can choose between them. Pryor (1977) proposed tests in a study of primitive and peasant economies that can be used to differentiate Polanyi's forms of integration. Pryor distinguished between what he called exchanges and transfers. Exchanges are balanced transactions where goods or services are exchanged for other goods or services of equal value. This is the kind of behavior most often observed in markets. Transfers are

one-way transactions where goods and services are given without a direct return. Grants, tributes, and taxes are all transfers. Pryor excluded "invisibles" from this accounting, so that taxes are considered to be transfers rather than an exchange of goods or money in order to purchase social order or military success. This exclusion is necessary because one can always hypothesize an invisible gain that makes all transactions balanced. In that case, there is no way to discriminate between different forms of behavior.

Pryor subdivided exchanges into those in which the ratio of goods or services exchanged can vary and those in which it cannot. The former may or may not involve money; the latter do not. He termed the former, market exchange; the latter, reciprocal exchange. The use of money is a good index of this distinction, as are changes in the exchange ratio over time. In the presence of money, changes in exchange ratios are expressed as changes in prices. Pryor divided transfers into centric and noncentric ones. Centric transfers are between individuals in a society and an institution or an individual carrying out a societalwide role. In the Roman context, large-scale centric transfers would be those with the Imperial authorities. If the grain to feed Rome were provided by taxes or tribute, this would be a centric transfer. If the grain were obtained by purchasing it with money, then this would be a market exchange.

These categories are observable, that is, they provide boxes into which activities and societies can be placed with confidence. They also correspond closely to Polanyi's forms of economic integration. Polanyi's first form, reciprocity, is composed of Pryor's noncentric transfers and reciprocal exchanges. His second form, redistribution, is accomplished by centric transfers. His third form, exchange, is characterized by what Pryor called market exchange. Pryor's project can be seen as a way to make Polanyi's classification empirically testable, not necessarily reaching Polanyi's conclusion that "price-making markets [are] the exceptional occurrence in history" (Neale 1957, 371).

This tripart schema corresponds also to a division of individual behavior (Temin 1980). People rely on a mixture of behavioral modes, choosing which one to use as a result of internal and external forces. These forces can be represented on two dimensions. One dimension measures internal forces along an index of personal autonomy. The other dimension indexes the rapidity of change in the external environment. When people are less autonomous and change is slow, they typically utilize customary behavior. When change is rapid and personal autonomy is neither very high nor very low, then people use command behavior. When personal autonomy is high and the pace of change is moderate, people employ instrumental behavior, that is, they have explicit goals in mind and choose actions that advance their plans. These different modes of behavior correspond to the three types of organization used in economic life. Customary behavior generally is used for noncentric transfers

and reciprocal exchanges, that is, in reciprocity. Command behavior is typical of centric transfers, that is, redistribution. And instrumental behavior is used in market exchanges.

There consequently are two types of tests we can use to discriminate between the various kinds of integration. Prices are used in market exchanges, but not in noncentric transfers. They may appear in reciprocal exchanges, although they will not vary in response to economic conditions in that context. Variable prices then can be used as markers for the presence of market exchange. Phrased differently, we can infer from the existence of prices that market exchange more closely describes the interaction containing the prices than reciprocity or redistribution. Of course, we will need to make sure that these prices can vary over time to make sure that the prices are not simply stable markers of a noncentric exchange, that is, a specific type of reciprocity.

In addition, people will behave instrumentally in market exchanges, not customarily or by command, since these two modes of behavior are typical of reciprocal and redistributive organizations. Thoughts are observed far less easily than prices, although ancient sources often report the former more volubly than the latter. Nevertheless, we can ask when ancient authors describe their activities if they are describing instrumental, customary, or command behavior. We do so by comparing how well each model of behavior fits the described actions or the imputed thoughts. Phrased differently, we look at the incentives people have to continue their behavior.

The analysis so far tells how to find market exchanges in the early Roman Empire. But how many market exchanges are needed to make a market economy where most resources are allocated by prices that are free to move in response to changes in underlying conditions? There is no general answer to this question, for most economists deal with market economies and have no need to test its very existence. It is necessary to compare Rome with other economies to see the nature and extent of market exchanges in market economies. England and Holland in the seventeenth and eighteenth centuries, shortly before the Industrial Revolution, had economies that everyone agrees were market economies based on agriculture (de Vries and der Woude 1997; Mokyr 2009). Yet even in these market economies, a substantial part of marketed output was allocated by centric transfers rather than by market exchanges. Taxes in Britain were more than 10 percent of national income, and taxes in Holland were more than 40 percent of the income of unskilled laborers, of which about half came from excise taxes on goods consumed by workers. Some market exchanges also had characteristics of reciprocity and customary behavior. Large public works in both countries, primarily to drain land and (in Holland) contain the sea, were paid for by wealthy men, mostly but not exclusively large landowners. Nominal wages stayed constant for many years at a time in the

market economy of early modern England, even though the price of grain fluctuated widely, suggesting that the "labor market" was at least partly an oxymoron; the employment relation often was reciprocal exchange (Phelps Brown and Hopkins 1981; O'Brien 1988; Floud and McCloskey 1997).

Even though there were extensive nonmarket transfers and exchanges, most resources in preindustrial Britain and Holland were allocated by markets. This can be seen by contrasting them with economies that were not primarily market economies. The feudal economy described by Bloch (1961) was a customary economy. Most transactions were made without prices as tenants worked on the lord's land part of the time and as vassals entertained lords to show their fealty. In addition, many transactions were centric transfers as tenants and vassals transferred resources—their labor or the produce of their tenants' labor—to lords in return for protection in the chaotic world of the medieval period. As obligations were written down and then commuted into money payments, the customary feudal economy developed into early modern market economies.

Centrally planned economies in twentieth-century Russia and China were command economies. Russian industries and Chinese farms were compelled to delivery quantities of goods according to a central plan. Prices in the Soviet Union were fixed for long periods of time. Planners expected firms to innovate out of the love of socialism. When that did not work, they set a higher, but still fixed, price for "new goods." Not surprisingly, many old goods were relabeled as new goods, and there was no increase in innovation (Berliner 1976). There were not even prices in the countryside of China until quite recently, as far as we can see, only production quotas. Only now that market reforms are being introduced are farmers selling produce for a price instead of delivering a quota.

There is no formal test to decide which kind of economy we are observing. The classification of these few economies should appear clear, which is why they were chosen. But for an economy about which we have fewer preconceptions we will need to ask several questions. Do the most important commodities, like food and lodging, have prices that move? Are there many transactions in which price appears to play a large part? Do prices move to clear markets? These questions will be answered affirmatively in succeeding chapters.

Before we get to that detail, we need to clarify the nature of what economists call markets and describe some useful economic tools. Markets were prominent in the ancient world; it will ease later discussions to clarify what a market is. The problem is that there is a popular definition and an economic definition, sowing confusion in historical discussions. The popular definition of a market is a place at which trade is conducted. The *Oxford English Dictionary* notes that the Roman forum was designated as a market in medieval writing. Markets now include fish markets, farmers markets, and supermarkets for

food. In the modern world, most trade is directed via stores—distinguished from markets by having uniform, posted prices. Department stores arose in the mid-nineteenth century, and the initial function of prices was to let the store know how much the customer had paid and therefore the amount to be returned, not to inform the customer how much he or she would have to pay.

The stock market is located in a specific place on Wall Street, even though news of stock-market activity is all around us. It is considered to be a paragon of markets by economists because stock prices change the way competitive prices are expected to behave. Current prices embody all information about the stock to date. Future prices depend on future information and cannot be predicted. The best prediction of tomorrow's stock's price therefore is today's price. In mathematical terms, stock prices move as a random walk, that is, tomorrow's price is today's price plus a random (with today's knowledge) movement. I show in chapter 3 that agricultural prices in Hellenistic Babylon moved as a random walk, that is, that they behaved like modern market prices.

Now think of selling a house. We speak of putting our house on the market, but there is no place to take a house—and, of course, no way to take it even if there were such a place. The market in this case is a virtual or disembodied market. It is defined by the nature of the goods or services being sold rather than by where they are sold. This is the key to the economic use of the term, which focuses on the items being sold rather than the method of selling them.

People who anticipate buying or selling a house want to think about its price. To find a suitable range of prices, they look at the sale prices of other, similar houses. But what makes another house similar to this one? It might be location, the prime characteristic of all real estate, so that only local sales are relevant. Local sales might be those on the same street, in the same neighborhood, the same city, or the same country. They might be houses of the same size, or of the same age, or with the same kind of garden. They might even be houses of approximately the same putative value.

This highly ambiguous description is a key to how economists use the term *market*. All houses are in some sense in the same market, but some are closer substitutes for the house being sold than others. Economists argue roughly that houses are in the same market if the price of one affects the price of the other. This is the general idea, but the statement is not quite accurate. On one hand, the price of any single house cannot affect the price of any other in a perfectly competitive market, to be defined shortly, because there are so many similar houses in this kind of market that the sale of any one house has no effect on the market as a whole. On the other, the price of nearby apartments might affect the price of houses. We do not have to be very precise here; we stay with the idea of a market consisting of goods and services that compete with each other. The boundaries of such a market are unclear, and setting them provides employment for economists, but not for ancient historians.

Some historical cases are clear. The Romans dealt with the Chinese over the Silk Road, but travel was hard and long to get from one place to the other. Some goods were exchanged, and some imperfect knowledge of each party about the other went along the road, but the goods that were transported were hugely expensive at their destination, and the information was distorted. It is interesting to know about the Silk Road, but Rome and Han China were not in the same market (Liu 2010). "The two world empires remained hidden to each other in a twilight realm of fable and myth" (Bang 2009, 120).

Conditions on the Silk Road can be illustrated by the writings of Ibn Battuta, a traveler from the fourteenth century CE. He observed that Turkish tribes exported horses to India. The horse sold for about one dinar apiece in Asia, but for more than two hundred dinars in India (Gibb 1986, 145). It is unlikely that a price rise in Asia would affect the price in India. If the price doubled from one to two dinars, the price differential would hardly change. Prices differed by two orders of magnitude between Asia and India, and that shows that the two places were not in a common market. As shown in chapter 2, wheat prices around the Roman Mediterranean were all of the same magnitude, and very unlike the conditions of the Silk Road or fourteenth-century Asia and India.

Going from markets to a market economy adds another level of complexity to the discussion. When Hopkins (1978) described Rome as a slave society, he did not mean that everyone was a slave. Similarly, not every resource in a market economy is allocated through a market. In both cases, the terms indicate that slaves and markets were important, even dominant, institutions. In twentieth-century America—arguably the purest market economy in history—economists have estimated that one-third of economic activity in the United States today takes place within households, that is, in householding (Eisner 1989: 26). The proportion was even higher in the ancient world, but I argue that the economy of the early Roman Empire was a market economy because of the importance and prevalence of market activity (Temin 2001).

The consideration of societies can be made sharper by use of the New Institutional Economics (NIE). This body of thought grows out of a belated recognition by economists that institutions affect economic activity—and are in turn affected by economic pressures. Douglass North (1981; 1990) won a Nobel Prize for making this point over and over again. A paragraph in the earlier of these books says that Rome fell when it could no longer maintain property rights. This paragraph illustrates a weakness of the NIE. No ancient historian can take such a paragraph seriously. Was a decline in property rights a cause or an effect of the "decline of the Roman Empire"? How do you define or measure either of these concepts to find out?

We should not throw the baby out with the bathwater. The New Institutional

Economics helps focus attention on the institutions that govern activities in the ancient world, and it has given rise to some basic hypotheses that may be useful to explore when considering ancient institutions. For example, property rights have been found to promote economic growth by more systematic studies than North's. Acemoglu, Johnson, and Robinson (2001) made this assertion for modern colonies. They argued that colonies differed initially by the healthiness of European colonists. Where the Europeans survived, they brought with them European institutions. Where Europeans died frequently from new (to them) diseases, colonial leaders instituted what are called extractive institutions that did not guarantee private property, condoned bound service of various types, and enriched a small elite at the extent of the general population. Acemoglu, Johnson, and Robinson found that the effects of these initial conditions, indexed by European mortality, explain a substantial amount of income differences in former colonies today. This paper spawned an enormous literature, both because of its ideas and of a new indicator of institutions that avoided the chicken-and-egg problem in North's paragraph. (Economists speak of this chicken-and-egg problem as the identification problem, that is, the problem of identifying which is chicken and which is egg; see chapters 4 and 6.)

Another aid to economic activity is education. Like property rights, it often is hard to determine whether education is a cause or effect of economic growth and prosperity. The same goes for governments that keep corruption at a minimum and for the protection of intellectual rights, that is, the application of property rights to new discoveries. While all of these institutional factors raise similar identification problems, it is useful to set them out separately in order to see what kind of institutions dominated ancient societies. For example, chapters in Scheidel, Morris, and Saller (2007, part VII) describe regions of the Roman Empire, distinguishing them by their initial institutional background and making progress toward solving the identification problem. The western provinces contained few cities before the Roman conquest, and their economies were redirected after integration into the empire. The eastern Mediterranean provinces by contrast built on previous urban patterns, and Roman Egypt developed from its previous well-developed organization and its peculiar geography. Both Cicero and Trimalchio were educated, and they both worked in activities based on the existence of private property.

More difficult to measure but perhaps more important is the culture in which people operate. The Stoic tradition in Rome valued reciprocity in all actions. It made the fulfillment of contractual obligations a matter of personal honor. The effects of laws therefore were amplified by the actions of individuals. Even today, this informal culture promotes the smooth running of economic activities. Verboven (2002, 349) emphasized the role of the "moral economy" in Rome: "While conceptually reciprocity and market exchange may

be opposed they not only coexist in reality but interact continuously. While the market economy profoundly influenced the operation of reciprocity relations and networks, the latter in turn influenced the market system." Reciprocity allowed people to engage in market activities in the expectation that the people they dealt with would fulfill their expectations and act to their mutual benefit. Although the NIE emphasizes the role of laws, the informal networks that underlie these laws are equally important. Wickham (2009, 31) records that these values were preserved by education as the Roman Empire declined, arguing that Roman literary culture held the empire together through shared knowledge and values. Laws, education, and culture are the institutions that make economies work well.

The importance of a shared culture in the modern world has been emphasized by Akerlof, a Nobel laureate for his work on asymmetric information, a key ingredient of the New Institutional Economics. He argued that people act to be connected to their chosen social group. They sometimes are forced into groups by virtue of their gender or race, but often people can choose their groups. Akerlof and Kranton (2010) illustrate this choice by an examination of high school students in the United States. The high school students divide into "jocks" and "nerds," who dress differently, talk differently, and associate largely with their own group. They argue that students have the ability to choose which group to join by considering the costs and benefits of the alternatives. Romans made similar choices when they chose to adhere to Stoic norms. The similarity is abundantly clear when comparing the Roman "economy of friends" and the efforts by the secretary of the U.S. Treasury in 2009, Henry Paulson, to work with his friends in the modern financial system to preserve their position as the global financial crisis spread (Verboven 2002; Paulson 2010).

There was far less information available to ancient people than to people in today's world. In fact, we may know more about the ancient economy than the ancients did, despite the paucity of evidence that has survived two millennia. The NIE focuses our attention on the lack of information and way that people try to deal with it. These concerns run through the following chapters, and in particular, chapter 5 on the grain trade is a contribution to the NIE. I will return to the problems of expensive information and asymmetric information—when one party to a transaction knows more than the other—many times. In order to explain a few basic economic tools, I assume in the rest of this chapter that information is freely available to all.

I also distinguish between personal and anonymous exchanges. The former is negotiated between a buyer and seller, possibly with a broker to facilitate the transaction. Most house purchases and sales, as well as most bazaar transactions, are of this type. Anonymous exchanges involve stated or posted prices that are available to any customers that come by. When we discuss the price

of wheat in ancient Rome, we are referring to anonymous exchanges. Only if wheat had been sold in a bazaar for a different price to each purchaser would it be classified as personal exchange. In the abstractions of modern economics, all exchanges are anonymous.

One of the foundations of economic analysis is the separation of supply and demand. Both terms refer to schedules or curves relating the quantity supplied or demanded as a function of the relevant price. We have evidence of prices in the ancient world, and many of them appeared to vary as a result of changes in supply and demand. Some prices were fixed by administrative fiat of some sort, and some people were not aware of prices. I will discuss how to deal with the former; the latter can be dealt with by interpreting prices as an incentive to buy or sell. Economists speak of prices as shorthand for factors that provide incentives to supply or consume. University professors, for example, perform academic and administrative services for their departments and universities even when there are no explicit prices. The incentives to do so are informal, signifying reciprocity and customary behavior. Nevertheless, if the burden of doing these jobs gets large, professors will do less. If the rewards for these activities increase—say by enhancing chances for promotion or getting a better office—they will do more. This kind of enhanced price is harder to observe than a market price, but it functions in the same way.

We distinguish between supply and demand because it often is the case that different people are behind them. This was true particularly in Roman cities, where food was brought from farms located in the countryside and sometimes far away. It was true within cities when craftsmen made clothing or oil lamps for others to utilize. Robinson Crusoe, alone on his island, was both supplier and demander, but it even makes sense to distinguish him as producer (determining supply) and consumer (determining demand). The distinction helps to clarify the role of different forces affecting the allocation of resources even in such a simple economy (Temin 2012).

The quantity demanded generally increases when the price falls. At lower prices, people can consume more; their resources (in whatever form they take) go farther. In addition, people often want more when the price is lower; they may shift between goods to use more of the cheapest goods and leave some money left over for other things. If prices get much lower, then people may even think of new uses for a commodity. For example, the price of cotton fell dramatically in the Industrial Revolution, leading people to think of putting washable cotton sheets on the beds and cotton curtains on their windows.

These factors will differ in intensity for different goods, and economists use the concept of price elasticity to describe the extent to which the quantity demanded rises when the price declines. Unitary elasticity is defined to be when the proportional increase in the quantity demanded just equals the proportional

decline in the price. Total expenditure stays the same. When the quantity demanded changes less than this, the demand curve is inelastic; when it changes more, demand is elastic. Demand is infinitely elastic if it is so elastic that even a very small change in price will lead to dramatic—even infinite—changes in the quantity demanded. In that case, the very high elasticity of demand keeps the price from varying. That is true in competitive markets, where the actions of any single person have no effect on the price. If the demand for houses, to return to the earlier example, is infinitely elastic, then the decision of any one person to put his or her house on the market will not have any effect on the price.

The quantity supplied generally increases when the price rises. As the price for a product increases, producers make and sell more. They can afford to use more inputs to produce their product, and they may enjoy greater return from the sale. The reasoning implicitly assumes that there are two inputs needed for production. Following a long tradition of classical economists, call them labor and land. If land is fixed, then increasing the number of workers will result in diminishing returns from each worker as more and more of them are added. It is diminishing returns that make the supply curve slope upward.

Supply and demand curves are shown in figure 1.1. Economists normally draw the quantity on the horizontal x-axis and price on the vertical y-axis, and I have followed that convention here. Since the demand curve slopes down and the supply curve slopes up, they generally cross. This is shown in the figure as happening at Q* and P*. What happens if the price is above P*? The quantity of this good that producers want to sell is larger than Q*, while the quantity that people want to buy is less than Q*. Some of the goods produced will remain unsold, and producers will try to get rid of them. The easiest inducement for consumers to buy more is to reduce the price, and the price will fall if it is above P*. Similarly, if the price is below P*, people will want to buy more of the good than producers want to sell. Producers will see that they can sell almost as much as before—each individual producer may expect to sell as much as before—if they raise the price. It will rise as long as the price is below P*. Only when the price equals P* will it stay at that level. We therefore speak of P* and Q* as the equilibrium level of this market.

Why do economists use this framework? The first reason is to understand *changes* in prices or quantities. For example, the production of wheat increased in Roman times. Looking at figure 1.1, we see that the quantity is not likely to differ much from Q* while the supply and demand curves stay the same. If the quantity of wheat produced rose substantially, we then can ask why it rose. We can ask if the supply curve, the demand curve, or both curves shifted to move Q* to a new, higher level. Archaeological debates about innovations in agriculture focus on the supply curve, while thinking about feeding the city of

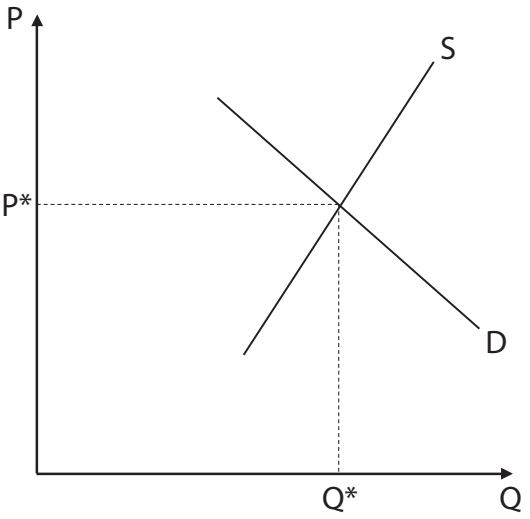

Figure 1.1. Supply and demand

Rome is concentrated on demand. Thinking about supply and demand enables us to integrate these disparate analyses.

Ancient historians may be wary of this framework because it appears to assume its conclusion. The motives imputed to buyers and sellers in the description of equilibrium assumed they were acting instrumentally and maximizing profits or satisfaction. Isn't that the same as assuming a market economy? No. Economic research into a variety of markets in industrial and agrarian economies has found that individuals today almost always want to better themselves. They act instrumentally to do so, although their actions often are constrained by the rules of society that are studied by the NIE so that they can improve their conditions a little, but not very much. The questions for ancient historians are largely whether the rules by which Roman society was organized were conducive or opposed to market activity. Supply and demand are useful even when rules did not seem directed toward economic affairs.

For example, a recent comparison of the supply and demand for wine and wheat in Republican Italy argued that there was not enough demand to support many large estates. It concluded that these markets were essentially competitive, earning limited profits for even large landowners and implying that "we must remove the aristocracy's formation of large, commercial estates from the central role they have long played in reconstructions of the social and economic developments in the middle and late Republic" (Rosenstein 2008, 23).

The forces of supply and demand operate even in reciprocity and redistribution. There are no explicit prices in these cases, but examples abound. The

Roman Senate gradually changed in the second century CE from a group of Italian senators to a group from the provinces (Eck 2000). The separation of supply and demand leads us to ask if this was due to conditions of supply (the scarcity of rich Italians) or instead to demand (a desire to have a wider representation in the Senate). Hopkins (1980) famously tried to estimate the GDP of the Roman Empire to show that the tax burden was light. He clearly was motivated by the presumption that rising taxation would have led to disaffection from the empire, that is, that it would have been harder to maintain the tax rate as its burden increased.

A second reason to use this supply and demand framework is to describe the way in which people made decisions. While the demand for Roman wheat might have risen, each Sicilian or Egyptian farmer would only have known what price—or tax rate—he faced. We have several surviving comments about the prevailing price of wheat, some in normal times and more in unusual ones. The presence of these prices indicates that both farmers and consumers knew what the price was. Since these prices typically were not for individual transactions, they also indicate the presence of anonymous exchanges. We have no way of knowing how widespread this information was, but the quotations suggest strongly that this was general information. It makes sense therefore to see farmers as facing a competitive market in which their output was too small to affect the price. They then made their choices on the basis of what they saw as a fixed market price, just as farmers do today. We can use the tools of a competitive market to analyze the behavior of Roman farmers, even though we do not presume that they—or many more recent farmers—consciously saw themselves in what we now call a competitive market.

A third reason is to examine administrative decisions to see if they were effective or not. For example, wheat was given away in early imperial Rome under the *annona*, the annual storage and distribution of wheat for the city of Rome, for free or a very low price. This price almost certainly was below P^*, the price that would have prevailed if the wheat was bought on some kind of market. In that case, following the analysis of equilibrium, we expect that there should have been pressure from consumers for more free distribution than the authorities planned to give away. The program expanded over time, and this analysis provides one reason why it did.

Two extreme cases are often spoken of by economists. The first one is the infinitely elastic demand curve. As noted already, this is a characteristic of a competitive market, where there are many producers all trying to sell their products in the same market. Transport and transaction costs in the ancient world kept many producers from competing head-to-head with others, but the abstraction gives us a benchmark against which to evaluate what we observe. Given that there were lots of farmers, vineyards, olive presses, makers of oil

lamps, etc., the assumption of a competitive market can be very useful. We can show that condition in figure 1.1 by making the demand curve horizontal.

The second extreme case is when supply is completely inelastic, that is, the supply curve is vertical. A vertical supply curve says that the amount supplied is independent of the price. Paying a high amount or almost nothing will not affect how much is supplied. The most prominent example of this condition is agricultural land. When the Antonine and Justinian plagues struck the ancient world, they decreased the number of farmers, but they had no effect on the quantity of farmland. With fewer farmers seeking to work on the same amount of land, the price of land fell. Since the fall did not affect the quantity of land, we speak of this price as a rent, that is, a price that does not affect the allocation of resources. The more inelastically a good is supplied, the more its price resembles rent.

Rent seeking in the NIE consists of activities designed to capture economic rents. They do not encourage productive activity, but rather contest the returns to inelastically supplied goods and services. A thief, for example, does not produce anything; he steals things. In other words, he changes the ownership of existing resources, which is known as rent seeking. If we undertake activities like locking our houses or hiring body guards to deter thieves or assassins, that also is rent seeking. These preventive activities redirect activities that could be productive into unproductive pursuits; locks and guards are only used if thieves try to steal our possessions or others want to harm us. The existence of rent seeking causes the costs of purchasing to exceed the return from selling it; this discrepancy gives rise to what we call transaction costs, which include both rent seeking and anything else—like transport or information costs—that introduce a gap between the selling and buying price.

The analysis so far has treated an isolated market. There are many markets, and we need to analyze what happens when different markets come into contact. Ricardo presented the theory of comparative advantage two hundred years ago; it has lasted as one of the most convincing argument in economics, showing how trade can benefit both partners. It is a simple theory, but it requires a little background to be understood. The theory of comparative advantage is so important that it has given rise to its own branch of economics: international economics. I will use the language of international economics here, talking of countries and regions trading with each other, but the analysis is only an extension of the supply-and-demand analysis already covered.

Every country has what economists call a production possibility frontier, or PPF. The PPF shows how much of any one good or service can be produced, given how much of the other goods and services are being made. This relationship is best seen in two dimensions, assuming that a country makes only two products. Let us call them wine and wheat. If we put wine on the vertical axis

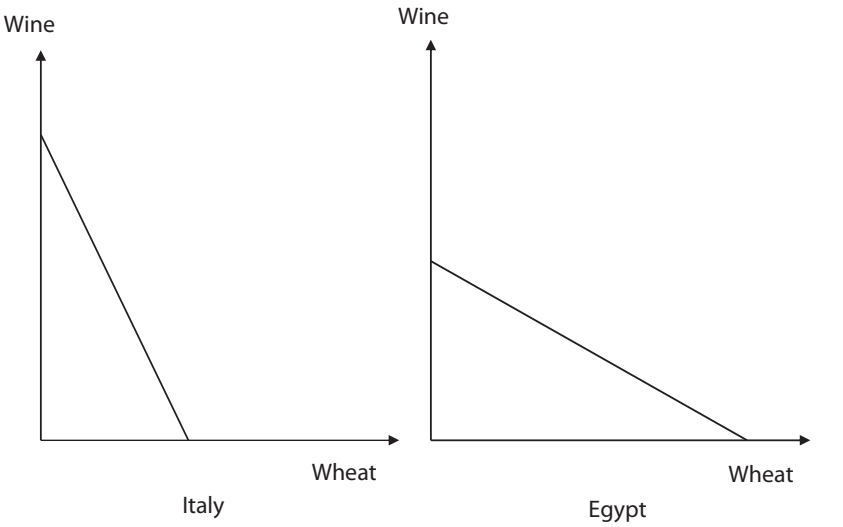

Figure 1.2. Production possibility curve

and wheat on the horizontal axis, we can draw a country's PPF. It will touch each axis where the country devotes all of its resources to the production of either wine or wheat, that is, if it specializes in one or the other. The PPF connects these two points. Ricardo assumed it ran in a straight line, assuming that the amount of wheat that needed to be given up to produce an extra unit of wine was not affected by the amount of wheat and wine being produced. He assumed there was a single input to production—call it labor—which was easily switched between the production of various goods. There was no second input like land and no diminishing returns like those introduced earlier to explain and upward-sloping supply curve.

This relationship is shown in figure 1.2. I show in this figure a PPF for each of two countries or regions that might trade with each other. The curves differ from one region to the other, even though both embody the same linear assumption. They differ in their slope. (The other possible difference—in height—will be discussed later.) One region, which we will call Italy, can make more wine more efficiently in terms of forgone wheat than the other region, which we will call Egypt. Egypt is well suited to growing wheat and needs to transfer a lot of resources from growing wheat to increase its wine production. The PPF for Italy therefore is steeper than the PPF for Egypt. (Note that another factor of production, land, has crept into the analysis to explain why countries differ.)

Consider the PPF for Italy. Where the PPF hits the vertical y-axis it shows how much wine would be produced in Italy if all the labor in Italy was used to

produce wine. Where the Italian PPF hits the horizontal x-axis, it shows how much wheat would be produced if all the labor was used to produce wheat. If Italian agriculture is not completely specialized in wine or wheat, then total Italian production is shown by a point on the PPF between these extreme positions. The slope of the PPF shows the (constant) amount of one product that has to be forgone to produce more of the other. The ratio of the prices of the two goods is the *inverse* of this slope. Since Italy can make so much wine if it chooses to specialize in wine production, wine is cheap in Italy. The same reasoning applies to Egypt, where the PPF is flatter because Egypt is more suited to growing wheat. Wine therefore is more expensive in Egypt than in Italy because wine is scarcer—as represented by the flatter PPF.

It is the difference in the steepness of the PPF between the two countries that allows them to have comparative advantages and gains from trade. I have drawn the curves about the same level, but nothing rests on that. Assume for a minute that Italy is more efficient at producing both wine and wheat than Egypt. If the two PPF curves have different steepness, it still will be worthwhile to trade. For example, consider a lawyer who is the best lawyer in town and also the best typist. She has an absolute advantage over her secretary, even though the secretary has a comparative advantage in typing. The secretary can do a lot of typing for each unit of law services he omits, even though he does less legal work and typing than the lawyer in any time period. It makes sense for the lawyer to specialize in doing law and delegate her typing to her secretary, even though she is better at both. Despite the lawyer's absolute advantage in both activities, she still can gain by exploiting her *comparative* advantage in legal services.

Return to figure 1.2. If there is a market, then the price of wine in terms of wheat will be higher in Egypt than in Italy, since the PPF is flatter. If farmers cannot sell wheat on any kind of market, they will make the choice of product by comparing the relative outputs they can get from their limited resources. We can express this choice as expressing what economists call the "opportunity cost" of producing wheat or wine. That is the amount of the product *not* grown in order to produce the one that is grown. The opportunity cost functions exactly the way the price does in a market, and I use price as a generic term to include both market prices and opportunity costs. Egyptian farmers would like to produce wine due to its high price; the flat PPF shows that they cannot do so with Egyptian resources.

Now assume that trade is introduced between Italy and Egypt. Wine is more expensive in wheat units in Egypt because the opportunity cost of producing wine is larger than in Italy. Egyptians then will want to export wheat to get wine, which is relatively cheaper in Italy. Italians face exactly the opposite incentives. Wine can be produced easily in Italy, and the Italians will be happy

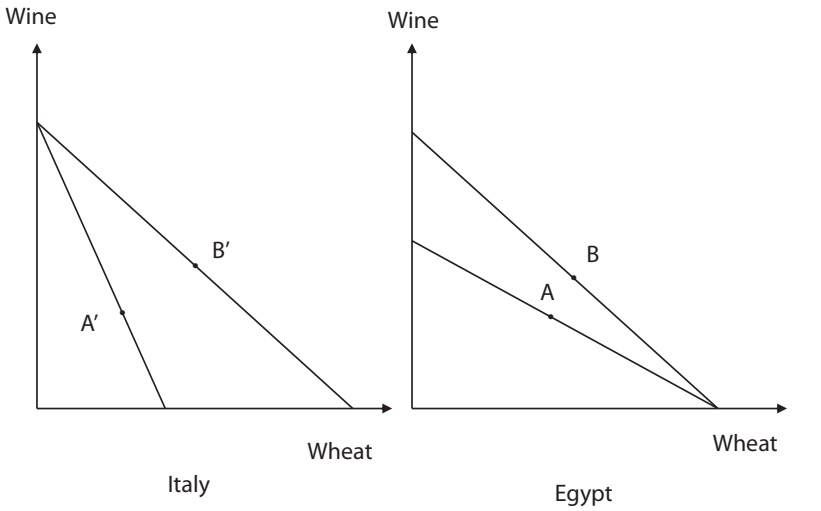

Figure 1.3. Effects of trade

to import wheat which is harder to grow (relative to wine). Trade will make both countries or regions better off.

The benefits are shown in figure 1.3. The price of wine was higher in Egypt before trade, and the price of wheat was higher in Italy. Once trade is allowed, both countries will have the same price ratio (in the absence of transport costs), which will be in between the initial price ratios in Italy and Egypt. The price of wine will fall in Egypt, allowing people there to get more wine for a given opportunity cost in wheat. Italy will use its resources to produce wine, getting its wheat by importing it. The initial consumption might have been at a point like A on the Italian PPF. With trade, Italy can now consume at point B, above the PPF and unobtainable without trade. Similarly, Egypt will use its resources to produce wheat and increase its consumption of wine and wheat from A' to B'. The price of wheat in terms of wine will fall in Italy, and rise in Egypt. The price of wheat in terms of wine, or of wine in terms of wheat, will be the same in both countries.

Adam Smith wrote that the division of labor was limited by the extent of the market. Trade extends the market between countries or regions and thereby promotes the division of labor. This is one way in which the extension of trade increases the earning of workers. Of course, if different regions or countries have resources unique to that locale, trade also allows these resources to be used for the benefit of the whole trading area (chapter 2).

Three extensions of this basic theory should be mentioned. First, what will be the new, common price of wine in terms of wheat? We know only that it must be between the original prices in Italy and Egypt, and the theory

explained here does not contain enough detail to demonstrate where it will fall in this range. The position depends on the volume and elasticity of supply and demand for the two goods in the two countries or regions. In particular, large countries or regions that have large supplies and demands have much more effect on the eventual price than small countries. (This is where the height of the PPF is important.) When Britain was brought into the Roman trade network, it got many more gains from trade than the rest of the Roman world. Interregional trade benefits both regions, but taxes may offset form of the gains. For example, much of the wheat sent to Rome from Egypt was tribute. We clarify the effects of this tribute by dividing it into two parts. Trade improved access to all products in both Rome and Egypt. Tribute transferred some—or perhaps all—of this gain from Egypt to Rome.

Second, the model as stated assumes that there are no transport costs when trade is allowed. That is why the price lines with trade in figure 1.3 have the same slope in both graphs, indicating that the relative prices of wine and wheat were the same in Italy and Egypt. In antiquity, transport costs often were quite high, both because of the cost of transporting goods and because of administrative costs like duties and verification. If there are significant transport costs, the price ratios in the two countries will not approach equality. Instead, they will remain apart by the cost of the transport. If this wedge is large enough, it may preclude trade even if the costs of production in the two countries are different.

Transaction costs never completely eliminate trade. Very rare and expensive goods can be traded profitably even if transaction costs are high. Before the Pax Romana, jewelry and royal objects were traded around the known world. But high transaction costs prevented trade in cheaper goods, like wheat. Only when costs were low did trade extend to bulk commodities and the articles of common usage. This kind of trade flourished in the early Roman Empire, but it had existed earlier across the Mediterranean Sea. Two Phoenician ships sank in deep water during the eighth century BCE, each carrying four hundred amphoras of wine. Their documentation has been lost, and we do not know why they were sailing, but it makes sense to infer that the people who sent eight hundred amphoras of wine into the center of the Mediterranean were engaged in interregional trade (Temin 2006c).

The New Institutional Economics reminds us that transaction costs may be affected by institutions as well as transport costs. Trade requires not only shops or carts, but also ways to compensate prospective merchants for their efforts in bringing goods to strangers. The means of payments, the security of contracts—even implicit ones—are aspects of the institutions that promote trade.

Third, Ricardo drew the PPF as a straight line, but economists now generally draw it curving above a straight line. A convex PPF describes an economy

in which there are diminishing returns to the production of wine and wheat. Here we consider two inputs to production, land and labor. If land cannot be transferred easily between different crops, there will be diminishing returns to labor in each activity. (This is the assumption that makes supply curves slope upward.) As the economy moves away from specialization in, say, wheat, it produces the first unit of wine by sacrificing only a tiny bit of wheat. In a position away from the axes where the economy is producing both wine and wheat, the economy has to give up a larger amount of wheat to free enough resources to make more wine. The gains from trade are the same as before with this complication, assuming that the internal price ratio of the goods differed initially in the two countries. The difference is that while countries will concentrate in the production of goods where they have a comparative advantage, they generally will continue to produce some of the other good as well. They will only specialize completely as shown in figure 1.3 if the cost structures in the two countries are very different.

Ricardo knew about diminishing returns; the rents to factors with inelastic supply curves often are known as Ricardian rents. But mathematics was not developed well enough two hundred years ago when Ricardo was writing for him to draw better diagrams. The necessary changes are shown in figure 1.4. In this diagram, the effect of land is shown directly as a cause of diminishing returns, not simply as a determinant of regional differences. The result is that the PPF for each region is curved. If all the labor is used for one or the other crop, there will be diminishing returns, and there will not be as much output as shown in figure 1.3. The initial price lines are now only tangent to the PPF at one point. That point shows where the PPF reaches the highest price line possible to maximize production at this relative price. As before, the initial points of production are labeled A and A′.

With diminishing returns, the effects of international trade are not as dramatic as before. As prices in both regions approach each other, each region moves along its PPF to reach the highest price line showing the new, international relative price. As before, consumption is now at B and B', above each country's PPF. Trade has allowed each region to benefit more than it could from using its resources in isolation. The basic insight of comparative advantage is maintained with this elaboration of Ricardo's theory.

There is, however, one important detail revealed by figure 1.4. Instead of going to a corner solution where each region produces only one product, each region goes only partway toward the relevant axis, to points C and C' in figure 1.4. Both regions specialize in the sense that they produce more of their export good, but they do not abandon production of their import goods due to diminishing returns in the export industry. In figure 1.3, each region was either isolated or completely specialized. In the more realistic figure 1.4, both regions

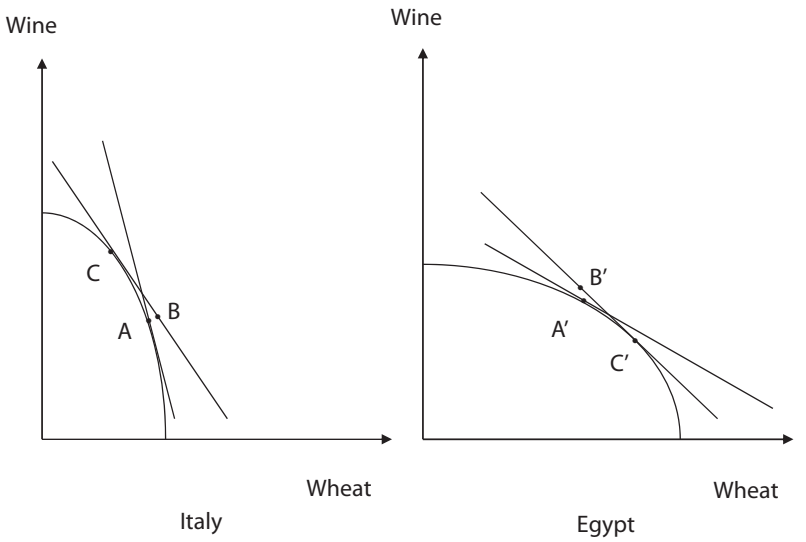

Figure 1.4. Effects of trade with diminishing returns

continue to produce both goods even after trade is introduced. Only if the production possibility frontiers are very different in the two regions or close to flat will there be complete specialization.

This is the form of comparative advantage taught universally today. Supply and demand curves provide tools for the understanding exchanges of individual commodities or services, whether through markets or other arrangements. The New Institutional Economics helps to evaluate the operation of markets. Comparative advantage provides a way to understand the economic interactions of regions, whether through markets or other kinds of transfers, illuminating the effects of the *Pax Romana*, the changing composition of production in Roman Italy, and economic expansion in the early Roman Empire. I show in the following chapters that a substantial part—perhaps most—of Roman exchanges were accomplished through markets, resulting in substantial improvements of living standards, particularly in Roman Italy.