

Introduction

International interdependence is a central feature of the world economy. The economic fortunes of countries are intertwined via trade, foreign direct investment, and financial capital flows. Production networks are spread across countries and continents, making the supply of products in one country highly dependent on economic activities in multiple foreign countries. The global crisis of 2008 illustrates this interdependence in a most vivid way; it led to a decline in the volume of world trade by more than one-quarter, negatively impacting countries whose *financial* systems remained sound. Moreover, the historical record shows that long-distance trade interacted in complex ways with economic development, and that it played a central role in the evolution of the world's economy. It is therefore important to understand what drives foreign trade and how trade affects economic outcomes. Building on a large research literature, this book offers this sort of understanding.

Unlike in the natural sciences, where important research objects do not change much over time, in the social sciences generally and in economics particularly, the objects of research alter and reshape. In this respect, international trade is no exception. When countries and regions transform as

a result of economic, technological, political, or institutional change, the nature of foreign trade changes too. Moreover, such changes are not rare in historical perspective, but rather frequent. As a result, the thinking on this subject has been repeatedly adapted to varying circumstances. This motif serves as an organizing principle of this book, which explains the evolution of scholarly research on the structure of world trade from its inception to its present form.

While long-distance trade plays an essential role in modern economies, it was also a salient feature of economic development after the Neolithic Revolution, as hunter-gatherers evolved into sedentary societies that specialized in food crops. The importance of trade further increased with the emergence of cities and early civilizations. Caravans traveled along the Fertile Crescent, trading between Mesopotamia and the Levant, and trading routes expanded over time to distant parts of Asia and Europe. The Roman Empire managed an extensive network of trade, which according to McCormick (2001, p. 778) bound together three continents: Europe, western Asia, and Northern Africa. Large volumes of goods traveled by sea and land, including oil and grain, with much of the trade being subsidized by the Empire. Merchandise moved across Europe primarily on north-south routes, and across the Mediterranean.

The collapse of the western part of the Roman Empire in the fifth century c.e. brought many of these developments to a halt. Ward-Perkins (2005) documents the decline of the standard of living in the Empire's regions. He reports archaeological research showing that the Romans lived in a "sophisticated world, in which a north-Italian peasant of the Roman period might eat off tableware from the area near Naples, store liquids in an amphora from North Africa, and sleep under a tiled roof" (pp. 87–88). Long-distance trade was instrumental in preserving this standard of living, which was not limited to the elite but filtered down to the masses through the availability of high-quality functional products. Ward-Perkins' map in figure 5-4 (2005, p. 98) illustrates the wide distribution, across all of Europe and North Africa, of one type of pottery

mass-produced in southern France. According to a long-prevailing view, the collapse of the Roman empire was followed by “dark ages.”¹

McCormick (2001) documents in great detail the evolution of communications and the mobility of people across distant regions. He argues that despite the lack of good data on commerce, these developments—which were particularly pronounced during the Carolingian Empire in the eighth century—point to the presence of extensive long-distance trade. European imports of spices were replaced by imports of exotic medicines and new drugs provided by Arab pharmacology, while silk continued to flow into northwest Europe. “To pay for these imports,” McCormick states, “Europe produced a rather narrow range of high-value, low-bulk goods. Some textiles, perhaps, and some tin seem plausible, if barely documented. Fur, probably, and Frankish swords certainly were exported to the Muslim world” (p. 791). But the largest exports around 800 were European slaves, who were in high demand in Spain and in the more advanced economies of Africa and Asia. This trade played a central role in the advance of the European economy.

A matrix of traded products across eight regions circa 1000 is provided by Findlay and O’Rourke (2007, table 2.1).² According to these data, for example, western Europe exported swords to eastern Europe and slaves and swords to the Islamic world, while eastern Europe exported slaves, furs, and silver to western Europe and the Islamic world, and furs and swords to Central Asia. The Islamic world exported pepper, spices, textiles, silk, and silver to western Europe, and textiles and silver to eastern Europe. It also exported textiles to Central Asia, and textiles, swords, and horses to Sub-Saharan Africa. As a final illustration, East Asia exported silk to the Islamic world, Central Asia, South Asia, and Southeast Asia; it exported porcelain to the Islamic world and South Asia, tea to Central Asia, and copper to Southeast Asia.

The Middle Ages saw an expansion of trade with the rise of city-states such as Venice and Genoa and the advent of the commercial revolution (see Findlay and O’Rourke, 2007, chapter 3). Although quantitative data

are hard to come by, Findlay and O'Rourke (p. 140) report Wake's (1986) findings about the spice trade. Pepper imports increased by a modest 20 percent between 1400 and 1500, while imports of spices other than pepper (e.g., cloves and nutmeg) more than doubled during this period. Venice imported about 60 percent of the pepper at the beginning of the period and less than half of the other spices.

Although there is no doubt that the discovery of America by Christopher Columbus in 1492 and the discovery of the passage to the East Indies via the Cape of Good Hope by Vasco da Gama in 1498 had monumental effects on world history and on long-distance trade, historians dispute the immediate impact of these discoveries. The Iberian states of Portugal, Castile, and Aragon had obviously been affected, and the rest of the world was influenced in the following centuries. But how important were these discoveries for the global integration of markets? Some historians argue that world markets were integrated before the age of discovery, while others argue that integration started in earnest only afterward. Surely, the number of voyages to the Americas and the Indies greatly increased as a result of these discoveries, and so did the tonnage of shipments. Yet O'Rourke and Williamson (2002) show that price convergence across regions—which is an important measure of market integration—did not take place before the nineteenth century.³

Findlay and O'Rourke (2007, figure 4.5) present data, based on de Vries (1993), on the number of ships sailing to Asia per decade and the tonnage returned by them to Europe. Panel (a) of their figure is reproduced in Figure 1.1.⁴ Evidently, Portugal dominated this route in the sixteenth century, after which the Netherlands took over as leader. Moreover, the number of ships that sailed from Portugal declined over time. Nonetheless, the tonnage they brought back to Portugal during the sixteenth century did not fall, because they became bigger and a larger fraction of the departing ships returned (see Findlay and O'Rourke, 2007, p. 185). Portuguese tonnage of shipments started to decline, however, after the rise of competition from the Netherlands, England, and France.

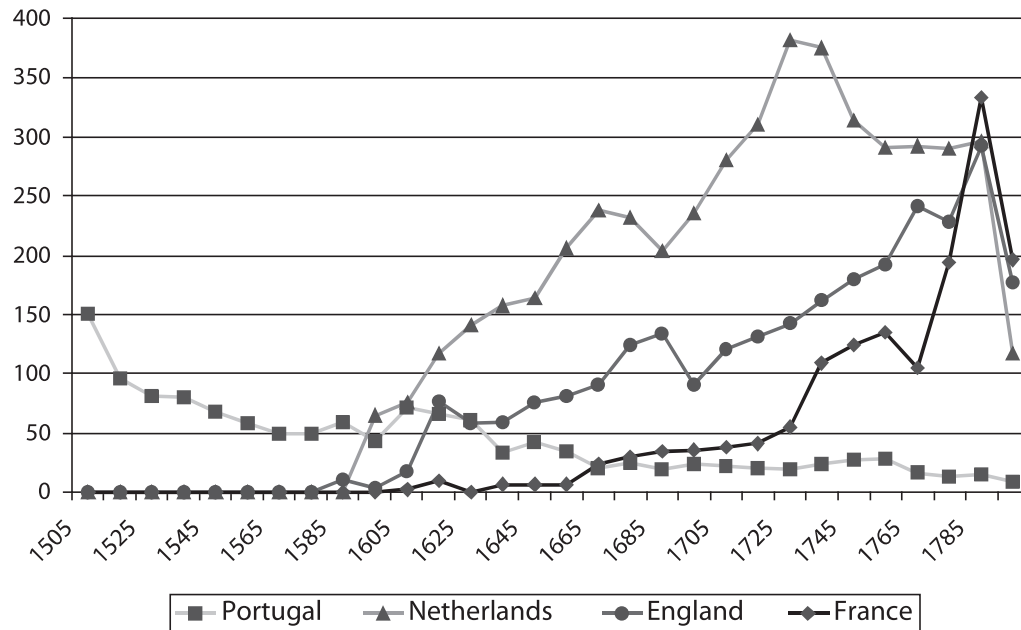


Figure 1.1. Number of ships sailing to Asia per decade. Data from Findlay and O'Rourke (2007, figure 4.5).

Although the discovery of the New World and the passage to the Indies played a prominent role in the evolution of the European economies in the centuries to come, the volume of world trade relative to income remained very small until the nineteenth century, in line with the price convergence argument in O'Rourke and Williamson (2002). According to Estevadeordal, Frantz, and Taylor (2003), imports plus exports reached only 2 percent of gross domestic product (GDP) in 1800, then increased to the first peak of 21 percent in 1913—just prior to World War I—and declined between the two world wars, as shown in Figure 1.2.⁵ After World War II trade rose faster than income, as shown in Figure 1.3, and the trade-to-income ratio climbed, surpassing the 1913 peak in the early 1970s. Today the ratio of trade to income is much higher than ever before.

According to O'Rourke and Williamson (2002), long-distance trade in the pre-eighteenth-century period consisted for the most part of non-competing products—that is, products that were not produced in the importing regions (e.g., spices, silk, woolens). In the early nineteenth century

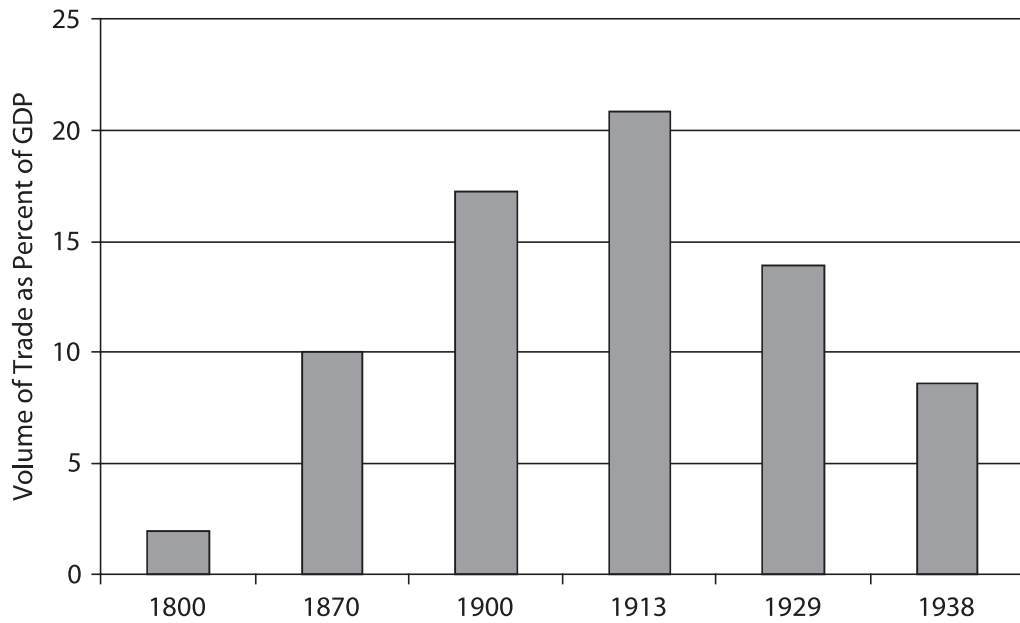


Figure 1.2. World imports plus exports as a percent of world GDP. Data from Estevadeordal, Frantz, and Taylor (2003).

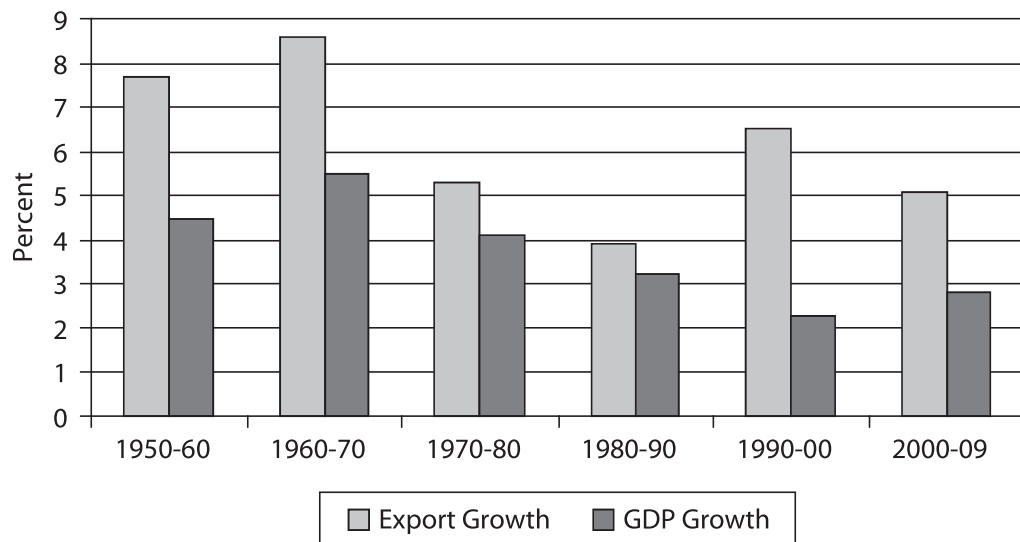


Figure 1.3. Average growth rate per decade of the volume of world exports and GDP. Data from the World Trade Organization, International Trade Statistics, 2009 (accessed online on April 6, 2010).

it also covered significant amounts of basic goods, such as wheat, and simple manufactures, such as textiles. During the nineteenth century, trade expanded rapidly, partly because of an astounding decline in transport costs and partly because of the rise of manufacturing. As a result, price gaps across markets were greatly reduced. The shift from noncompeting to basic goods also enabled trade to expand at a rapid pace. The last phase, which consisted of additional growth and diversification of manufacturing industries and the rise of product differentiation, further enhanced the growth of world trade. That is, the initially low volume of trade and its subsequent growth were materially influenced by the transformation of production and consumption.

While the evolution of long-distance trade was related to economic development, the interdependence between them was complex. In particular, one cannot argue that the effects were unidirectional (i.e., that economic development brought about trade expansion), because long-distance trade affected economic development and economic development affected trade. Moreover, the influence of trade on economic development operated through multiple channels, including institutional and political means.

Long-distance trade in the aftermath of the discovery of the Americas is often cited as a major event that contributed to the divergence in economic conditions between Europe and China. Although Europe and China were similarly advanced in the mid-eighteenth century, the Industrial Revolution took place in Europe, as a result of which Europe grew faster than China. This eventually led to large gaps in income per capita (see Pomeranz, 2000). Naturally, the Industrial Revolution was not driven by trade per se; trade was, rather, a contributing factor. According to Allen (2009), the availability of cheap coal and high wages in England induced the development of technologies that substituted machines for labor, and these technologies fueled British economic growth.⁶ As the new technologies spread to continental Europe, they prompted economic growth in the Netherlands first and then in other western European countries.

Acemoglu, Johnson, and Robinson (2005) point out that European growth in the post-1500 period was concentrated in countries with access to the Atlantic Ocean: Britain, France, the Netherlands, Portugal, and Spain, countries that engaged in trade with the New World and acquired overseas colonies. These commercial opportunities strengthened the political power of merchant groups and entrepreneurs and weakened the power of monarchs. As a result, constraints on the executive were broadened and property rights became more secure for a larger segment of society. These unintended consequences of trade with the New World enabled the Atlantic traders to forge ahead of other European countries.⁷

Yet trade with the New World also had negative effects in Spain, where Castilian institutions proved to be inadequate in limiting the power of Philip II. The flow of silver from the Americas encouraged Philip II to engage in wars that eventually became too expensive and required domestic taxes and large loans from foreign bankers. According to Drelichman (2005) and Drelichman and Voth (2008), the ensuing struggles between the Crown and the Cortes weakened domestic institutions, and this had unfortunate consequences for Spanish economic growth.⁸

Although long-distance and international trade are not exactly the same phenomena, they are closely related in that much of long-distance trade is also international trade. True, trade between California and Massachusetts (two states of the United States) is long-distance, as is trade between British Columbia and Quebec (two provinces of Canada), yet neither is international. And trade between Turkey and Syria or Israel and Jordan, which is international, is short-distance in comparison. Nevertheless, our discussion will focus on international trade, emphasizing the movement of products across national borders.

Countries differ in national features that shape the structure of foreign trade, while geographic attributes are more important in shaping trade flows across regions within a country. Moreover, regions within a country are more integrated than regions of different countries. Nevertheless, the

fields of international trade and regional economics share common traits that were emphasized already by Ohlin (1933), such as transport costs, agglomeration of economic activity, and the unlinking of production from consumption.

My exposition follows the historical evolution of the field of international trade, highlighting the interplay between theory and evidence. In particular, I explain theoretical arguments in the context in which they were developed, the evidence that was amassed to test or challenge them, and modifications of the theoretical arguments that were developed in order to accommodate new evidence. This chain of theorizing, empirical investigation that confirms parts of the theory and contradicts others, and updating of the theory in view of new evidence has been a fruitful pattern in understanding international trade. Moreover, this pattern has been unavoidable in view of the changing nature of international economic interactions. In other words, theories that had been suitable at one time became less appropriate as national economies—and with them patterns of international specialization—changed.

Chapter 2 discusses the two major paradigms of foreign trade that were developed in the early parts of the nineteenth and twentieth centuries, respectively—the former by David Ricardo, the latter by Eli Heckscher and Bertil Ohlin. Each one was cultivated in the context of its time; the former explained trade flows by differences in labor productivity across countries, while the latter explained foreign trade by differences in factor endowments (i.e., the availability of productive resources such as labor, capital, and land). In each case the underlying causes of trade were designed to address specific issues. These two paradigms were extensively studied during most of the twentieth century, and they were applied to a host of issues, such as gains from trade, the conflict of interest between different groups in society concerning the desirability of open markets, the impact of trade policies—including free trade agreements and multilateral trade negotiations—and the relationship between international trade and economic growth. Gains from trade and distributional conflicts

are examined in Chapter 3. The impact of trade policies on national economies is a large subject that deserves a book-length treatment of its own, and I have chosen not to cover it in this deliberately short presentation. The relationship between trade and growth is an important subject that is dealt with in detail in Helpman (2004, chapter 5), and since I believe that there is not much new that can be added to that treatment, I have not included it in this book.

While the neoclassical trade theory that dominated the thinking about foreign trade during most of the twentieth century has many merits, it proved inadequate in addressing a number of phenomena that became particularly salient in the post–World War II period. This led to the first major revolution in trade theory—in the early 1980s—and to the development of what was then dubbed the “new” trade theory, which I discuss in Chapter 4. The motivation for this revolution was empirical, and the “new” trade models—which emphasize economies of scale and monopolistic competition—triggered new empirical work. Yet as the nature of world trade kept changing and new data sets became available in the 1990s, the inadequacy of the theoretical models from the 1980s became visible. As a result, a second revolution took place in the early 2000s, this time focusing on characteristics of individual firms and how they engage in international transactions. The first stage of this revolution is discussed in Chapter 5, focusing on trade in goods. The second stage is discussed in Chapter 6, focusing on foreign direct investment (i.e., the ownership of subsidiaries in foreign lands), offshoring, and outsourcing.

One may wonder why foreign direct investment (FDI) has not been mentioned so far. Was it not important in the more distant past and became important only in the 1990s? The answer is, of course, that international capital flows, including FDI, played a major role in economic history, and especially so with the expansion of colonies after the discovery of the New World. Foreign asset holdings, in the form of FDI and foreign financial assets, grew fast in the nineteenth century in tandem with the growth of trade, and their size relative to GDP expanded at the

end of the nineteenth century. Though foreign asset holdings collapsed after World War I, like foreign trade, they expanded rapidly again after World War II (see Obstfeld and Taylor, 2004, table 2.1).

Much of the expansion of foreign asset holdings in the postwar period has been portfolio investment, with which we shall not deal in this book. The discussion of FDI, which is intimately related to foreign trade, is relegated to Chapter 6, where it is possible to build on the knowledge acquired in the previous chapters. FDI flows grew fast after World War II, reaching a peak before the collapse of the dotcom bubble in 2001, when they also collapsed. They recovered, regaining their 2001 peak between 2006 and 2007 (see UNCTAD, 2009). Importantly, the emergence of worldwide networks of production, in which multinational corporations play center stage, can be understood only within the broader frameworks discussed in Chapters 4 and 5.

The last chapter provides a brief discussion of two ongoing research programs that are not treated in the main chapters of this book, along with reflections on future directions. Since the study of international trade and foreign direct investment has become highly specialized, I very much hope that this nontechnical exposition of what is known on this subject will help the reader better understand the world around us.