THE POLITICAL ECONOMY OF THE WORLD TRADING SYSTEM

THE WTO AND BEYOND

Third Edition

BERNARD M. HOEKMAN
and
MICHEL M. KOSTECKI

OXFORD
UNIVERSITY PRESS

INTRODUCTION

ESTABLISHED in 1995, the World Trade Organization (WTO) administers the trade agreements negotiated by its members, in particular the General Agreement on Tariffs and Trade (GATT), the General Agreement on Trade in Services (GATS), and the Agreement on Trade-related Intellectual Property Rights (TRIPS). The value of total world trade in goods and services, including payments for intellectual property was some US$16 trillion (thousand billion) in 2007. The WTO's rules and principles establish a legal framework for much of this exchange.

The WTO builds upon the organizational structure that had developed under GATT auspices. At its creation in 1947, the GATT was essentially a tariff agreement. Over time, as average tariff levels fell as a result of periodic rounds of negotiations, the focus expanded to cover nontariff barriers (NTBs) and domestic policies with an impact on trade. A complex patchwork of policy-specific agreements emerged. Participation in the GATT expanded steadily. By the end of the Uruguay Round (1994), 128 countries had joined the GATT. Since the entry into force of the WTO in 1995, another two dozen countries acceded, bringing the total to 153 as of 2008. Suggestions made during the Uruguay Round negotiations that 'GATT is dead' and more recent criticisms of the WTO sit oddly with these signs of popularity.

The underlying philosophy of the WTO—as was the case for the GATT that preceded it—is that open markets, transparency and nondiscriminatory trade policies are conducive to the national welfare of all countries. A rationale for the way the organization works—through reciprocal negotiations that define enforceable commitments and mutually agreed rules of the game for trade-related policies—is that the prospect of better access to export markets helps governments overcome political constraints that prevent the adoption of more efficient trade policies.

Although there are many similarities with the GATT, the WTO differs in a number of important respects from the institution that preceded it. These differences have potentially important implications for the functioning of the trading system, in particular for developing economies. The GATT was a rather flexible institution. Bargaining and deal-making lay at its core, with significant opportunities for countries to 'opt out' of specific disciplines. This is much less the case today. The WTO rules apply to all members, and are subject to binding dispute settlement procedures. This is attractive to groups seeking to introduce multilateral disciplines on a variety of subjects—ranging from the environment and labour standards to competition and investment policies to animal rights. But it is a source of concern to groups who perceive the (proposed) multilateral rules to be inappropriate or worry that the adoption of specific rules may affect detrimentally the ability of governments to regulate domestic activities and deal with market failures in ways that they deem most appropriate.

Almost from the start of its existence the WTO attracted a significant amount of critical attention. Public concerns are to some extent a reflection of the increasing speed at which global integration is occurring. Between 1900 and 2000 the value of international trade doubled. The cross-border flow of foreign direct investment (FDI) expanded even more rapidly, growing 10 times faster than world production. Since 2000, trade and FDI flows have continued to grow at a blistering pace, with world trade doubling once again and the stock of outward FDI passing the $13.5 trillion mark in 2007—accounting for some 10 per cent of world output (UNCTAD, 2008). In the post 2000 period, the growth rates of developing country trade have significantly exceeded those of high-income countries, helping to sustain high rates of economic growth and reductions in poverty rates. These positive trends coincided with more liberal trade policies and market-oriented reforms. Only one employee in ten is currently working in countries that are largely separated from the world market, compared to two-thirds some three decades ago (Dicken, 1998). Multinational corporations have assumed a much greater role in the world economy. In 2006, some 75 million people were employed by foreign affiliates of multinational companies, a threefold increase compared to 1990, with much of the growth in developing countries. One-third of all affiliates of multinational companies are located in China, although most of these tend to be small in scale or joint ventures (UNCTAD, 2007).

The high and sustained rates of growth in trade and cross-border investment flows have been beneficial from a global economic point of view, with the increase in developing countries' market shares improving the global distribution of income and economic activity. Developing countries as a group now account for 35 per cent of world trade, up from only 20 per cent in the early 1990s. The most impressive performer has been the Chinese economy, more than tripling its share in world exports between 1990 and 2007 and on track to become the largest exporter of goods in 2008 (WTO, 2008). 'Fear of China' is common among producers of

manufactures around the globe, not least in other developing countries, as is 'fear of India' among producers in rich countries of what used to be nontradable services supplied to businesses and households. The absence of significant across-the-board imposition of protectionist measures in importing countries is as much a testimony to the success of the trading system as is the observed growth in trade it has helped generate. Whether the open global trading system that now exists can be sustained depends on how effective societies will be in undertaking and facilitating the adjustments that are needed to accommodate the economic expansion of developing countries. To some extent those adjustments may imply the use of trade policies that are permitted by WTO rules—such as antidumping, which has been increasingly used by many countries—but mostly they require domestic policies that target the affected groups directly.

The need for liberalization is greatest in agriculture where policies in rich countries impose significant negative spillovers of many developing countries—a central issue in the Doha Development Agenda (DDA), and the one that is largely responsible for the slow progress of the negotiations. As important as managing adjustment, is for countries that have seen their trade shares increase greatly to play a more prominent role in contributing to the public good of an open multilateral trading system. This implies a greater willingness to lock in liberalization through the WTO and accede to requests for fuller reciprocity in trade negotiations. Such fuller participation in the WTO by the emerging market economies is one of the major challenges confronting the multilateral trading system. Without greater reciprocity, pressures will grow for protection in importing countries that are being forced to adjust as their domestic industries contract as a result of competition from (more) efficient exporters. Increasingly this is a services agenda, intermediated by foreign direct and portfolio investment (including sovereign wealth funds).

The growth of developing country trade since the mid-1990s is both an example of the importance of the WTO and the challenge confronting the organization. New trade powers need to be accommodated and integrated into the system, which requires greater attention to be given to addressing the concerns of low-income countries that are lagging behind. The enormous heterogeneity of the WTO membership—which can no longer be characterized by a North–South divide—complicates the needed agreement on where the boundaries of the institution lie. Views differ significantly on what the objectives of the WTO are or should be.

A number of the ministerial meetings of the WTO post-1995 were accompanied by demonstrations by groups spanning the nongovernmental organization (NGO) community, farmers and labour unions seeking to limit or to expand the reach of multilateral disciplines. High-profile and sometimes violent street protests during the 1999 Seattle WTO ministerial helped scuttle efforts to launch the so-called millennium round and marred ministerial meetings in Cancun (2003) and Hong Kong (2005). In Cancun, a Korean farmer committed suicide in front of TV cameras for the world to see. Many citizens of member countries have a very

limited and often skewed understanding of what the WTO is and does, and what it is not and does not do. The extent of the disconnect that emerged in the late 1990s is exemplified by a 1999 Swiss TV programme in which a small boy is scared to go to sleep because 'there is a WTO under my bed'. Although opposition to the GATT and the Uruguay Round was quite intense at times—giving rise to posters representing the institution as a 'GATTzilla' (referring to the cartoon monster Godzilla)—it never reached the point where a TV producer could feel comfortable assuming it impacted on the fears of children. Matters have improved substantially since 2000 as a result of an active outreach effort by the WTO Secretariat, but NGO concern and opposition remains prominent, as illustrated by a steady drumbeat of WTO criticism at the World Social Forum—a rival convention to the WTO-friendly World Economic Forum held annually in Davos, Switzerland.

Although efforts to liberalize trade have always been opposed—sometimes very vocally—by domestic groups who stand to lose from greater competition (for example farmers in high-income countries), the terms of the debate surrounding the WTO now extend well beyond the traditional trade liberalization agenda. Understanding how the WTO works, its strengths and weaknesses, and what might be done to make the institution a more effective tool of multilateral cooperation is vital. Many of the WTO's critics continue to have serious misconceptions about the organization, while many of those who are seeking to expand the WTO's mandate often appear to ignore basic principles of economics and risk the continued viability of an institution that has played a key role in sustaining the open economic system, which has helped raise per capita incomes in much of the world to levels never seen before in history. At the same time, some of the criticism reflects deeply held beliefs and concerns. Some of the subjects that are a bone of contention cannot or should not be dealt with by the WTO and claims of sins of commission or omission are therefore often inappropriate. But some matters can and should be laid at the door of the WTO.

Our goal in this book is to provide a succinct description of the principles, rules and procedures of the multilateral trading system, as well as a political economy-informed discussion of how it functions. This book does not provide a detailed negotiating history—who did what and when—although the results of negotiations and ministerial meetings are discussed at some length, including the subjects that were on the table in the DDA. Being an introduction, this book cannot be more than a starting point. Guides to further reading are provided at the end of every chapter. Readers interested in pursuing specific subjects in greater depth should consult the works recommended there as well as the bibliography.

The book is organized into five parts. Part I provides a brief historical overview of the evolution of the multilateral trading system, major developments in world trade and introduces the basic functions of the trade regime (Chapter 1). Part II deals with the WTO as an institution. Chapter 2 describes the organizational structure of the WTO, its scope and functions. Chapter 3 discusses WTO enforcement and dispute

settlement provisions, and summarizes the case load to date. Chapter 4 analyses the role of the WTO as a forum for negotiations. Special attention is given to the concept of reciprocity, as this is a key element of multilateral trade negotiations (MTNs).

Part III discusses the core disciplines of the WTO, which are contained in three multilateral agreements. Chapter 5 describes the GATT rules for merchandise trade—disciplines on tariffs, quotas, subsidies, customs procedures and product standards, among others. In each instance we discuss the political economy rationale underlying the rules, using cases and examples drawn from practice to illustrate their relevance and operation. Chapter 6 turns to the major sector-specific agreements that have been negotiated under GATT auspices, the three most important being the Uruguay Round Agreements on Agriculture and on Textiles and Clothing, as well as the Information Technology Agreement (ITA). Both sectors have a long history of protectionism in many countries, and both continue to have higher levels of protection in many countries than other sectors. Chapters 7 and 8 discuss the two major additions that were made to the trading system in the Uruguay Round: disciplines on policies affecting trade in services as embodied in the GATS, and the agreement on TRIPS respectively.

In Part IV, we describe and assess the major 'holes and loopholes' in the WTO. The various mechanisms allowing for the re-imposition of trade barriers are discussed in Chapter 9, which summarizes the rules on—and the economics of—the use of instruments of contingent protection. These have been very important in dealing with domestic political pressures for (re-)imposition of protection. Although often abused to the detriment of both national and global welfare, recent research has shown that they can also play a constructive political function by helping governments that decide to undertake far-reaching liberalization to implement and sustain economy-wide policy reforms. Chapter 10 deals with one of the most important exceptions to the most favoured nation (MFN) rule allowed by the WTO: preferential trade agreements (PTAs). Almost all WTO members are participants in one or more PTAs, raising serious questions about the practical relevance of the WTO nondiscrimination principle. Since the creation of the WTO the number of PTAs has increased steadily, as have nonreciprocal duty-free access schemes for least developed countries (LDCs). As a result, multilateralization of preferential trade is one of the major challenges confronting the WTO. Chapter 11 discusses the provisions of the WTO allowing for the negotiation of so-called plurilateral agreements, which apply only to those members that sign them. The most important of these is currently the Agreement on Government Procurement. The use of such agreements may well increase in the future, as it allows for subsets of members to move forward in areas where consensus cannot be obtained.

Part V addresses recent trends and challenges confronting the WTO. Chapter 12 discusses the evolving role of developing countries and former centrally planned economies in the multilateral trading system, and the concerns that these countries

have regarding its operation. Chapter 13 deals with a number of subjects that are likely to be on the negotiating agenda for some time to come, including competition (antitrust) policy, labour standards, investment and environmental policies. Chapter 14 turns to the question of governance of the trading system, the role of NGOs and the importance of ensuring domestic transparency of trade and investment policies.

The concluding chapter briefly summarizes some of the major themes that emerge from previous chapters and discusses possible futures for the WTO and the challenge of sustaining international cooperation in the trade area post-Doha.

The volume includes two annexes. Annex 1 provides a listing of WTO members and some of the key characteristics that help determine their influence and participation in the institution. Annex 2 summarizes the economics of major trade policy instruments. It covers tariffs, quotas, trade in services, subsidies, externalities and market failure, price discrimination (dumping), FDI, trade preferences, preferential public procurement and rent seeking. Although the discussion in the volume is mostly nontechnical, we hope inclusion of the material in Annex 2 will assist students of international relations, economics and business, as well as the interested reader, to relate basic economic concepts and analytical frameworks to the trade policy instruments that are the subject of WTO disciplines.

CHAPTER 1

THE TRADING SYSTEM IN PERSPECTIVE

ECONOMIC theory suggests that countries should pursue liberal trade policies and exchange goods and services on the basis of their comparative advantage. In practice, however, most nations actively intervene in international trade. Since 1947, the GATT has been the major focal point for industrialized country governments seeking to lower trade barriers. Progress towards liberalization of trade was fitful at times, often involving two steps forward and one step back. Nonetheless, recurring MTNs and the positive demonstration effects of the success of outward-oriented development strategies aimed at integration into the world economy resulted in a steady decline in the average level of protection in most countries. The processes and disciplines of the GATT helped governments to liberalize trade and to resist pressures for protection. This in turn helped foster ever-greater integration of the global economy through trade. The extent to which world trade has grown since the 1950s is truly phenomenal, especially when put in historical perspective. The volume of trade increased 27-fold between 1950 and 2006, three times more than the growth in global gross domestic product (GDP) (WTO, 2007). The GATT and, since 1995, the WTO played an important role in creating the multilateral framework that has supported this trade expansion.

1.1. Trade, Growth and Global Integration

The value of global trade in goods and services passed the US$15 trillion (thousand billion) mark in 2006. At US$12.5 trillion, trade in goods accounted for the lion's share of global flows, followed by trade in commercial services, which had grown to US$2.7 trillion in 2006 (WTO, 2007). Reported data on trade in knowledge, as measured by payments of royalties for use of trademarks, patents and so forth, added up to some US$140 billion in 2006. As data on both trade in services and the arms-length exchange of knowledge are incomplete (we return to this in the chapters on services and intellectual property), the US$15 trillion figure is an underestimate of the value of total cross-border flows of goods and services.

With only a few exceptions, notably in 2009 as a result of the global financial crisis and recession, trade has grown more rapidly than output each year since 1950 (Figure 1.1). The more rapid growth of trade as compared to GDP—by a factor of 2 during the post-1990 period—has resulted in rising merchandise trade-to-GDP or openness ratios for all regions (Figure 1.2).

Trade growth has been driven by a mix of technological and policy changes that reduced trade costs. These in turn have generated changes in the organization of production, stimulating a great increase in so-called vertical specialization, with

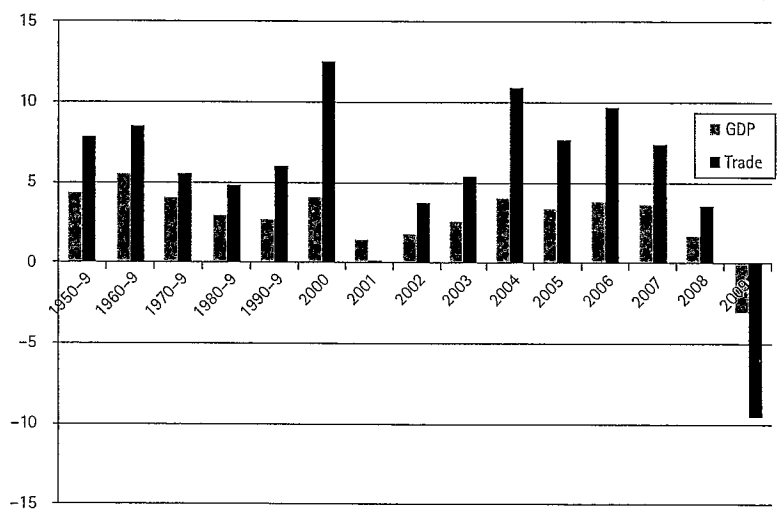

Fig. 1.1. Growth in the volume of world merchandise trade and GDP (per cent)

Source: WTO.

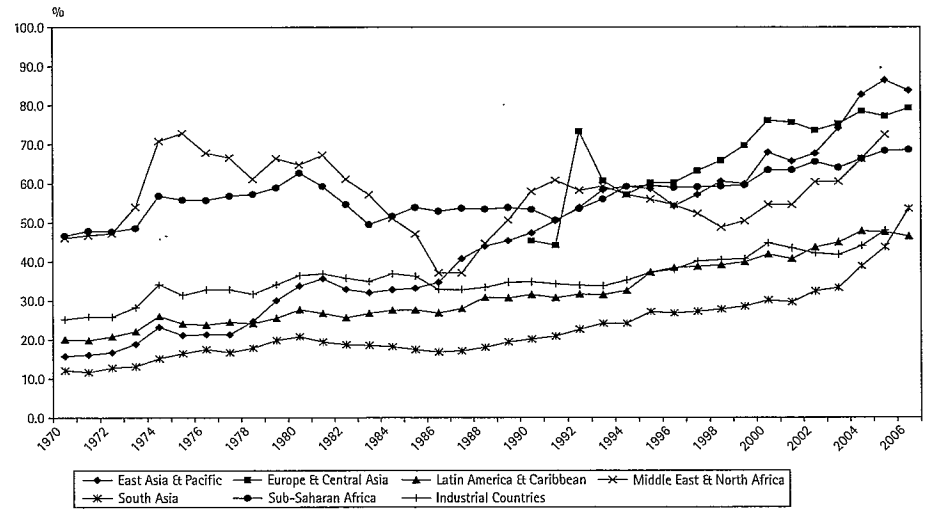

Fig. 1.2. Trade openness (ratio of trade to GDP), 1970–2006

Source: World Bank, World Development Indicators.

firms concentrating on (specializing in) specific bits of the production chain for a product. All these changes are inter-related, with policy reforms helping to stimulate technological change, and technological and managerial changes in turn putting pressure on policies. As discussed later in this book, the GATT played at best a marginal role in the trade policy reform process in developing countries—its impact was largely restricted to Organization for Economic Cooperation and Development (OECD) countries, as these were the nations that participated most actively in the institution. Developing countries began to liberalize their trade unilaterally during the 1980s and 1990s, supported by the international financial organizations, in particular the World Bank and the International Monetary Fund (IMF). The significance of the trading system for developing country members' policies only began to rise after the creation of the WTO in 1995.

Starting with average tariffs in the 20–30 per cent range around 1950 (WTO, 2007), complemented by a variety of NTBs that were often more binding (including quantitative restrictions and exchange controls), over time and in large part through recurring MTNs, average levels of protection of industrialized countries were lowered. As of 2006, the average uniform tariff equivalent of OECD merchandise trade policies was only 4 per cent (Kee, Nicita and Olarreaga, 2008), mostly reflecting protection of agriculture. Imports of many manufactures are

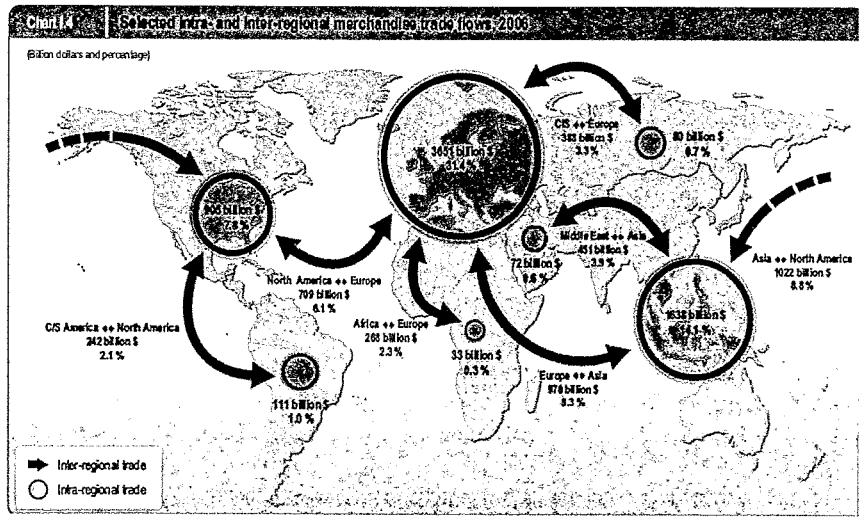

Fig. 1.3. Inter- and intra-regional trade flows, US$ billion and per cent, 2006

Source: WTO (2007).

now duty-free. In contrast to developed economies, most developing economies maintained relatively high barriers to trade for much of the post-Second World War period. Only in the 1980s did many low-income countries start to liberalize their trade. These differences in policy stances help in understanding the prevailing patterns of trade.

Global trade flows are dominated by exchanges within and between the three major regions of the global economy (the so-called triad): Europe, North America and East Asia. In 2006, intra-East Asia and intra-North American trade—represented by the circles in Figure 1.3—accounted for 53 per cent of world trade and about two-thirds of the total merchandise trade of the three regions. Trade in goods between members of the EU accounts for about one-third of global merchandise trade. Some 40 per cent of all developing country exports are destined for other developing countries.

All 49 LDCs[1] together accounted for only 1 per cent of world trade in 2006, reflecting the small size of their economies and very low per capita incomes. Their share has actually fallen over time—it stood at 1.7 per cent in 1970. South Asia and

[1] A country is defined as an LDC on the basis of a set of criteria applied by the United Nations Economic and Social Council (ECOSOC). The list is reviewed every three years to determine whether countries should be added or graduated from the list. In December 2007 Cape Verde was graduated, reducing the number of LDCs to 49. The next country to graduate will be the Maldives (in 2011).

Sub-Saharan Africa (including South Africa) together represent some 5 per cent of world trade; Latin America (excluding Mexico) another 3.6 per cent. The low trade shares of many of the poorest countries is in part a reflection of the very rapid increase in trade of East Asian countries, as well as the sustained growth in trade between high-income countries. In absolute terms most developing countries have seen the value of trade increase over time. The weaker relative trade performance of many developing countries has implied that the interests of individual developing countries are increasingly distinct, with some having benefitted greatly from the open global trade regime, but many others not perceiving significant benefits associated with WTO membership. As a result, the institution has become more concerned with the question of how to enhance the benefits of membership and assist poor countries in harnessing potential trade opportunities.

Developing countries have increasingly become producers and traders of manufactures. The share of manufactures in total exports of developing countries increased from just 30 per cent in 1980 to some 70 per cent in 2005—almost as high as in high-income countries (Figure 1.4). A substantial proportion of this global trade in manufactures, especially between OECD countries, comprises intra-industry trade—the exchange of similar, differentiated products. Intra-industry trade ratios are frequently above 60 per cent for OECD countries. Since the 1990s they have risen to similar levels for dynamic developing and transition economies. The corollary of the great increase in the share of manufactures in developing country exports is a fall in the share of agriculture and natural resource products. Of course, there are substantial variations in the share of manufactures in exports across countries and regions. Sub-Saharan African countries remain heavily dependent on (specialized in) natural resources and agriculture—e.g. oil, cocoa, cotton or coffee—as do many countries in Latin America (e.g. beef, sugar, grains) and the Middle East (oil). For the world as a whole, however, trade in merchandise is

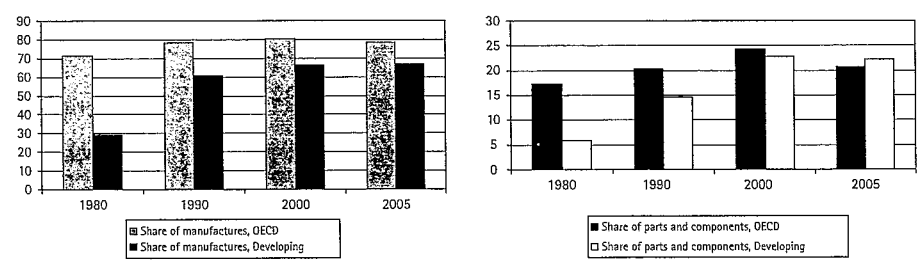

Fig. 1.4. Share of manufactures and parts and components in total exports, 1980–2005

Source: UN Comtrade database.

mostly trade in manufactured products. Agricultural products accounted for only 8 per cent of world trade in 2006—with the three largest exporters being developed economies: the EU, the US and Canada.

The rapid increase in the share of trade in manufactures, and within this, trade in components and parts, is one striking illustration of the process of globalization of production (Box 1.1). Among developing countries, East Asian economies took the lead in specializing in labour-intensive manufactures. Initially concentrating on simple products such as garments and footwear, these countries now produce a

Box 1.1. Changes in Global Production Sharing

The geographic fragmentation of manufacturing processes has long been a feature of world trade. One of its earlier forms involved the production of primary commodities in developing (and some developed) countries, shipment of these goods to (largely) industrial nations for further processing, and then the re-exportation (in part) of the processed product back to the primary commodity producer or third countries. For example, tin ore might be mined in Thailand or Malaysia, shipped to Japan for refinement and further manufacture and re-export. Such production sharing trade flows were based in part on comparative advantage, but policies—such as tariff escalation (tariffs that increase with the degree of processing of a good)—also contributed to this pattern of exchange.

The magnitude of such traditional production sharing trade has been eclipsed by international exchange of manufactured parts and components. A comparison of the value of East Asian trade in traditional inputs—agricultural raw materials, ores, minerals and nonferrous metals and unprocessed foodstuffs like cocoa and coffee beans—with manufactured components reveals that in 1984, Asian imports of traditional inputs were more than double those of manufactured components. By 1996, component imports were US$67 billion higher. Ten years later, at $781.8 billion, imports of components were over three times larger than traditional inputs. At 15.1 per cent, the average annual growth rate for component imports over this period was almost double that of traditional products. A similar pattern is observed for East Asian exports during this period, the ratio rising from 2 to over 8 (see Box Table).

Trade flow and product	Value (US$ bn)			Growth (%)
	1984	1996	2006	1984–2006
Imports				
Traditional inputs	39.2	98.9	247.4	7.9
Manufactured components	17.9	165.6	781.8	15.1
Exports				
Traditional inputs	17.0	27.6	110.0	8.2
Manufactured components	33.0	177.8	942.4	13.1

Source: 1984 and 1996, Ng and Yeats (1999). 2006 data provided by Francis Ng.

diversified mix of manufactured goods and participate very intensively in the process of global production sharing. East Asian global exports of components grew at an annual rate of 15 per cent during 1984–2006, more than four percentage points above the growth rate for all trade. Exports of components to other East Asian markets grew even faster (about 21 per cent per year). As a result, the share of all parts and components exports destined for regional markets almost doubled from 25 to 46 per cent. The corresponding figures for Latin America, for example, were only 17 per cent and 14 per cent respectively (Aminian, Fung and Ng, 2007). This illustrates another phenomenon—the increasing pace of regionalization of dynamic economies. The 2008–9 global recession revealed that fragmentation of production increases the sensitivity of trade to income: as demand fell, trade fell much faster. Regional vertical specialization did not insulate countries from the business cycle.

Although potentially incompatible with the process of globalization of production if associated with formal preferential trade agreements that create red tape such as restrictive rules of origin (see Chapter 10), the East Asian experience illustrates that regional integration reinforces the process of globalization if driven by market forces and complemented by openness to trade with the rest of the world. Trade expansion and the growth in production sharing have played an important role in the reorientation and expansion of trade between Central and Western Europe, and have also been increasing for a number of Latin American and North African countries, although in these regions the share of intra-industry trade is much lower than it is for East Asia and Central Europe.

An implication of the rising share of manufactures in global trade is that the factor content of trade has changed. As recently as the late 1980s many developing country exports were predominantly natural resource and unskilled-labour intensive. In all regions with the exception of the Middle East, a counterpart of the increase in production and exports of manufactures has been greater use of technology and skilled labour (human capital). Overall the skill-intensity of production and trade has been rising in both developed and developing countries, including in Sub-Saharan Africa (Table 1.1). This does not fit the prediction of the standard trade model where international exchange is driven by differences in factor prices that in turn are a reflection of differences in endowments of countries.

The 'standard' prediction from endowment-based theories of comparative advantage (Heckscher-Ohlin) is that as OECD countries have a more educated and skilled labour force, they should specialize in products that use such factors relatively intensively. The relative prices of goods that use less skilled labour more intensively should then fall as trade is liberalized (and those of skilled goods increase), which in turn should reduce the relative wages of the factors used in producing these goods domestically. At the same time, as unskilled labour-intensive activities are downsized and relative wages fall, there should be an

Table 1.1. Factor intensity of merchandise exports, 1988 and 2006

	Natural resources		Unskilled labour		Technology		Human capital	
	1988	2006	1988	2006	1988	2006	1988	2006
Industrial countries	24.4	23.1	9.4	7.3	38.9	44.6	27.3	25.0
Developing countries:								
East Asia & Pacific	29.0	15.1	29.2	17.3	23.4	50.6	18.5	16.9
Europe & Central Asia	59.4	46.5	15.1	10.3	14.0	20.7	11.5	22.5
Latin America & Caribbean	67.7	54.6	5.7	5.1	13.1	21.6	13.5	18.7
Middle East & North Africa	77.3	85.6	5.3	2.9	13.5	8.8	4.0	2.8
South Asia	51.0	42.5	35.1	24.5	8.0	16.8	6.0	16.2
Sub-Saharan Africa	74.0	72.3	19.5	4.2	1.9	11.7	4.6	11.8

Source: UN Comtrade and World Bank.

expansion in the demand for such labour in all parts of the economy. Conversely, developing countries should specialize in goods that use less skilled labour more intensively, as that is a factor with which they are well endowed, implying that liberalization should boost unskilled wages. It turns out that neither the product price effects nor the economy-wide expansion in unskilled labour intensity is observed in the data. Indeed, a large literature has concluded that the observed rise in the capital and skill-intensity of production and trade is mainly a reflection of technical change that is 'biased' towards (benefits) the more skilled. Skill-biased technical change (SBTC) is one explanation for the declining share of unskilled labour in total exports of almost all countries.

The increasing fragmentation or splintering of the production chain has been driven in part by technological changes that have lowered the costs of communication and transport, as well as by rapid growth in FDI flows. The latter in turn are in part also a result of the information and communication technology (ICT) revolution, which greatly facilitates control and communications and the far-reaching liberalization of policies towards FDI in developing countries, as well as other policy reforms that have improved the investment climate (Figure 1.5). Although data suggest that starting in 2005, policies towards FDI may have started to become somewhat more restrictive on the margin, the dominant trend has been for governments to take action to reduce barriers to FDI. The response to both technological changes and policy reforms was spectacular. The global value of stocks of FDI rose sixfold between 1990 and 2006, substantially faster than the growth in trade, which increased 'only' 35 times over the same period. As discussed in Chapter 7, much of this FDI is associated with services and is in part a response to decisions by many governments around the world to privatize state-owned utilities (such as telecommunications) and to open access to foreign provision of

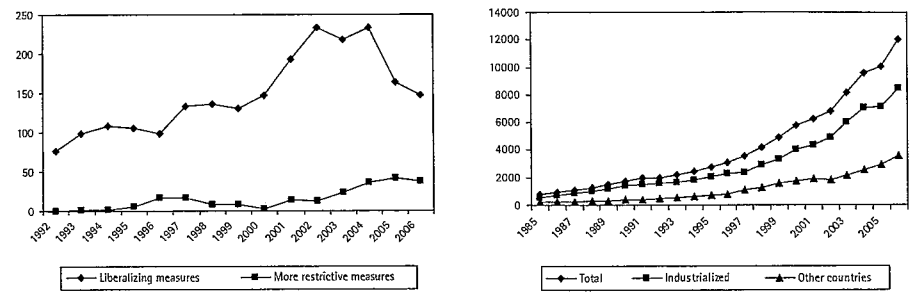

Fig. 1.5. Changes in FDI policies (number of measures per year) and value of FDI stocks, US$ million

Source: UNCTAD (2007).

services. However, much of it is in natural resources and manufacturing and related activities—such as energy services and distribution.

Foreign direct investment in manufacturing-related activities may be either horizontal or vertical in nature. Vertical FDI involves the establishment of facilities that specialize in specific parts of the production chain, with location of affiliates depending on the comparative advantage (comparative costs) of the host country. Horizontal FDI entails a firm essentially replicating plants that produce similar goods, implying that FDI and trade are substitutes for the parent firm. An example of horizontal FDI is the creation of Japanese car factories in the US the 1980s and 1990s, driven by protectionist American trade policies (so-called tariff jumping FDI). An example of vertical FDI would be BMW setting up a plant in South Africa to produce car seats that are incorporated into the cars it makes in Germany and other parts of the world. Vertical FDI often entails outsourcing energy- or labour-intensive parts of the production process to affiliates in countries that are well endowed with the required inputs. In principle such production could also be outsourced to unaffiliated firms located in such countries, but these may not exist or, if they do, lack the capacity to satisfy the required production standards. There may also be concerns regarding the ability to enforce contracts or potential leakage of technology that induce multinational firms to keep production 'in-house'. The relative importance of horizontal and vertical FDI varies depending on many factors. On average horizontal FDI tends to be relatively more important for developed countries, reflecting their similarity in factor endowments. Vertical FDI tends to be relatively more important for North–South FDI flows.

Of the US$12+ trillion global trade in goods, a significant share is intra-firm, involving flows between affiliates and parent firms. In the case of the US, some 45 per cent of total merchandise imports are intra-firm. Such trade is highly

correlated—by definition—with FDI flows. For example, Bernard et al. (2007) note that only 2 per cent of US imports from Bangladesh in 2000 were intra-firm, as compared to 75 per cent of imports from Ireland and Japan. In the case of Ireland the trade is driven by exports from US-owned affiliates, whereas in the case of Japan it reflects imports by Japanese-owned affiliates that have been established in the US. US imports from China are still mostly arms-length—intra-firm transactions accounted for 18 per cent of total imports in 2000. Understanding the determinants of observed differences in FDI inflows, and the magnitude of intra-firm trade versus arms-length outsourcing and offshoring of tasks has become a major focus of research in international economics.

Global integration and growth

The expansion of trade and FDI are just two dimensions of the multifaceted process of global integration that has been occurring in recent decades. Cross-border trade and investment flows have been a major engine of the process—'machines' that allow countries to transform one set of goods and services into another set that they value more highly. The increase in trade openness and cross-border investment is beneficial to the world as a whole. Empirical research by economists has shown a significant positive relationship between openness and economic growth (e.g. Greenaway, Morgan and Wright, 2002). Figure 1.6 charts the association between trade and growth by looking at a sample of 72 developing countries, divided into two groups: 'globalizers'—the 24 countries that saw the greatest increase in their ratio of trade to GDP between the mid-1970s and mid-1990s—and 'nonglobalizers'. Globalizers grew faster than both the nonglobalizers and the industrial countries over the period considered.

In a widely cited article, Sachs and Warner (1995) conclude that open developing countries grew by an average of 3.5 percentage points faster than a comparator group of closed economies. The sustained rise in post-Second World War European export–output ratios (compared to the pre-war period) was associated with a sustained increase in the average growth rate (Ben-David and Loewey, 1997). Sustained economic growth is critical in lowering poverty rates. East Asia, the developing region that has relied on trade the most intensively as part of its development strategy, has seen a number of countries catch up with the industrialized nations in terms of per capita income and significantly reduce the number of people living in poverty.

There is a vigorous academic debate on the relationship between openness and growth. For example, Dani Rodrik has argued that data and methodological weaknesses do not allow strong conclusions to be drawn (Rodrik, 1997; Rodriguez and Rodrik, 2001). Sceptics agree there is a positive association between openness and growth, but are not convinced the direction of causality is correct: they argue

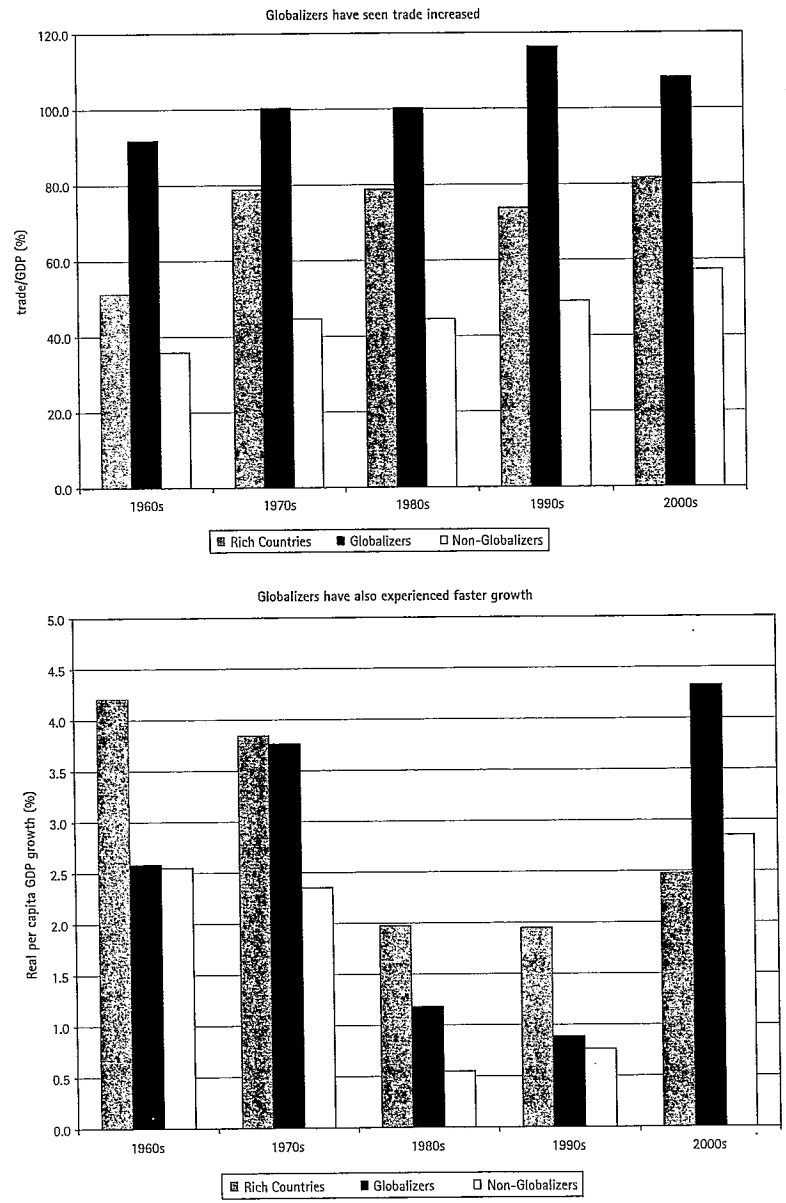

Fig. 1.6. Trade and growth go hand in hand

Source: Updated data, following Dollar and Kraay (2004).

that it may be it is growth that leads to openness rather than the other way around. One approach to disentangling causality is to examine the effects of those dimensions of openness that cannot be influenced by policy—that is, factors that are truly exogenous. Finding such factors is not easy. Examples are population, land area, borders and distances between countries. However, as noted by Winters (2004), such variables could have indirect effects on growth. Thus, geography may influence health, endowments or institutions, any one of which could affect growth directly as well as via trade performance.

While the technical academic debate continues, Frankel and Rose (2002) have shown that openness does indeed play a role even after allowing for geography and its possible indirect effects, while Wacziarg and Welch (2008) demonstrate that even if the cross-country evidence was not robust over the 1990s, if one takes countries' histories individually, the dates of trade liberalization do characterize breaks in investment and GDP growth rates. Specifically, for the 1950–98 period, they find that countries that liberalized their trade (raising their trade-to-GDP ratio by an average of 5 percentage points) enjoyed on average 1.5 percentage points higher growth in GDP per year compared with their pre-reform growth rate.

Because of the methodological problems that affect cross-country empirical analysis of the effects of trade liberalization, research has also focused on the *indirect* channels through which openness can affect growth (Winters, 2004). For example, liberalization will increase the variety of imports available to firms and households, allowing for a better match with the preferred characteristics of goods that are sought by buyers. Opening may also increase the variety of exports through greater specialization of production aimed for the world market. Both can contribute to productivity growth and, hence, aggregate growth. Feenstra and Kee (2004) show for a sample of 34 countries that more than 50 per cent of country productivity differences in the 1982–97 period can be explained by the differences in industry export variety.

Openness is also an important channel for the diffusion of knowledge. The research and development that is embodied in imported goods—especially capital goods—offers a specific mechanism through which the total factor productivity of an economy can be increased. More open economies tend to be more innovative, because of greater trade in knowledge (a greater quantity and variety of information, ideas and technologies associated with product and process innovations), and because greater competition spurs innovation, leading to higher rates of capital accumulation and productivity growth.

Empirical research has found that more productive firms are innately better at exporting, so that opening an economy leads to their growth and the demise of the least-productive firms. In the short run greater openness to trade and FDI will imply that local firms lose market share to foreign producers, impeding them from exploiting economies of scale and thus lowering their productivity. As less efficient

firms are forced to shut down, survivors will lower their cost base and/or upgrade their production processes. Over time, therefore, the average productivity of domestic firms will increase.

The existence of such competition effects has been documented by many empirical studies. Pavcnik (2002) is a representative example focusing on the trade liberalization undertaken in Chile. This involved the abolition of most NTBs and a reduction of import tariffs from over 100 per cent for some products to a uniform rate of 11 per cent across all industries. Using firm-level panel data, Pavcnik found that the productivity of plants in the import-competing sectors grew 3–10 per cent more than in the nontraded goods sector, which suggests that the exposure to international competition forced previously shielded plants to improve their performance. Exiting plants were on average 8 per cent less productive than plants that continued to operate.

Thus, openness leads to better exploitation of comparative advantage in terms not only of industries but also of firms within each industry. As noted by Anderson and Winters (2008), if the more productive firms are also foreign owned—as often will be the case given that foreign firms must be efficient enough for FDI to be worth the extra fixed costs involved in establishment abroad—then being open to FDI multiplies the gains from trade openness. Lower productivity firms may also gain, as the fixed costs of investing in newly opened foreign markets may become justifiable as a result of the prospect of larger sales volumes that come with exporting. Lower foreign tariffs should induce these firms to simultaneously export and invest in productivity improvements. In short, trade liberalization—both at home and abroad—can lead not just to a one-off increase in productivity and efficiency gains but also to higher rates of capital accumulation and productivity growth in the reforming economy because of the way reform energizes entrepreneurs.

No country has developed without being open and engaging with the world economy. Those that did not—such as India and China earlier in the twentieth century—paid a very high price in terms of continued poverty and low growth. Indeed, the importance of openness for growth is perhaps best illustrated by the experience of these two countries in the 1990s: liberalization and its associated integration into the world economy was associated in both cases with very significantly higher and sustained growth rates.

Managing global integration

Globalization, like any major technological change, gives rise to adjustment costs. It opens countries to exogenous shocks such as the 2008–9 global recession, which had major implications for trade because of the greater specialization in tasks. Global integration has cultural and social ramifications as well as economic

dimensions, and these must be recognized and managed. On the other hand, there is an enormous opportunity for continuing to use global integration to reduce poverty, hunger and economic injustice. Attenuating the negative effects of integration in instances where there are cross-border spillovers and assisting disadvantaged groups within and across countries is an important task for governments. The gains from trade generate additional resources that can be used by governments to assist in achieving such noneconomic objectives.

Trade liberalization is not a panacea that will automatically generate large growth benefits. Ben-David and Papell (1998) examine the post-Second World War growth path of 74 countries and conclude that 46 experienced a significant slowdown in economic growth rates during the period, even though openness ratios were rising. Relatively few countries have been able to attain and sustain growth rates that were high enough to result in convergence with the per capita income levels of industrialized nations. Indeed, country groups with the largest income gaps in 1960 have not shown any 'catch-up' convergence. Income gaps between many countries appear to be increasing rather than decreasing. Reasons for this are complex, but one common factor that characterizes incidences of convergence is the intensity (depth) of trade integration. Countries that trade intensively with each other tend to exhibit a relatively high incidence of income convergence (Ben-David and Loewey, 1997).

Of course, there is much more to growth than trade and trade policy. Greater trade integration is associated with faster growth, but complementary measures are needed to realize its full potential—including management of fiscal and monetary policy, public investment in human capital (education) and infrastructure, and the quality of public and private sector governance and contract enforcement. Of particular importance is that the overall policy regime is not biased against exports—which includes ensuring a competitive real exchange rate. For example, the magnitude of the beneficial knowledge spillovers from trade and inward FDI discussed above depends importantly on skill levels in importing countries and the ability to absorb and adapt the technologies. Much also depends on geography. Small landlocked countries surrounded by other low-income countries will inherently face much greater challenges than countries that are in close proximity to large industrialized economies.

Thus, for trade reforms to promote growth through improved price incentives, other policies are needed to ensure that investment flows into internationally competitive sectors and that weak or corrupt institutions do not undercut the positive incentives. For liberal trade policies to have a sustained effect on growth, they must be combined with other policies that encourage investment, allow effective conflict resolution and promote human capital accumulation. Competition policy, broadly defined to include the labour market and the regulation of entry, is particularly important: firms must be able to enter new areas and to exit declining, unprofitable activities. Research by Bolaky and Freund (2008) shows that

increased openness to trade is associated with a *lower* standard of living in economies with high barriers to the creation of new firms and restrictive labour market regulation that inhibits hiring. This is because such policies may prevent the expansion of the most productive firms by inhibiting the downsizing or exit of less profitable industries. As discussed at greater length in Chapter 7, policies affecting the performance of service sectors are also important in this regard—the availability, quality and cost of services are a key determinant of the competitiveness of firms.

Despite the fact that the term globalization is used incessantly, the world economy still is far from being integrated. Science, technology and a growing component of cultural life have become genuinely borderless. Advances in telecommunications and informatics industries and steadily decreasing transportation costs reduce the tyranny of distance. English has become the first second-language of the world. That said, borders continue to exist. Intranational transactions in goods and services continue to be a multiple of international transactions for all countries. This holds even for tradable goods. Many researchers have found that the so-called border effect is a robust feature of the world economy. Even in the case of what should be markets that are highly integrated—such as Canada and the US; or individual members of the EU, where governments have abolished trade barriers—bilateral trade is much less than equivalent trade within each country.[2]

Although barriers to trade and investment have been declining, for many sectors and activities policy continues to discriminate against foreign producers. Most glaringly, the global economy remains characterized by severe restrictions on the international movement of labour and highly distorting agricultural support policies in OECD countries. There are few signs that government policies are becoming significantly more welcoming towards liberalization of the temporary movement of service providers, let alone reducing barriers to labour mobility more generally. The unwillingness to liberalize trade in agricultural products was a major factor impeding more rapid progress in the Doha Round negotiations. Truly global industries such as electronics and aerospace coexist with a large set of industries that retain a regional or purely national character. Many of these are services that despite advances in technology remain effectively nontradable, and thus require FDI for international provision to occur.

[2] The classic paper in this area is McCallum (1995), who found that trade between Canadian provinces was some 20 times larger than trade between Canada and the United States. Subsequent research suggests that estimates of the border effect and 'home bias' are very sensitive to the accuracy of the characterization of internal trade costs and depend on the size of countries—the larger the country the lower the border effect. Most analysts find 'border effects' in the range of 5 to 15. Although these large effects cannot be attributed to the existence of a border alone, they do provide a measure of the extent to which markets are not integrated.

1.2. Trade Policy and Trade Agreements

International trade has been a feature of the world economy for millennia. The volume and pattern of whatever trade has taken place has largely been determined by trade costs. Such costs are in part physical—starting with the technical feasibility of transporting a good from A to B and, if feasible, the cost of doing so—and in part 'financial'—a function of the taxes or tribute that must be paid to those with the power to levy them, and the probability of complete expropriation (through theft and piracy) or loss (due to breakage, spoilage or natural calamities).

Technology and power have been the major forces determining trade flows over time, defining at any point in time the ability of regions to exploit their comparative advantages. Technological and institutional innovations that reduced transaction costs have had enormous impacts on what can be and is traded. Major innovations included 'hard science' inventions such as the sailing ship, the steam engine, development of railroads, aircraft, container shipping and refrigeration, as well as 'soft' inventions such as mechanisms to extend credit to traders and the development of contracts and procedures to enforce them.

As stressed by Findlay and O'Rourke in their excellent survey of the history of world trade since the Middle Ages, the exercise of power has had equally important impacts on trade:

> ... the greatest expansions of world trade have tended to come not from the bloodless tâtonnement of some fictional Walrasian auctioneer but from the barrel of a Maxim gun, the edge of a scimitar or the ferocity of nomadic horsemen. When trade required more workers, [these] could always be enslaved. When trade required more profits, these could be earned via plunder or violently imposed monopolies. For much of [history] trade can *only* be understood as being the outcome of some military or political equilibrium between contending powers... Politics thus determined trade, but trade also helped to determine politics, by influencing the capacities and incentives facing states. (Findlay and O'Rourke, 2007: xviii–xix)

From a historical perspective the policy stance advocated by many economists—unilateral free trade—has been applied relatively rarely, most notably by Great Britain in the second half of the nineteenth century. As free trade has been the exception rather than the rule, it is not surprising that trade agreements between sovereign states have frequently been used to overcome barriers to trade. Even in the case of imperial expansion and the pursuit of formal or informal empires by metropolitan powers, trade agreements sometimes were an important instrument. Examples in the nineteenth century were trade treaties negotiated between Britain and Latin American countries such as Brazil and Argentina (Gallagher and Robinson, 1953). Sometimes trade agreements have been a key element in the process of economic integration of independent territories—a noteworthy example was the German customs union (the Zollverein), which was a key building block of what is now the Federal Republic of Germany.[3]

A characteristic of colonial expansion was the application of metropolitan systems of law and protection of property rights to 'associated' territories—indeed, a defining characteristic of an empire is that control extends beyond foreign to domestic policy (Doyle, 1986). This was a fundamental dimension of the Roman Empire and helped create the pre-conditions for a single, integrated economy. Piracy was suppressed, roads built, and with sea and land routes substantially secure, commerce spread throughout the Mediterranean. The pottery, bronze, wine and oil of Italy were exchanged for African grain and eastern spices. Economies of scale led to large productive enterprises scattered throughout what was otherwise an overwhelmingly agricultural world (Gibbon, 1776).

The Iberian, Dutch and English empires of the sixteenth century and thereafter were of a different character in that the depth of integration was less. More important were discriminatory trade policies that sought to monopolize trade or to restrict competition. For example, not able to compete with more efficient Dutch shipping technology and constituting a less attractive market for some colonial products, seventeenth century England imposed trade restrictions on its colonies. The trade of American colonies was often subjected to exclusivity requirements—through a ban on trade with other states or through mandatory use of metropolitan shipping services—and regulated through restrictions on colonial production. Often, regulations prohibited local processing of goods or production of goods that could compete with output produced by the colonial power (Davies, 1997).

Trade relations between European powers and Asian territories initially tended to be less dictated by the former, reflecting more powerful local states. The latter produced goods (such as spices) that were sought after in European markets, forming a natural basis for trade. Often European traders sought to obtain agreement on (or to impose) extra-territorial application of home country law to commercial transactions and the protection of property rights. Local rulers who sought to limit the impact of a foreign presence on their control of society frequently were willing to accept such extra-territoriality. One form this took was through establishment of so-called treaty ports. Examples were Macao, Nagasaki and Goa. These served as an 'air lock' between international commercial relations and the control of civil society more generally:

[3] Keller and Shiue (2007) analyse the impact of removing borders between German states participating in the Zollverein. They conclude that this had a significant impact in integrating markets (as measured by convergence in prices of grains), but that much of the integration was also due to changes in technology—specifically the introduction of railroads connecting the various states.

From St. Paul's claim of civis Romanus sum against the subordinate patrimonial kingdom of Herod and the steelyard of the Hanse in London to the immunities of European settlers in Alexandria, Tunis, Constantinople and Shanghai, foreign powers have demanded extra-territorial application of their law over their nationals (both natural and legal persons). The outcome has often been the establishment of a regulated treaty port. (Doyle, 1986: 202)

Trade cannot prosper without legal security of property rights and mechanisms to enforce contracts. One lesson from international trade relations between states throughout history is that traders will seek to ensure that such mechanisms are applied. This can be achieved through a variety of means—full-fledged integration into a formal empire being the most far-reaching one; and free trade agreements and treaty ports between sovereign states being alternative solutions.[4] At an even broader and more general level history clearly shows that geopolitical stability matters critically for trade—political turmoil and major conflicts are associated with a decline in trade. Periods of hegemonic dominance have been associated with trade expansion because of the associated decline in uncertainty and trade costs, as the hegemonic power controlling a specific set of trade routes or region provided the peace and security as well as the institutional infrastructure needed to enforce contracts and protect property rights. In more recent times, this infrastructure has been provided in part through explicit cooperation between states, the GATT/WTO being one important vehicle for such cooperation.

From the perspective of exporters it is of little import what motivates a government to restrict trade. What matters is to induce governments to lower trade barriers. Abstracting from the exercise of military force or the threat thereof, formal trade agreements generally are the tool that is used to do so. As mentioned, the alternative—convincing governments to adopt a unilateral free trade stance—has only rarely been observed, making moves by countries to voluntarily pursue unilateral trade liberalization in recent decades somewhat exceptional in historical perspective.[5] As already mentioned, a major exception in the nineteenth century was Great Britain. It repealed its so-called Corn Laws in 1846 (which restricted imports of wheat and other grains and had been imposed in 1815 at the end of the Napoleonic wars) and moved to essentially a unilateral free trade stance at home

[4] There is an interesting literature exploring the emergence and maintenance of legal norms in the absence of central authority. A conclusion that emerges from these studies is that the threat of ostracizing a member of a club who is reliant on repeated interaction with other members can have a powerful impact as an enforcement device. Government involvement in contract enforcement is not necessarily required. Milgrom, North and Weingast (1990) and Greif (1993) discuss historical examples. Similar dynamics have been shown to prevail in modern Sub-Saharan Africa—see Fafchamps (2004).

[5] Many countries and colonies had low tariffs in the nineteenth century, but this was essentially imposed on them by the hegemonic/colonizing powers (Bairoch, 1989).

and in the overseas territories it controlled. This free trade policy applied to all sources of supply, not just British goods.[6]

Other major powers also liberalized trade during this period, but did so through the negotiation of trade treaties. The conclusion of the Cobden–Chevalier Treaty between Britain and France in 1860 created the equivalent of a free trade zone between the two countries and was followed by a series of trade agreements. During 1862–7, France concluded commercial treaties with virtually every major trading power in Europe (with the exception of Russia) as well as with the United States. All these treaties included a most-favoured-nation clause, following the lead of the Cobden–Chevalier agreement. As in each case the countries involved also negotiated treaties with each other and Great Britain, the trade concessions granted were multilateralized. As of the late 1860s, France was at the centre of an impressive network of trade agreements that substantially reduced protectionist trade barriers throughout Europe (Curzon, 1965). Average tariffs in Europe fell to some 9–12 per cent in the mid-1870s as a result of these treaties (Bairoch, 1989).

A key outlier during this period was the United States, which maintained high tariffs on manufactures to support its industry. Much of this industry was located in the North of the country, which implied that the agricultural sector—concentrated in the South—effectively was obliged to transfer a share of its income to the North as it was forced to pay more for machinery and consumer goods. This is an example of trade diversion that can be associated with the formation of a customs union—see Chapter 10 and Annex 2. A doubling of average tariffs in 1861 to 47 per cent helped set off the civil war: an objective of the South was to escape tariffs through secession from the Union (Adams, 1993: 330).

The nineteenth century was the period during which much of the intellectual debate about free trade emerged. There were two clear camps. Those in favour of free trade included Adam Smith (*The Wealth of Nations*, 1776) and David Ricardo (*On the Principles of Political Economy and Taxation*, 1817). Others argued that trade barriers were required to support infant industries. Influential contributions here were Alexander Hamilton's *Report on the Subject of Manufactures* (1791) and Friedrich List's *National System of Political Economy* (1841). The ideas of Smith and Ricardo on the benefits of free trade and the principle of comparative advantage provided the intellectual support for the free trade movement in Europe—both on the European continent and in Britain. Writings by Hamilton and List constituted a source of inspiration for those who favoured protection of infant manufacturing industry in the United States and Germany respectively. As is often the case, there was a time lag between the development of the theories and government action inspired by them. The British free trade movement emerged

[6] British industry helped enforce this free trade stance. For example, when the British government in India attempted to impose a small revenue tariff in 1853–4, the British textile industry ensured that an equivalent excise tax was levied on Indian textiles (Doyle, 1986: 264).

half a century after the publication of Smith's works. Full-fledged US infant industry protectionism materialized a quarter of a century after the publication of Hamilton's Report.

Despite the rise of infant industry protection in the major powers during the latter part of the nineteenth century, the global economy became significantly more integrated. Global trade expanded much faster than global output, driven by major reductions in transport costs as a result of technical changes (railroads, steamships), increased demand for commodities such as cotton, and large-scale migration into the Americas. This expansion in trade and factor flows generated significant adjustment pressures. In the case of Britain, for example, the rapid growth in New World agricultural production and exports led a large decline in the profitability of British agriculture. Real land rents fell by over 50 per cent between 1870 and 1913 (Findlay and O'Rourke, 2007: 396). Although industrialists benefitted greatly from both the rise in industrial output and the increased demand derived from export opportunities, and producers of inputs in trading partners profited from demand for their goods, British and European agricultural interests lost. The resulting lobbying for protection led to gradually increasing protection of agriculture on the European continent. Average tariffs rose from essentially 0 to 20/40 per cent between 1880 and 1910 in countries such as France and Germany. However, Great Britain maintained its free trade stance until the outbreak of the First World War.

After the First World War, restrictive trade policies became the norm. To some extent this was in response to the United States, which was unwilling to participate in efforts during the 1920s to re-establish a more open global economy following the disruption to trade that had been caused by the war and war-time policies. As the US economy moved from recession to depression following the 1929 stock market crash and subsequent monetary policies, the US Congress adopted the infamous Smoot-Hawley Tariff Act, raising average US tariffs on dutiable imports from 38 to 52 per cent. This led US trading partners to impose retaliatory trade restrictions and engage in rounds of competitive devaluation of their currencies. A domino effect resulted, as trade flows were diverted to relatively unprotected markets, forcing down prices, giving rise to protectionist pressures there, and thus leading to higher trade barriers.

At the end of the Second World War, statesmen such as Presidents Roosevelt and Truman and, particularly, Cordell Hull, the US Secretary of State, were deeply influenced by the lessons of the post-First World War period. They perceived the need for establishing cooperative mechanisms to avoid both competitive devaluation and the excessive use of trade barriers to guarantee the national market to domestic producers (Gardner, 1969). The negative consequences of the beggar-thy-neighbour policies of the early 1930s were still very vivid in 1945. They inspired the US willingness to pursue the type of international cooperation it had spurned in the 1920s and early 1930s and actively support multilateral liberalization efforts,

including efforts to negotiate the International Trade Organization (ITO) and the GATT. In the Anglo-American view, the post-war international economic system was to be constructed in such a way as to remove the economic causes of friction that were believed to have been at the origin of the Second World War. An important element in this vision was the establishment of a stable world economy that would provide all trading nations with nondiscriminatory access to markets, supplies and investment opportunities.[7] There was a strong perception that there was a positive correlation between trade and peace, and, as important, between nondiscrimination and good foreign relations (Bailey, 1932).[8] In the US, the Reciprocal Trade Agreements Act of 1934 had already initiated a shift to a more liberal trade policy stance through the adoption of the unconditional MFN principle, albeit firmly grounded in the principle of reciprocity. This policy was extended after the Second World War and incorporated into the draft charter of the ITO and the GATT.

1.3. Functions of the Multilateral Trading System

Multilateral cooperation among sovereign nations often occurs through the creation of institutions. Because a central authority is absent in international relations, political scientists have developed the concept of a regime, defined as 'sets of implicit or explicit principles, norms, rules and decision-making procedures around which expectations converge in a given area of international relations' (Krasner, 1983: 2). The principles and procedures imply obligations, even though these are not enforceable through a hierarchical legal system. Regimes reflect patterns of cooperation over time among members that are based on the existence of shared interests. The multilateral trading system is a good example of a regime.

Two viewpoints are helpful in understanding the role of the trading system. The first is to regard it as a mechanism for the exchange of trade policy commitments. The second is to consider it as a mechanism through which the resulting code of conduct in implemented and enforced, that is, to focus on the result of the

[7] Although there were major differences between the US and the UK regarding the latter's insistence that the system of Commonwealth preferences be maintained.

[8] The academic literature on the relationship between trade and the probability of war has argued that this may go either way. For example, two countries that are on opposing sides of the globe and do not trade at all are less likely to go to war than two neighbouring states that trade a lot. However, Mansfield (1994) has concluded that, controlling for such factors, there is a robust negative relationship between the volume of trade between country pairs and the probability of a war between them.

exchange. Much of this book focuses on the outcome of negotiations and the disciplines that members agree to apply. What follows first briefly discusses the system as a forum for exchange, a subject that is explored in greater depth in Chapter 4. We then summarize the main elements of the system as a code of conduct: the nondiscrimination rule (MFN and national treatment), transparency, enforcement and flexibility (as exemplified by a variety of safety valves and vagueness in some disciplines).

The system as a market

The WTO is a forum for the exchange of liberalization commitments. That is, it is a market. Bargaining and negotiation are the main instruments used to reduce barriers to trade and agree to rules of behaviour. Multilateral trade negotiations are mechanisms through which governments exchange market access and other policy commitments.

In any country the structure of protection at any point in time is the result of the interaction between the demands expressed by various interest groups in society and the responses by governments and legislatures. Attempts to alter this equilibrium and move towards a national welfare-increasing reduction in protection will generate opposition by those groups that expect to lose from liberalization. Such losses are usually concentrated in import-competing industries, while the gainers—consumers of the products concerned—tend to be much more diffuse. This gives rise to a political economy problem. Those facing losses have a much greater individual incentive to organize and invest in lobbying against liberalization than those that gain from reform have to lobby for liberalization (Olson, 1965). Individual gains are relatively small and dispersed among a large number of voters, while losers are more concentrated. This is the main reason why trade restrictions are imposed in the first place.[9]

A MTN can solve this problem by confronting those who gain from protection with another lobby that may be equally powerful: the set of firms that benefit from greater access to foreign markets. Similarly, through reciprocally reducing trade barriers, the prisoners' dilemma that confronts large countries can be overcome, again improving world welfare. Moreover, by encompassing many products, a MTN can generate some automatic compensation for those who lose protection for their sector by lowering the average price of consumption and investment goods by providing access to cheaper imports.

A MTN is akin to a market in the sense that countries come together to exchange market access commitments on a reciprocal basis. It is a barter market. In contrast to the markets one finds in city squares, countries do not have access to a medium

[9] In developing countries without an effective tax administration, tariffs frequently have an important revenue rationale as well.

of exchange: they do not have money with which to buy, and against which to sell, trade policies. Instead they have to exchange apples against oranges: tariff reductions for iron against foreign market access commitments for cloth. This makes the trade negotiation market relatively inefficient, and is one of the reasons that MTNs can be a tortuous process.

Why do countries use trade policy?

To understand the role of the WTO as a market for the exchange of trade policies it is useful to first consider the rationale for trade restrictions. Motivations for activist trade policy can be divided into a number of types. First, revenue: governments need income, and taxing trade is often the easiest method of collecting it. Taxation of trade for revenue purposes has been a hardy perennial throughout recorded history, and remains important for many developing countries. Of course, those who are subject to the tax have an incentive to lobby for exemptions and invest resources to induce the authorities to lower the tax burden. Taxes imposed by rulers can constitute an important motivation for conquest or, more peacefully, for cooperation, such as the negotiation of tax treaties. Tax policy can have important effects on trade patterns. For example, in the fourth century BC, Rhodes was a major commercial power in the Eastern Mediterranean, controlling the neighbouring seas and with a vibrant port. Rhodes charged a two per cent tax on the value of cargo carried on all ships entering its harbour, including transit cargo. To divert shipping, Roman traders lobbied for the creation of a free port in Delos. Once established, trade rapidly shifted away from Rhodes, and the port lost most of its harbour tax revenues. This tax competition proved very costly from a social welfare point-of-view: Rhodes used part of its tax proceeds to police the sea-lanes and prevent piracy. Without the revenue, these activities declined, piracy increased significantly and trade became more costly (Adams, 1993: 83–4).

Another motivation for trade policy is to improve the terms of trade—the ratio of the prices they get for their exports and the prices they pay for imports. This rationale applies only to countries that have the power to influence world market prices because of their economic size or market power. Such 'large' countries can use trade policy either to reduce the prices of imports and/or to increase the prices of exports. (Large countries that use trade policy for revenue purposes will *ipso facto* affect the terms of trade as well.) Economic theories that allow for imperfect competition, product differentiation and increasing returns to scale have potentially expanded the number of situations under which countries can in practice affect their terms of trade. Thus, a country does not have to be 'large' in an absolute sense to be able to affect its terms of trade for a given product.

A third motivation is mercantilist—a belief that imports are bad and exports are good. This belief is generally based on the observation that imports must be paid for and thus imply the transfer of foreign exchange abroad (historically specie—gold or silver), whereas exports bring in foreign exchange. The objective of mercantilist policy is a trade surplus—ensuring that the value of exports exceeds the value of imports. Mercantilism is often driven by nationalism, the perception being that trade surpluses and political power are closely linked. Mercantilist policy therefore tends to favour direct promotion of exports and restrictions on imports through tariffs, quotas, prohibitions or state monopolies. The policy makes no economic sense. Starting with philosophers and economic thinkers such as David Hume, Adam Smith, John Stuart Mill and David Ricardo, it has been pointed out that imports are desirable and that exports are simply a way to pay for imports. Moreover, a trade surplus will have macroeconomic effects that will act to push the balance of payments back into equilibrium.[10] The theory of comparative advantage and gains from free trade was developed largely in reaction to mercantilist thought and practice.

Fourth, trade barriers frequently have been used as instruments for agricultural and industrial development. This was an important factor in the latter part of the nineteenth century, with continental European powers and the United States pursuing activist trade policies to protect infant industries. French colonies relied heavily on discriminatory trade policies such as tariff walls against the rest of the world, keeping British goods out of these markets. With France, Germany and the United States becoming increasingly industrialized, British trade dominance was eroded and British goods came to be diverted away from traditional export markets, initially the newly industrializing markets, and subsequently rest-of-the-world colonial territories. These policies eventually helped induce Britain to abandon the free trade policy it had adopted in the mid-1850s and begin to pursue preferential trade regimes with its own territories. This in turn led to the adoption of a system of imperial preferences that became a major bone of contention between the UK and the US in the negotiations on the GATT and ITO.

Finally, trade policy is a source of rents for specific groups in society. Protectionist policies have the effect of redistributing income from consumers of the affected goods to those that produce them or to those that control the right to import. By imposing barriers to trade, some segments of society gain at the expense of other groups. It is for this reason that protectionism can constitute good politics. It is a mechanism through which interest groups can be rewarded for political support in relatively nontransparent ways. Groups seeking protection from

[10] The fallacy of mercantilist thought regarding the need for a positive balance of trade inspired David Hume to develop his famous 'price–specie flow' mechanism. This illustrated the point that trade surpluses and associated inflow of specie would drive up prices and result in a loss of export competitiveness.

Box 1.2. Political economy drivers of trade policy

Economists have developed two broad types of analytical frameworks to reflect the fact that policy is endogenous—the result of a political process in which groups seek to maximize their utility or welfare. Both proceed by embedding either a voting or lobbying model of the political process into an economic model. The former often focus on the 'median voter', whose preferences will determine outcomes in two-party elections. The latter start from the presumption that interest groups will lobby politicians for specific policies that benefit them, and offer political (and financial) support for their election conditional on their preferred policies being pursued. These models help to understand why countries adopt policies that do not maximize national welfare: policies may not be economically efficient, but they are 'politically' efficient—they emerge as the equilibrium outcome of a specific political process.

There is strong empirical support for the view that trade policy formation is driven by political economy forces. However, empirical research on the political economy of trade policy has long had only a tenuous connection with underlying theoretical frameworks that generated clear predictions that could be tested. This changed with the development of a formal theoretical political economy framework by Grossman and Helpman (1994, 2002). This allows for formation of lobbies and is based on the simple precept of a government that maximizes a weighted sum of welfare (W) and lobbying contributions (C): $G = aW + C$, where a is the weight the government puts on a dollar of welfare relative to a dollar of contributions from special interests. Free trade would be the efficient outcome if the government maximized welfare, that is, if the objective function G admitted only W. In order to induce the government to set positive tariffs government must be compensated, via contributions, for the loss in consumer welfare weighted by a. A tariff t_i on good i raises its price p_i above the world market price, while an import subsidy lowers it. Assuming that individuals own capital that is specific to a sector, increasing the price of the good produced by that sector raises the return to the specific capital used to produce it. Owners of sector-specific capital in an import-competing sector thus have a strong incentive to politically organize and offer the government contributions in return for a tariff—with higher tariffs eliciting a higher 'payment (contribution). The Grossman-Helpman model yields a precise testable implication about the cross-sector pattern of protection and has generated a cottage industry of empirical applications and tests.

The model predicts that in politically organized import-competing sectors (those that form lobbies) trade protection is positively related to the ratio of domestic production to imports. The intuition here is that large domestic sectors make the largest lobbying contributions, while the lower the import volume, the lower the social cost of protection, thus diluting the opposition of consumers to tariffs for that sector. Thus, the model predicts there is a tradeoff between additional profits for specific factors employed in an industry and consumer surplus. Empirical studies find strong support for this prediction, but also conclude that governments appear to place great weight on social welfare: estimates of the a parameter are invariably very high, implying a weight on welfare that is 50–100 times greater than the weight given to lobbying contributions (Gawande and Krishna, 2004). Potential explanations for this are that there is likely to be substantial uncertainty about whether protection will in fact be delivered, thus lowering the effective impact of contributions, and that the Grossman-Helpman model does not take into account that many goods are inputs into the production of other goods, as a result of which lobbies work against each other (Gawande, Krishna and Olarreaga, 2005; Gawande and Hoekman, 2006).

import competition or the right to control imports often offer political support to the government (or to challengers in elections) as a quid pro quo (Box 1.2). Government officials may benefit directly from trade restrictions by capturing the rents associated with control over goods that can be sold in domestic markets at prices above their world market rate (cost).

It is often difficult to distinguish between the motivations for restricting trade. For example, trade policies that are part of an industrial policy may create rents and affect the terms of trade. There are similarities between mercantilism and infant industry protection—both have strong nationalistic connotations, and both rest on weak economic foundations (the economics of infant industry protection is discussed further in Chapters 5 and 9). However, in principle a major difference is that infant industry protection can (and should) be pursued in a nondiscriminatory manner. Given the objective of protecting local economic activity, this is most efficiently done in a nondiscriminatory way if governments decide to use trade policy instruments. Mercantilism in contrast is essentially bilateral in nature—what matters is the bilateral trade balance.

Historically, revenue considerations have figured almost universally—even free trade Britain imposed significant revenue tariffs. One implication is that one cannot necessarily determine from the average tariff or the magnitude of tariff revenue collections how high trade barriers are. What matters is the difference in the extent to which domestic and foreign products are taxed. If this difference is small, a country can be characterized as maintaining a liberal trade policy, even if tariffs are imposed.

Impacts of trade policy on welfare

From a national welfare perspective, the utility of trade policy depends largely on the market power of a country. A small country that cannot influence prices on world markets will generally lose from imposing trade barriers. Protection gives rise to both production and consumption distortions: producers confront artificially high domestic prices that encourage them to produce 'too much' of the protected products, while consumers consume 'too little'. Producers gain at the expense of consumers, and the deadweight losses associated with the transfer from the latter to the former imply that overall welfare is reduced. The elimination of these distortions is, therefore, a major source of the gains from liberalization (Box 1.3). Trade liberalization helps nations to realize a more efficient utilization of their resources (production capacities). Trade liberalization has two essential effects. First, it brings about a reallocation of resources towards those activities in which the country has comparative advantage. The economy becomes more productive on average as those industries in which the country has a comparative advantage expand by drawing resources from previously protected or subsidized industries

Box 1.3. Gains from specialization

The central concept underlying trade is opportunity cost. Producing (consuming) something comes at the cost of not producing (consuming) something else. An important economic theorem states that there are gains from trade associated with minimizing opportunity costs through the division of labour (specialization). Consider a simple example. Suppose the people of Plains, who are good at raising animals (say cows), must also spend time growing wheat (at which they are less good than raising cows). Each hour spent growing wheat has a high opportunity cost in terms of cows forgone, but there is no choice but to devote the time required to grow wheat. Suppose the people of Agria are good at farming, but do not have much aptitude for raising cows. Agria will then have a high opportunity cost in terms of time not spent farming. If these two countries/groups of people could trade with each other, they could concentrate on what each one does best. Economists say that they would specialize according to their comparative advantage. This will ensure that total output produced expands in both regions, and that each is able to consume more wheat and beef and milk than would be possible without trade.

The decision what to specialize in depends on what one does best compared with the other things that could (or would have to) be done. The people of Plains might be better farmers than those in Agria, in that for every hour invested in farming they get a larger harvest. However, as long as an hour spent by the people in Plains on farming has a higher cost in terms of forgone cows than does an hour spent on farming in Agria, Plains should specialize in cows. What matters is not *absolute*, but *comparative* advantage. International trade provides nations with the opportunity to specialize in production according to their comparative advantage. A country may be better at everything than another country in absolute terms, but by definition it cannot have a comparative advantage in everything.

(which either grow more slowly or contract following liberalization). Second, trade liberalization expands the consumption opportunities of countries, as more efficient production generates greater income and increased opportunities to buy goods and services from other countries (see Annex 2 for a graphical illustration and brief discussion of the standard mechanics of the gains from trade and the effects of trade policies).

The inter-industry reallocation and adjustment process that is the basis of the standard theory of comparative advantage and the gains from trade is replicated *within* industries as well: the more productive domestic firms in an industry expand by drawing resources from less productive firms that contract or go out of business. Recent theoretical developments and empirical analysis have emphasized the importance of recognizing that there is much heterogeneity of firm performance and efficiency/productivity within industries, and that this is a significant source of the welfare gains from trade liberalization (Melitz, 2003). Many empirical studies have shown that much if not most of the adjustment

associated with trade reform involves shifting of resources within an industry rather than across industries (Hoekman and Winters, 2007). Recognition of the heterogeneity of firms within and across industries helps to understand this empirical observation, and helps to understand why trade liberalization is important for economic growth over time. As the more efficient firms expand and the less efficient ones contract and either exit or are taken over, the overall productivity of the economy increases. If there are scale economies and imperfect competition, liberalization will allow more efficient firms to further reduce unit costs as their market expands.

Recent theorizing that stresses heterogeneity of firms also helps to explain why some firms in an industry export while others only sell on the domestic market (Tybout, 2003). It also provides a much better understanding of the forces that result in intra-industry trade. If there are fixed costs associated with contesting international markets, only the more efficient firms will be able to export their products, and different firms will specialize in different varieties of (differentiated) products. Thus, liberalization, by allowing the more efficient firms to expand, not only will promote the overall export performance of an economy, but also much of the resulting trade will be of the intra-industry type. Consumers gain not just because of the elimination of the traditional production and consumption distortions but also because they get access to a much wider range of (differentiated) goods and services—many of which may not be produced at all in autarky.

As a result of the technological changes discussed previously and the resulting increase in scope to separate in time and space the various productive tasks along a value chain, including not just goods (components) but also services such as design, marketing and back-office administrative transactions processing, liberal trade policies allow firms to exploit factor cost differences across countries for specific tasks. Given that gains from liberalization are larger, the greater the variance of rates of protection across tasks, and that protection of some tasks or activities—e.g. services—is low or nonexistent (see Chapter 7), technological changes that permit trade in tasks increase the gains from liberalization of trade in goods, even if tariffs are relatively low (Anderson and Winters, 2008).

Motivations for international cooperation

In contrast to small countries, large countries may be able to change the terms of trade—the price of their exports relative to the price of imports—in their favour by restricting trade. However, for the world as a whole the imposition of trade restrictions by one or more countries can only reduce welfare. Large countries thus may find themselves in a so-called Prisoners' Dilemma situation: it is in each country's interest to impose restrictions, but the result of such individually rational

behaviour is inefficient (see Chapter 4). All countries end up in a situation where their welfare is lower than if they applied free trade policies.[11] Both small and large countries, therefore, have an incentive to cooperate and agree to reduce or abolish trade barriers. Trade and trade liberalization is a positive-sum game.

Although basic trade theory suggests that small countries that are price-takers on world markets and that want to maximize their wealth should not impose trade barriers, a major reason why free trade is rarely observed is that some groups in a society will gain from protection (at the expense of others). As costs of liberalization generally are concentrated in specific industries, usually those that have invested resources in (lobbied for) protection, they will oppose liberalization. Potential losers are concentrated and often already organized—as organization will have been required in order to obtain the protection in the first place. The overall benefits of a liberal trade regime are in the aggregate usually greater than losses accruing to those who gained from protection. However, these benefits accrue to a large and diffuse set of agents. On an individual or household level basis, the benefits of liberalization are in most cases small, creating only weak incentives for the potential winners to organize themselves politically. In principle, the losers can be compensated, as the removal of the inefficiencies caused by protectionist policy will, once the economy has adjusted, increase total output and consumption by more than the (transitional) losses incurred by those who must change the economic activity they are engaged in. Actually compensating the losers is not always easy, however, and in practice occurs only rarely, and, if so, is generally partial.

One reason for this is that compensation is difficult—governments may not have the instruments needed. Trade integration may affect the redistributive capacity of governments by changing the structure of the economy and, therefore, the tax base, and by affecting the distribution of political power. The capacity and willingness to provide for domestic redistribution and compensation cannot be analyzed separately from the decision to open the country to trade and foreign direct investment flows (Verdier, 2005). This suggests that policymakers may need to provide insurance mechanisms in order to secure national welfare gains. To minimize distortions, any such instruments should not involve manipulation of relative factor and goods prices (which, of course, is exactly what trade policies do). Examples of such instruments are lump-sum, one-off payments and mechanisms that provide insurance against declines in the value of key assets such as land and human capital. The latter is particularly important in rural communities as land values may be a primary base for local tax revenues, and thus the provision of public goods and services.

[11] That is, large countries need to take into account the possibility of retaliation. Another problem is that if tariffs are not set at the optimal level, large countries may easily lose from activist trade policy—even if other countries pursue free trade.

The imbalance in the strength of political forces favouring and opposing liberalization provides a possible rationale for the pursuit of reciprocal trade negotiations. Rather trivially, although a (small) country will benefit from liberalizing its trade, it is even better if trading partners do the same. More important from a political economy perspective is that by making liberalization conditional on greater access to foreign markets, the total gains of liberalization increase and in the process liberalization becomes more feasible politically. Being able to point to reciprocal, sector-specific export gains may be critical in mobilizing domestic political support for liberalization at home. By obtaining a reduction in foreign import barriers as a quid pro quo for a reduction in domestic trade restrictions, specific export-oriented domestic interests that will gain from liberalization have an incentive to support it in domestic political markets. This political economy rationale for reciprocal negotiations is now generally accepted as a basic explanation for the existence of trade agreements and the WTO.

Economists often stress the importance of the terms of trade in providing a theoretically consistent rationale for the formation of trade agreements. The argument is that countries negotiate away the negative terms-of-trade externalities that would be created by the imposition of trade restrictions in partner countries (Bagwell and Staiger, 2002). Questions can be raised regarding the empirical relevance of this explanation for small countries that cannot affect world prices (in the terms of trade sense). Part of the answer may be that most products that are traded are differentiated, potentially giving small countries some market power (as what matters is not the size of the country, but the degree to which the product(s) of the country are substitutable and the number and cost of alternative suppliers of substitutes). However, for low-income countries that export mostly commodities the empirical relevance of such product differentiation-based market power is likely to be very limited. More important, governments of a small country may want to be a member of the WTO because its exporters will benefit from the low tariffs that large WTO member countries negotiate reciprocally with one another but must then extend to all other members under the MFN rule.

This explanation can only be partial, however, because it does not explain why large countries want small countries to join the WTO. It may be that in practice large countries simply do not care, as small countries cannot affect the terms of trade. An implication is that trade agreements will tend to reflect the concerns of large countries, and that reciprocal exchanges of trade policy commitments will be concentrated among large countries. To a significant extent this is indeed what occurs. However, at the same time large countries have supported expansion of the membership of the WTO, and negotiated bilateral trade treaties and preferential access arrangements with small countries. This is difficult to square with the terms-of-trade explanation for trade agreements, suggesting other motivations must be relevant as well.

The term-of-trade rationale has also been criticized in the specific context of the WTO because the GATT does not discipline the use of export taxes, which can be used to affect the terms of trade (Ethier, 2001b, 2004; Regan, 2006). If terms-of-trade considerations were indeed the sole driver of trade agreements, governments would want to discipline all border policies that can influence the terms of trade. In the WTO this is not the case. Nor can terms-of-trade theories explain why small country governments negotiate limits on their own use of import tariffs and other policies when joining trade agreements.

Another strand of economic theory (e.g. Tumlir, 1985; Staiger and Tabellini, 1987; Maggi and Rodriguez-Clare, 1998, 2008) provides an alternative rationale: trade agreements may offer a mechanism to governments that want to commit to a set of policies that may not be (politically) feasible to adopt or maintain. This line of theory has trade agreements serving as a lock-in mechanism or anchor for trade and related policy reforms. By committing to certain rules that bind policies, a government can make its reforms more credible: officials can tell interest groups seeking the (re-)imposition of trade policies that doing so will violate their commitments and generate retaliation by trading partners.

This rationale for trade agreements is conditional on agreements being enforced. In practice, agreements may not be enforced against small countries because the incentives for trading partners to invest the required resources may be too weak, that is, costs exceed expected benefits. If this is the case—as is suggested by the evidence summarized in Chapter 3—this weakens the commitment explanation for cooperation, making it conditional on there being a terms-of-trade externality needed to induce compliance (generate the credibility). In addition, the large number of holes and loopholes that are embodied in the WTO weaken the credibility-cum-commitment that is implied by membership—as governments still have great leeway to (re-)impose protection. As is often pointed out in the economic literature on the WTO, it is an incomplete contract.

A third perspective on the rationale of trade agreements has been developed by Ethier (2004, 2007), who categorically rejects the 'real world' validity of terms-of-trade driven explanations. Ethier stresses that WTO members retain access to instruments through which they can affect their terms of trade, starting with export taxes—which, as mentioned, are not subject to disciplines. Instead, Ethier stresses domestic political economy dynamics, and builds on—is consistent with—a long tradition that starts from the premise that governments seek to maximize political support: their concern is to get re-elected or to remain in power. This in turn implies that they will respond to and seek to satisfy the domestic constituencies that they need to stay in power. Taking as given that governments are conservative in the sense that they put greater weight on prospective losses for groups in society than on the expected gains from liberalization (which is realistic as losers can be identified and will mobilize whereas many of the beneficiaries of greater exports do not know who they are), governments have incentives to impose or maintain

protection because this raises the incomes of the groups from which they derive political support. If foreign governments could be induced to liberalize, however, that provides a direct gain for existing exporters. This in turn changes the government's incentives as it affects the balance of political support. A more liberal stance becomes optimal as the government will benefit from reducing import tariffs on a quid pro quo basis (see also Grossman and Helpman, 2002).

An interesting distinct feature of Ethier's analysis is that it provides an explanation for *gradual* liberalization: trade reforms generate higher levels of political support if spread out over time. Gradualism is a standard feature of virtually all trade agreements, in that they tend to be implemented in stages. Usually this is explained on the basis of adjustment costs. Ethier (2004) offers another motivation for gradualism: it has a political support rationale.

Although the formal theoretical frameworks that have been developed by economists in recent years have helped clarify the possible rationales for trade agreements, the economic literature can only offer a partial perspective. Complementary explanations for the formation of trade agreements have been offered in the international relations and political science literature. These disciplines place more emphasis on the role of power, on domestic political considerations and the structure of institutions, and on 'noneconomic' objectives and values such as the avoidance of war and ideology. The stress on power and foreign policy considerations is clearly historically relevant given the impact of the exercise of power on trade flows (Findlay and O'Rourke, 2007). In practice, as stressed by the WTO (2007), the huge differences between countries and their underlying interests imply that there can be no single, formal 'grand theory' of the GATT/WTO.

In our view, although the terms-of-trade (market access-cum-cost shifting) framework is elegant and generates important insights into the factors that will support trade agreements, it is too abstract to help understand the actual process of multilateral cooperation on trade. The genesis of the GATT reveals rather unambiguously that terms-of-trade considerations did not drive negotiations or determine the final outcome (Curzon, 1965; Jackson, 1969; Dam, 1970). In practice, the political economy-based frameworks provide greater insights into the design and mechanics of cooperation in the GATT/WTO.

Reciprocity

For a nation to negotiate, it is necessary that the expected gain from doing so is greater than the gain available from unilateral liberalization. By obtaining reciprocal concessions, these gains are ensured (Box 1.4). More technically, what reciprocity in trade negotiations does is to help to offset the externalities (economic inefficiencies) that are imposed by countries as they implement trade policies,

Box 1.4. Political economy forces and reciprocal liberalization

Hillman and Moser (1996) argue that a useful way to understand the role of reciprocity is to start from the premise that import-competing industries have property rights to their home markets, a right that has been acquired as a result of past lobbying or political support granted to governments. In the same way that protection can be explained as the outcome of a political process where governments seek to maximize political support—taking into account the fact that tariffs are often used for revenue purposes and tend to persist after alternative tax bases are developed—reciprocal liberalization can be explained as the outcome of a political process. In this case the interests of the domestic right-holders (the import-competing industries) are balanced with those of domestic export industries seeking equivalent rights in foreign markets (and lower input costs). If the latter group offers enough political support, erosion of the former group's rights may prove politically rational. For a discussion of the resulting dynamics in the context of US trade policy, see, for example, Destler (2005) and Devereaux, Lawrence and Watkins (2006).

Whatever is offered by one country (the *demandeur*) in a MTN as a quid pro quo for a demand by a trading partner must be of interest to the government asked to alter its policies. Thus, to be effective the offer must help meet the objectives of influential foreign lobbies that will then push for the desired change in policy in their country. Alternatively, offers might be designed to help the government compensate groups that are likely to lose significantly from a reduction in protection. Options here include a gradual reduction in the level of protection and acceptance of safeguard mechanisms—as discussed below, two 'principles' that characterize the WTO.

Although export interests are the primary players in supporting liberalization in the MTN context, other groups favouring liberalization may also play a role. Examples include consumer or economic-development lobbies (the effect of development aid is frequently offset by protection against developing country exports, an example of incoherent policies to which we return in Chapter 12). To mobilize such groups they must be aware of the detrimental impact of trade policies on their objectives, and these impacts must be large enough to induce them to organize. The provision of information on the effects of protectionist policies is, therefore, of great importance. Indeed, the need for such information is quite independent of the MTN process, given that in many instances a unilateral change in policy would be welfare-improving. The main point, however, is that what counts is political support. If consumer and other groups favouring a liberal trade policy do not mobilize and exercise political influence, they generally will be irrelevant.

generally driven by a desire to respond to interest groups that seek protection and have supported the election (or selection) of a given government. In effect, by insisting on reciprocity countries may be able to ensure that their 'terms of trade' are not affected detrimentally as a result of own liberalization, in the process counterbalancing the resistance by losing lobbies with the support generated by those that benefit.

Reciprocity in trade negotiations comes in many guises. It may be diffuse or specific (Keohane, 1984). If specific, it may be expressed in quantitative or qualitative terms, and may apply to levels or to changes in protection (Winters, 1987*a*). Although the GATT and the GATS have as underlying goals a broad balance of market-access commitments, by requiring reciprocity, nations attempt to minimize free riding. In the case of bilateral negotiations, this is done by a suitable choice of products on which concessions are offered and sought; in the case of multilateral across-the-board negotiations, it is done by a suitable choice of products to be exempted from liberalization (see Chapter 4).

Generally, nations are quite successful in minimizing free riding. For example, internalization, defined by Finger (1974, 1979) as the sum of all imports originating in countries with whom a country exchanges concessions as a percentage of total imports of goods on which concessions are made, was about 90 per cent for the US in the Dillon (1960–1) and Kennedy (1964–7) Rounds. Allen (1979), focusing explicitly on bilateral bargains made in the Kennedy Round, showed that there was a relationship between the size of concessions made on commodity tariffs and the degree of bargaining power a country had on a commodity vis-à-vis its major trading partners. Thus, reciprocity is in part a function of the weight a country can bring to bear in a negotiation.

Reciprocity also applies when countries accede to the WTO. Given that new members obtain all the benefits in terms of market access that have resulted from earlier negotiating rounds, existing members invariably demand that potential entrants pay an 'admission fee'. In practice this implies not only that upon joining the WTO a country's trade regime must conform with the rules of the GATT, GATS and TRIPS, but also that the government will be asked to liberalize access to its market. Accession modalities are discussed further in Chapter 2.

For reciprocity to work it is important that lobbies favouring open markets do not have other means of getting what they want. Finger (1991) has pointed out that large countries increasingly negotiate increased market access for their exporting firms bilaterally. Such bilateral alternatives weaken the power of reciprocity in the multilateral context, as they reduce the incentives for export interests to support liberalization during MTNs. If true, this would constitute a major systemic downside of regional integration. As discussed in Chapter 10, other analysts take an opposite view and argue that PTAs may create political economy forces that generate support for expanding preferential liberalization to nonmembers and thus eventual multilateralization (Ethier, 2004; Baldwin, 2006*a*).

A code of conduct for trade policy

The trade policy exchange market (MTNs) generates specific commitments by the participants. These commitments pertain to market access—specific liberalization

promises—and to certain rules of the game that all agree to abide by. The WTO encompasses a complex set of specific legal obligations regulating trade policies of member states. These are embodied in the GATT, the GATS and the TRIPS agreement. The rules and principles of the WTO constrain the freedom of governments to use specific trade policy instruments, and are largely motivated by a desire to constrain the ability of signatories to re-impose protection through the 'back door'.

As mentioned previously, one view of the role of the WTO is that is analogous to a mast to which governments can tie themselves to escape the siren-like calls of various pressure groups (Roessler, 1985). It is a mechanism through which the political market failure that is inherent in many societies—both industrialized and developing—can be corrected, at least in part, because reneging on liberalization commitments requires compensation of affected trading partners. However, much depends on the will of governments to tie themselves to the mast and on the strength of the rope used. WTO rules and disciplines—discussed at length in later chapters—embody many holes and loopholes that governments can invoke if they desire to. Much also depends on whether it makes economic sense to tie oneself to the mast. A necessary condition is that abiding by the rules is in the national interest of members. As discussed subsequently, a number of existing WTO rules arguably do not meet this test.

The WTO embodies a rule-oriented approach to multilateral cooperation. This contrasts with what can be characterized as a results-oriented or managed-trade approach—agreements on trade flows, market share or international prices. The WTO establishes a framework for trade. It does not define or specify outcomes. Four principles are of particular importance in understanding both the pre-1994 GATT and the WTO code of conduct: (1) nondiscrimination; (2) transparency (3) accountability; and (4) flexibility. Each of these is discussed at length in subsequent chapters; what follows briefly summarizes the main features of each.

Nondiscrimination: MFN and national treatment

The principle of nondiscrimination has two components, the MFN rule and the national treatment principle. Both components are embedded in the main WTO rules on goods, services and intellectual property. However, their precise scope and nature differ across these three areas, especially national treatment (see later chapters). The MFN rule requires that a product made in one member country be treated no less favourably than a 'like' (very similar) good that originates in any other country. Thus, if the best treatment granted a trading partner supplying a specific product is a 5 per cent tariff, then this rate must be applied immediately and unconditionally to the imports of this good originating in all WTO members.

Most favoured nation applies unconditionally. It cannot be made conditional on considerations of reciprocity, which is a principle that applies in negotiations, not in the application of negotiated rules. However, exceptions are made for the

formation of free trade areas or customs unions and preferential treatment of developing countries. Upon accession of a new member, an existing member may also invoke the WTO's nonapplication clause (Article XIII). These exceptions to MFN are discussed in subsequent chapters.

Most favoured nation is a fundamental rule for the WTO for a number of reasons. It ensures that deals that are struck between two countries to lower tariffs are not 'undone' subsequently by one of the parties offering better terms to another country. That is, MFN is an instrument that helps make reciprocity 'work' (Bagwell and Staiger, 2002, 2004). It provides insurance against so-called concession diversion (Schwartz and Sykes, 1997; Ethier, 2004). Most favoured nation also reduces overall negotiating costs—once a negotiation has been concluded with one country, the results extend to all. This obviates the need for other countries to initiate discussions to obtain similar treatment. Instead, negotiations can be limited to the principal suppliers of specific products. Most favoured nation also provides smaller countries with a guarantee that larger countries will not exploit their market power by raising tariffs against them in periods when times are bad and domestic industries are clamouring for protection, or alternatively, give specific countries preferential treatment for foreign policy reasons. Most favoured nation raises the costs of lobbying for protection by ensuring that all exporters to a market will be affected by an increase in protection. Most favoured nation therefore helps in the enforcement of multilateral rules by raising the costs to a country of defecting from the trade regime to which it committed itself in an earlier MTN or upon accession. If it desires to raise trade barriers it must apply the new policies to all WTO members. This increases the political cost of reneging on prior commitments because it implies higher economic costs for importers, who then have stronger incentives to object to the policy change. Finally, from a consumer welfare perspective, if policy does not discriminate between foreign suppliers, importers and consumers will continue to have an incentive to source from the lowest cost foreign supplier.

The national treatment rule is the second leg of the nondiscrimination principle. It requires that foreign goods—once they have satisfied whatever border measures apply—be treated no less favourably than like or directly competitive goods produced domestically in terms of internal (indirect) taxation (Article III: 2 GATT). That is, goods of foreign origin circulating in the country should be subject to the same taxes and charges that apply to identical goods of domestic origin. A similar obligation applies to nontax policies (regulations) (Article III: 4 GATT). In both cases, the obligation is to provide treatment 'no less favourable'. A government is free to discriminate in favour of foreign products (against domestic goods) if it desires, subject, of course, to the MFN rule—all foreign products must be given the same treatment.

National treatment is a virtually all-encompassing discipline. The potential reach of the national treatment provisions in WTO agreements is far-reaching: they span virtually *all* governmental policies that affect the conditions for sale and

distribution, widely interpreted, of imported products (Horn and Mavroidis, 2004). Moreover, the rule is not limited to explicitly discriminatory measures, but also spans any policy that *indirectly* has the effect of discriminating against imports. The rationale for national treatment is to preclude the use of domestic regulatory or tax policies to nullify a negotiated tariff concession. The reach of the principle is, therefore, limited to the impact of specific policies on (very) specific products, with much depending on whether domestic and imported products are 'like' each other.

The provision has, not surprisingly, given rise to a substantial number of disputes and case law, which is discussed in Chapter 5.

Although the nondiscrimination rules are invariably regarded as fundamental and defining principles for the trading system, the theoretical rationale for MFN remains a matter of debate and research by economists. Although it is clear that the policymakers who designed the GATT placed great weight and importance on the principle of nondiscrimination—strongly influenced by the inter-war experience—exactly how MFN helps to sustain cooperation and what its role is in moving countries to adopt lower tariffs than they otherwise would is less clear. Much of the literature on this question—which is surveyed in WTO (2007) and Horn and Mavroidis (2001)—has tended to focus on analysing situations where countries are symmetric. More recent analyses that allow for the types of asymmetry that characterize actual trade relationships may help in deepening the understanding of the role played by nondiscrimination (Box 1.5).

Transparency: information and communication

Ensuring that commitments are implemented requires access to information on the trade regimes that are maintained by members. Numerous mechanisms are incorporated into the agreements administered by the WTO to facilitate communication between members on the policy areas covered by agreements. A large number of specialized committees, working parties, working groups and councils meet regularly in Geneva. These interactions allow for the exchange of information and views, concerns and disagreements to be aired, and potential conflicts to be defused in an efficient manner.

World Trade Organization members are required to publish their trade regulations, to establish and maintain institutions allowing for the review of administrative decisions affecting trade, to respond to requests for information by other members, and to notify changes in trade policies to the WTO. These internal transparency requirements are supplemented by multilateral surveillance of trade policies by WTO members, facilitated by periodic country-specific reports (Trade Policy Reviews) that are prepared by the secretariat and discussed by the WTO Council—the so-called Trade Policy Review Mechanism (see Chapter 2). This external surveillance also fosters transparency, both for citizens of the countries

Box 1.5. Understanding the role of MFN

Economic first principles suggest that the optimal tariff policy of a country is likely to be discriminatory. One reason already alluded to for this is whether or not a country can affect the terms of trade. Another is whether or not the government imposes tariffs to collect revenue. In both cases, according to the so-called Ramsey pricing rule (Ramsey, 1927), the level of the tariff should be higher on producers (consumers) that have less elastic supply (demand). If the demand for a good is uniformly less elastic than that for another good, the optimal tax rate is higher for the first good due to the lower deadweight loss from taxing it rather than the second good. If the first good is totally inelastic there is no deadweight loss from taxing it, and the first best can be reached by taxing just this good. Broda, Limão and Weinstein (2006) provide evidence that countries that are not bound by the GATT/WTO systematically set higher tariffs on goods that are supplied inelastically, and that those that can affect the terms of trade do indeed levy higher tariffs, as predicted by the theory.

The various potential reasons motivating the use of MFN mentioned in the text taken together suggest that MFN is important in supporting the use of trade agreements by governments. Given the myriad differences across countries, it is quite unlikely that the nondiscrimination rule affects all countries in a similar fashion. Research has begun to emerge that puts country heterogeneity at centre stage. Saggi and Sengul (2008) argue that useful insights regarding the role of the GATT/WTO system in world trade can be achieved by formally analysing GATT as a club whose only requirement is that members grant MFN to each other. They show that the desirability of such an MFN club from a country's perspective depends on how its production cost compares to others. In their model, receiving MFN from others is of greater value to countries that have relatively lower costs of production. In related work, Saggi (2009) concludes that adoption of MFN by a country hurts the smaller exporter to its market while benefitting the larger one. Thus, the application of MFN by a country does not necessarily benefit all of its trading partners.

Saggi and Yildiz (2005) note that when market structure is asymmetric across countries, MFN does not necessarily dominate tariff discrimination even from a world welfare perspective An intriguing result of their analysis is that a high-cost country may choose to join an MFN club even though its welfare as a member is lower relative to a world in which no such club exists (i.e. a scenario where all countries pursue tariff discrimination). This result obtains because the fate of a high-cost country as a nonmember can be even worse than that as a member. This result may shed some light on the role played by special and differential treatment (SDT) in the multilateral trading system (discussed in Chapter 12). Saggi and Sengul (2008) suggest that such exceptions to MFN may be necessary to undo some of the adverse distributional effects of an MFN club on high-cost members. In their analysis the adoption of SDT helps ensure that the MFN club benefits all members.

concerned and for trading partners. It reduces the scope for countries to circumvent their obligations, thereby reducing uncertainty regarding the prevailing policy stance.

Transparency is a basic pillar of the WTO. It is a legal obligation, embedded in Article X GATT and Article III GATS. Transparency is important for several

reasons. It reduces the pressure on the dispute settlement system, as measures can be discussed in the appropriate WTO body. Frequently, such discussions can address perceptions by a member that a specific policy violates the WTO—many potential disputes are defused in informal meetings in Geneva. Transparency is also vital in terms of ensuring 'ownership' of the WTO institution—if citizens do not know what the organization does it will have less legitimacy and political support for it may erode. The Trade Policy Reviews are a unique source of information that can be used by civil society to assess what the implications are of the overall trade policies that are pursued by a government. From an economic perspective, transparency can also help reduce trade-policy-related uncertainty. Countries with policy regimes that are perceived by investors as unstable are generally associated with higher capital costs—investors will demand a risk premium on funds invested in such countries to take into account the probability of losses due to policy reversals. Such premia can be high. Mechanisms to improve transparency can help lower risk perceptions by reducing uncertainty. World Trade Organization membership itself, with the associated commitments on trade policies that are subject to binding dispute settlement, can also have this effect.

Accountability: enforceable commitments

Liberalization commitments and agreements to abide by rules of the game will have little value if they cannot be enforced. The nondiscrimination rules play an important role in ensuring that market access commitments are implemented and maintained. The tariff commitments made by WTO members in a MTN and upon accession are enumerated in schedules (lists) of concessions. These schedules establish so-called ceiling bindings—the member concerned cannot raise tariffs above bound levels without negotiating compensation with the principal suppliers of the products concerned. The MFN rule then ensures that such compensation—usually reductions in other tariffs—extends to all WTO members, raising the cost of reneging. Once tariff commitments are bound, it is important that other, nontariff, measures that can hollow out the value of the tariff concession are not used. A number of GATT provisions, including a ban on the use of quantitative restrictions on imports and exports and the rules on subsidies, essentially serve this purpose (see Chapter 5).

If a country perceives that actions taken by another government have the effect of nullifying or impairing negotiated market access commitments or the disciplines of the WTO, it may bring this to the attention of the government involved and ask that the policy be brought into conformity with its obligations. If satisfaction is not obtained, WTO dispute settlement procedures may be invoked. These involve the establishment of a panel of impartial experts who are charged with determining whether a contested measure violates a member's commitments under the WTO. Because the WTO is an inter-governmental agreement, private

parties do not have legal standing before the WTO's dispute settlement body. Only governments have the right to bring cases. The existence of dispute settlement procedures precludes the use of unilateral retaliation. For small countries in particular, recourse to a multilateral body is vital, as unilateral actions will be ineffective and thus not be credible. More generally, small countries have a great stake in a rule-based international system, as this constrains the likelihood of being confronted with bilateral pressure from large trading powers to change policies that are not to their liking.

Flexibility: calibrated commitments and (conditional) safety valves

A final principle characterizing the WTO is flexibility. This manifests itself in a number of forms. One is that governments may, if they desire, re-impose trade restrictions in specified circumstances. There are three types of provisions in this connection: articles allowing for the use of trade measures to attain noneconomic objectives, articles aimed at ensuring 'fair competition', and provisions allowing for intervention in trade for economic reasons. The first include provisions allowing for policies to protect public health or national security, and to protect industries that are seriously injured by competition from imports. The underlying idea in the latter case is generally that governments must be able to use trade policy instruments when competition from imports becomes so vigorous that domestic competing industries confront major adjustment pressures, with consequent political and social problems. The second type of measures include the right to impose countervailing duties on imports that have been subsidized and antidumping duties on imports that have been dumped—sold at a price that is below that charged in the home market. The objective of 'fair competition' is often in direct conflict with market access, as the instrument used by governments to attain 'fairness' is usually a trade barrier. Such barriers are, however, perfectly legal and permitted as long as they satisfy the criteria laid down in the relevant WTO provisions. Finally, the third type of 'safety valve' allows for actions to be taken if there are serious balance-of-payments difficulties, or if a government desires to support an infant industry.

1.4. From GATT to WTO

The General Agreement on Tariffs and Trade was not formally an international organization (that is, a legal entity in its own right), but an inter-governmental treaty. As a result, instead of member states, GATT had contracting parties.

This changed with the establishment of the WTO, which is an international organization that administers multilateral agreements pertaining to trade in goods (GATT, 1994*a*, as well as numerous issue-specific agreements on antidumping, subsidies, import licensing, and so forth), trade in services (GATS), and trade-related aspects of intellectual property rights (TRIPS). To reflect the fact that the WTO is an organization, in this book we will generally use the terms 'contracting parties' to refer to signatories of the pre-1994 GATT, and 'members' to refer to signatories of the WTO. We also make a distinction between the GATT 1947 (the old GATT) and the GATT 1994 that is embodied in the WTO. The old GATT was both a set of rules and an institution; the new GATT is just one of three multilateral agreements that are overseen by the WTO.

The WTO applies to agreements between nation-states and customs territories that address government policies. The WTO deals predominantly with the actions of governments, establishing disciplines on trade policy instruments such as tariffs, quotas, subsidies or state trading. Thus, the WTO is a regulator of regulatory actions taken by governments that affect trade and the conditions of competition facing imported products on domestic markets. In this it is no different from the old GATT.

A fundamental perception of the founders of the GATT was that multilateral institutions facilitating cooperation between countries were important not only for economic reasons, but also that the resulting increase in interdependence between countries would help to reduce the risk of war (Meade, 1940; Hull, 1948; Penrose, 1953; Hirschman, 1969). The expected increase in real incomes following trade liberalization and nondiscriminatory access to markets was expected to reduce the scope for political conflicts. The increase in transparency and the availability of a forum in which to discuss potential or actual trade conflicts was expected to reduce the probability of these spilling over into other domains. The Preamble of the GATT 1947 states that its objectives include raising standards of living, ensuring full employment and a large and steadily growing volume of real income and effective demand, developing the full use of the resources of the world and expanding the production and exchange of goods (GATT, 1994*a*: 486). It goes on to say that reciprocal and mutually advantageous arrangements involving a substantial reduction of tariffs and other barriers to trade, as well as the elimination of discriminatory treatment in international trade, will contribute to the realization of these objectives. Nowhere is any mention made of free trade as an ultimate goal. This continues to be the case under the WTO.

The GATT emerged from the negotiations to create an ITO after the Second World War. The negotiations on the charter of such an organization, although concluded successfully in Havana in 1948, did not lead to the establishment of the ITO because the US Congress was expected to refuse to ratify the agreement. The GATT was negotiated in 1947 between 23 countries—12 developed and 11

developing—before the ITO negotiations were concluded.[12] The countries involved in the 1947 exchange of tariff reductions were anxious that implementation of liberalization not be conditional upon the conclusion of the ITO talks. Therefore, they created the GATT as an interim agreement. As the ITO never came into being, the GATT was the only concrete result of the ITO negotiations.

Although the GATT incorporated the provisions of the commercial policy chapter of the ITO, having been conceived as a temporary trade agreement, it lacked an institutional structure. In the first years of its operation it did not even exist as an entity except once or twice a year when formal meetings of the contracting parties were held (Curzon and Curzon, 1973). Its organizational structure emerged only gradually. Although major decisions were taken at the sessions of the CONTRACTING PARTIES,[13] it rapidly became obvious that a standing body was needed. An inter-sessional committee was formed in 1951 to organize voting by airmail or telegraphic ballot on issues relating to import restrictions justified for balance-of-payments reasons. This committee was replaced in 1960 by a Council of Representatives, which was given broader powers and responsibilities for day-to-day management. Throughout the 1947–94 period, the GATT secretariat was formally known as the Interim Commission for the International Trade Organization (ICITO), created during the negotiations on the ITO. It was technically a United Nations (UN) body, as the ITO negotiations occurred under UN auspices. Because the ITO never came into existence, the formal relationship between the GATT (a treaty) and the UN was always tenuous.

Over the more than four decades of its existence, the GATT system expanded to include many more countries. It evolved into a de facto world trade organization, but one that was increasingly fragmented as 'side agreements' or codes were negotiated among subsets of countries. Its fairly complex and carefully crafted basic legal text was extended or modified by numerous supplementary provisions, special arrangements, interpretations, waivers, reports by dispute settlement panels and council decisions. As of the early 1990s, a well-oiled GATT machine existed, helping contracting parties manage developments in the trading system, including through surveillance of trade policies and assisting conflict resolution through consultations, negotiations, mediation and dispute settlement.

Some of the major milestones are summarized in Table 1.2. The early years of the GATT were dominated by accession negotiations, a review session in the mid-1950s

[12] The founding parties to the GATT were Australia, Belgium, Brazil, Burma, Canada, Ceylon, Chile, China, Cuba, Czechoslovakia, France, India, Lebanon, Luxembourg, Netherlands, New Zealand, Norway, Pakistan, Southern Rhodesia, Syria, South Africa, the United Kingdom and the United States. China, Lebanon and Syria subsequently withdrew.

[13] The term CONTRACTING PARTIES, in capital letters, was used to denote joint actions taken by all signatories to the agreement.

Table 1.2. From GATT to WTO: a chronology

Date	Event
1947	Tariff negotiations between 23 founding parties to the GATT concluded.
1948	GATT provisionally enters into force on 1 Jan. 1948. Delegations from 53 countries sign the Havana Charter establishing an ITO in March 1948.
1949	Annecy round of tariff negotiations.
1950	China withdraws from the GATT. The US Administration abandons efforts to seek Congressional ratification of the ITO.
1951	Torquay round of tariff negotiations. Germany (Federal Republic) accedes.
1955	A review session modifies numerous provisions of the GATT. A move to transform GATT into a formal international organization (an Organization for Trade Cooperation—OTC) fails. The US is granted a waiver from GATT disciplines for certain agricultural policies. Japan accedes to the GATT.
1956	Fourth round of multilateral negotiations held in Geneva.
1957	Creation of the European Economic Community.
1960	A council of representatives is created to manage day to day activities. The Dillon Round is launched.
1961	Dillon Round concluded. The 'Short-Term Arrangement' permitting quota restrictions on exports of cotton textiles agreed as an exception to GATT rules.
1962	The Short-Term becomes the Long-Term Arrangement on Cotton Textiles.
1964	The Kennedy Round begins.
1965	Part IV (on Trade and Development) is added to the GATT, establishing new guidelines for trade policies of—and towards—developing countries. A Committee on Trade and Development is created to monitor implementation.
1967	Kennedy Round concludes.
1973	The Tokyo Round starts.
1974	The Agreement Regarding International Trade in Textiles, better known as the Multifibre Arrangement (MFA) enters into force, replacing the Long-Term Agreement. The MFA restricts export growth to six per cent per year. It is negotiated in 1977 and 1982 and extended in 1986, 1991 and 1992.
1979	Tokyo Round concludes. Includes a set of 'codes of conduct' on a variety of trade policy areas that countries can decide to sign on a voluntary basis. Most codes predominantly attract OECD membership.
1982	A GATT ministerial meeting—the first in almost a decade—fails to agree on an agenda for a new round.
1986	After lengthy preparatory work, including national studies on trade in services, the Uruguay Round is launched in Punta del Este, Uruguay.
1990	A ministerial meeting in Brussels fails to conclude the Uruguay Round. Canada formally introduces a proposal to create a Multilateral Trade Organization that would cover the GATT, the GATS and other multilateral instruments agreed in the Uruguay Round.
1993	In June the US Congress grants fast-track authority to the US Administration—under which it cannot propose amendments to the outcome of negotiations—setting a 15 December deadline for the Uruguay Round to be concluded. Three years after the scheduled end of negotiations, the Uruguay Round is concluded on 15 December in Geneva as a 'Single Undertaking'.
1994	In Marrakech, on 15 April, ministers sign the Final Act establishing the WTO and embodying the results of the Uruguay Round.
1995	The WTO enters into force on 1 January. Financial services agreement concluded but US does not sign.

(cont.)

Table 1.2. (*Continued*)

Date	Event
1996	Maritime services talks collapse. The first WTO ministerial conference hosted by Singapore creates working groups on trade and investment, trade and competition policy, transparency in public procurement and trade facilitation. Integrated Framework for Trade-related Technical Assistance created.
1997	Forty governments agree to eliminate tariffs on computer and telecommunication products by the year 2000 (the Information Technology Agreement).
	Negotiations on an Agreement on Basic Telecommunications and a Financial Services Agreement are concluded under GATS auspices.
1998	The second WTO ministerial conference commemorating the fiftieth anniversary of the multilateral trading system takes place on 18–20 May in Geneva.
1999	Ministerial meeting in Seattle fails to launch a new round.
2000	Negotiations start on the so-called built-in agenda determined at the end of the Uruguay Round—agriculture and services.
2001	China accedes to the WTO. A new round is launched in Doha, Qatar, the Doha Development Agenda, spanning trade in agriculture, manufactures, and services. EU puts in place the 'Everything But Arms' initiative granting LDCs duty and quota free access to its markets.
2003	Establishment of the 'G20' group of developing countries. The 'mid-term' review Ministerial meeting in Cancun collapses amid disagreement on whether to launch negotiations on the four so-called Singapore issues, as well as differences on agriculture—including an African Heads of State call for accelerated reductions in cotton subsidies.
	General Council Decision allowing WTO members to grant compulsory licences to import pharmaceutical products if there is insufficient local manufacturing capacity.
2004	In July a negotiating framework is agreed that includes only one of the four Singapore issues—trade facilitation, paving the way for continued negotiations.
	EU expands to encompass 25 member states.
2005	The final stage of the Uruguay Round Agreement on Textiles and Clothing is implemented, abolishing remaining quantitative restrictions on imports imposed by WTO members.
	Ministerial meeting in Hong Kong makes little progress beyond agreement to abolish export subsidies, agreement on duty- and quota-free market access for LDCs.
	TRIPS Agreement amended to formalize 2003 decision on compulsory licensing—the first ever amendment to the WTO; Aid for Trade taskforce established.
2006	The inability of the major protagonists to make concessions leads the Director General to suspend the Doha negotiations in mid year.
2007	Vietnam becomes the hundred-and-fiftieth member of the WTO.
	Expiry of US Trade Promotion Authority in June reduces prospects of timely conclusion of Doha talks.
	Deadline for conclusion of Economic Partnership Agreements between EU and ACP countries expires.
2008	The EU concludes a series of EPAs with ACP countries.
	In July another mini-Ministerial effort to agree on negotiating modalities for the Doha Round fails.
	In December, the Director-General of the WTO decides not to call for a Ministerial meeting to push forward Doha negotiations, citing a lack of demonstration.
2009	WTO launches an initiative to monitor and report on the use by Members of trade-related policy responses to the global financial crisis and recession.
	Ministerial conference planned for November 2009—the first in four years.

that led to modifications to the treaty, and the creation of the European Economic Community (EEC) in 1957. In 1962, derogations from the GATT rules in the area of trade in cotton textiles were negotiated. This developed into successive Multifibre Arrangements (MFA-I through MFA-IV; see Chapter 6)—a complex system of managed trade that was inconsistent with the basic principles of the GATT, but that benefitted producers in OECD countries as well as many of the developing countries that were granted a minimum level of guaranteed access to rich country markets. Starting in the mid-1960s, recurring rounds of MTNs gradually expanded the scope of the GATT to a larger number of nontariff policies. Until the Uruguay Round, effectively no progress was made on liberalization of trade in agricultural products and textiles and clothing. The deal that finally subjected these sectors to multilateral disciplines included agreement on the creation of the GATS, TRIPS and the WTO itself.

There are many similarities between the old GATT and the WTO. The basic principles remain the same. The WTO continues to operate by consensus and continues to be member-driven. However, a number of major changes did occur. Most obviously, the coverage of the WTO is much greater. Moreover, in contrast to the old GATT, the WTO agreement is much more of a 'single undertaking'—most of its provisions apply to all members. Thus, the WTO has many more implications for developing countries than did the GATT, where participation was more à la carte as well as being limited to trade in goods. In the dispute settlement area, the process became more 'legalistic' with the creation of a standing Appellate Body. Finally, much greater transparency and surveillance functions were granted to the secretariat through the creation of the Trade Policy Review Mechanism.

1.5. Challenges for Global Cooperation on Trade

The GATT proved a very successful instrument through which industrialized countries gradually lowered and bound their tariffs. The idea that a rule-based approach is superior to an outcome- or results-based trading system steadily gained adherents during the GATT years. Whereas many governments in the 1960s and 1970s were engaged in efforts to manage trade—through central planning, barter, or commodity agreements—this approach proved unsuccessful. Commodity agreements were difficult to enforce and generally failed. Central planning and centralized trade proved to be an unsuccessful system of economic management and was abandoned following the dissolution of the Council of Mutual

Economic Assistance (CMEA) and the Union of Soviet Socialist Republics (USSR), and the opening of the Chinese economy to international trade and private sector participation.

Over time the agenda of MTNs grew to include various nontariff policies. In part this reflected the expansion in use of instruments that circumvented GATT disciplines—voluntary export restraint agreements being an important example (Nogues, Olechowski and Winters, 1986). In the 1990s, the focus of attention began to turn to domestic regulatory regimes. However, tariffs have not become irrelevant. In OECD countries, tariffs on agricultural products are a multiple of those applied to manufactures, and within manufacturing, there are tariff peaks exceeding 15 per cent on many labour-intensive products in which developing countries have a comparative advantage. Developing countries tend to have barriers against imports of manufactures that are much higher than those prevailing in OECD countries. They also have high rates of protection on imports of many agricultural goods. Barriers to trade in services are more difficult to measure, but the consensus view is that these tend to be higher than those prevailing for trade in goods.

Although a significant tariff negotiating agenda still exists, future MTNs will revolve increasingly around nontariff measures (NTMs) and domestic policies that are deemed to have an impact on trade. Table 1.3 reports a measure of the overall level of trade restrictiveness implied by policies. The Overall Trade Restrictiveness Index (OTRI) is defined and calculated as the uniform tariff equivalent of observed policies on a country's imports. That is, they represent the tariff that would be needed to generate the actual level of trade reported for a country in 2006. The OTRI captures all policies on which information is reported by the United Nations Conference on Trade and Development (UNCTAD) (*ad valorem* tariffs, specific duties and NTMs such as price control measures, quantitative restrictions, monopolistic measures and technical regulations and mandatory product standards). As many NTMs are not necessarily protectionist in intent or effect, the OTRI is not

Table 1.3. Overall trade restrictiveness index, 2006 (per cent)

	Total Trade	Agriculture	Manufacturing
High Income (tariffs only)	7.0	43.1	4.3
	2.1	12.4	1.4
Upper Middle Income (tariffs only)	13.0	29.3	11.8
	4.6	6.6	4.5
Lower Middle Income (tariffs only)	11.8	26.5	10.6
	6.5	11.5	6.0
Low Income (tariffs only)	17.7	26.6	16.7
	10.8	15.3	10.4

Source: World Bank and IMF (2008).

necessarily a good measure of the level of protection that a government seeks to provide domestic industry. However, it is a good measure of the level of trade restrictions that are implied by policy, whatever the intent. Table 1.3 also reports the OTRI using only tariff data (including the *ad valorem* equivalent of specific duties). The data reveal that there is still a significant tariff negotiating agenda confronting WTO members, especially in agriculture, but that NTMs account for a major share of the overall level of trade restrictiveness.

This has implications for international cooperation: the interface between trade policy and economic policy more generally defined has become increasingly blurred. Agreeing on the elimination or reduction of NTMs is more difficult than negotiating downward the levels of tariffs. One reason for this is that it is much less obvious that specific NTMs are detrimental to a country's welfare. For example, attitudes towards environmental quality or product safety differ across countries, and this may be reflected in differences in environmental or product standards or in targeted subsidy programmes. Economic theory suggests that under certain conditions intervention will be called for (see Annex 2). Negotiations on regulatory issues, therefore, may be zero-sum games (some countries may lose), in contrast to tariff reductions, which are positive-sum (all countries gain, even though certain groups in each country will lose unless they are compensated). Another problem, again in contrast to tariffs, is that it can be difficult to agree on what constitutes a NTM. Even if agreement is reached on what types of policies are trade-distorting, incrementally reducing their negative impact may not be feasible. For many NTMs, all that may be possible is to agree to apply basic principles of transparency, national treatment, and MFN, and to seek to adopt procedural rules. However, pressures for harmonization of policies have been mounting. Although the GATT traditionally shied away from attempts to agree on common policies, differences in nontrade policies—regarding the environment, labour standards or antitrust—are increasingly leading to claims that these result in unfair competition and should be countervailed. A major challenge for WTO members is to deal with these pressures.

Experience has amply demonstrated that calls for protection and incentives to renege on liberalization commitments will inevitably arise. The Uruguay Round negotiations were a response to the managed trade and new protectionism that had proliferated during the late 1970s and early 1980s. The extensive recourse made by OECD governments to trade-distorting NTMs (antidumping, export restraint agreements, subsidies) was in part driven by exogenous shocks. These included the collapse of the Bretton Woods system of fixed exchange rates, and successive price hikes for crude oil imposed by the OPEC cartel, which helped give rise to stagflation (a mix of rising prices, weak output growth and rising unemployment). Matters were compounded by international political developments such as *détente* that reduced the primacy of foreign policy considerations in maintaining cooperation in trade.

As in the inter-war period, trade restrictions formed part of an inappropriate policy response to structural adjustment pressures, which were augmented by the emergence of East Asian countries as competitive suppliers of labour-intensive manufactures. The difference with the inter-war period was that multilateral cooperation did not break down. Although GATT rules were frequently ignored and circumvented, more often than not the letter, if not the spirit, of the rules of the game was honoured. The explosion of grey area measures, especially voluntary export restraints (VERs), constituted a major challenge to the system, but as discussed at greater length in subsequent chapters, VERs emerged in large part because of GATT disciplines on the use of emergency protection. The launch and successful completion of the Uruguay Round revealed that the major trading nations were willing to maintain multilateral cooperation and strengthen disciplines regarding the use of NTMs. The system proved robust during the 1997–8 financial crises—there was no significant increase in protectionism in East Asia or the OECD. Greater use of protectionist policies was observed in the 2008–9 global recession, but most countries did not significantly raise trade barriers. Those that invoked trade policy tended to use contingent protection mechanisms permitted by the WTO (antidumping, safeguards).

The World Trade Organization members confront a very different world from that existing in the immediate post-Second World War period. Although the US continues to be the dominant economy of the world, it is no longer a public-spirited hegemon willing to tolerate free riding and deviations from multilateral rules by trading partners for foreign policy reasons. Many of the trade disputes and the recourse to NTMs that emerged in the 1980s were in part a reflection of what Bhagwati (1991) has called the diminished giant syndrome of the US. Since then, the relative decline of the US in economic terms has continued, with the expansion of the EU to encompass 27 countries as of 2008, and the very rapid growth of China. The world economy is ever more multipolar. Instead of one dominant economic and political power (the US), there are now at least three major players—the EU, the US and China. None of the three can be relied upon to take up the type of leadership role provided by the US at the end of the Second World War. At the same time, the WTO as an international organization cannot take the lead—it is a membership-driven (controlled) institution, with a secretariat that has no power to self-initiate action or to make decisions. At the end of the day what matters is the continued willingness of WTO members to abide by the negotiated rules of the game, and to use the multilateral institution as a mechanism to liberalize trade further and pursue cooperation in areas that give rise to disputes and friction. This requires there to be clear-cut gains for all members—something that is becoming more difficult to achieve as talks confront thorny issues of domestic regulation. However, much still needs to be done on the 'traditional' agenda—the potential gains from further liberalization of trade in goods and services are still very large, for both OECD countries and for developing economies.

1.6. Further Reading

Ronald Findlay and Kevin O'Rourke, *Power and Plenty: Trade, War and the World Economy in the Second Millennium* (Princeton: Princeton University Press, 2007) is a fascinating, highly informative account of global trade and its determinants for most of recorded history. *Against the Tide: An Intellectual History of Free Trade* (Princeton: Princeton University Press, 1996), by Douglas Irwin, is a masterful tour de force that is required reading for anyone with an interest in the case that has been made for and against free trade. David Mansfield, *Power, Trade and War* (Princeton: Princeton University Press, 1994) is a careful empirical analysis of the relationship between an open international system, bilateral trade flows and the probability of war.

For an appraisal and history of negotiations of the Havana Charter and the General Agreement on Tariffs and Trade, see William Brown, *The United States and the Restoration of World Trade* (Washington, DC: The Brookings Institution, 1950); and William Diebold, *The End of the ITO* (Princeton: Princeton University Press, 1952). Richard Gardner, *Sterling-Dollar Diplomacy: The Origins and the Prospects of Our International Economic Order* (New York: McGraw-Hill, 1969, 2nd edn) is an excellent discussion of the motivations and processes underlying the construction of the post-war international economic institutions, including the GATT. An early study of the GATT system that continues to be well worth reading is Gerard Curzon's *Multilateral Trade Diplomacy* (London: Michael Joseph, 1965).

There is a large literature on the political economy of trade policy decisions and institutional design issues. I. M. Destler, *American Trade Politics* (Washington, DC: Institute for International Economics, 2005), now in its fourth edition is a classic and regularly updated book on the politics of US trade policy. Arye Hillman, *The Political Economy of Protectionism* (New York: Harwood, 1989) surveys the economic literature.

Robert Keohane, 'Reciprocity in International Relations', *International Organization*, 40 (1986): 1–27, discusses the notion of reciprocity from a political science and international relations perspective. L. Alan Winters, 'Reciprocity', in M. Finger and A. Olechowski (eds), *The Uruguay Round: A Handbook* (Washington, DC: The World Bank, 1987) does so from the perspective of an economist. S. H. Bailey, 'The Political Aspect of Discrimination in International Economic Relations', *Economica*, 12 (1932): 96–115, is an often-cited contemporary assessment of the costs of discrimination in trade.

Those interested in the theoretical framework underpinning the terms-of-trade view of the rationale for the WTO can do no better than consult Kyle Bagwell and Robert Staiger, *The Economics of the World Trading System* (Boston: MIT Press, 2002). For a theoretical analysis of the WTO that combines the terms-of-trade rationale with a political commitment motivation on the part of governments, see

Andres Rodriguez-Clare, 'A Political Economy Theory of Trade Agreements', *American Economic Review*, 97 (4) (2007): 1374–406. The various theories that have been developed to explain the role of the WTO by economists, political scientists and legal scholars are surveyed and summarized in WTO, *World Trade Report 2007* (Geneva: WTO, 2007).

A clear and accessible introduction to the legal and institutional aspects of the world trading system is presented in John H. Jackson, *The World Trading System: Law and Policy in International Relations* (Cambridge: MIT Press, 1997). John Croome, *Reshaping the World Trading System* (Leiden: Kluwer, 1999) is a detailed negotiating history of the Uruguay Round, written by a GATT insider. The prevalence of NTBs at the beginning of the 1980s is documented and quantified in Julio Nogues, Andrej Olechowski and L. Alan Winters, 'The Extent of Nontariff Barriers to Industrial Countries Exports', *World Bank Economic Review*, 1 (1986): 181–99. Patrick Low, *Trading Free: The GATT and US Trade Policy* (New York: Twentieth Century Fund, 1993) discusses the evolution of US trade policy thinking in the 1980s, the use of contingent protection and US attitudes towards the GATT.

Gilbert Winham, 'GATT and the International Trade Regime', *International Journal*, 15 (1990): 786–822, is a leading political scientist's view of the GATT and its role in international relations. Frieder Roessler, 'The Scope, Limits and Function of the GATT Legal System', *The World Economy*, 8 (1985): 287–98, discusses the role of GATT rules as constraints on governments. Alan Deardorff, 'An Economist's Overview of the World Trade Organization', in G. Flake and F. Myeong-Hwa Lowe-Lee (eds), *The Emerging WTO System and Perspectives From East Asia* (Washington, DC: Korea Economic Institute of America, 1996) provides an insightful and accessible economist's view of the WTO, emphasizing the importance of the institution as a forum for communication and information exchange.

A special issue of *The World Economy* (volume 23, April 2000) 'Developing Countries and the Next Round of WTO Negotiations' reviews many of the issues that were to figure on the agenda of the Doha Development Agenda. Bernard Hoekman, Aaditya Mattoo and Philip English (eds), *Development, Trade and the WTO: A Handbook* (Washington, DC: The World Bank, 2002) is a compilation of short papers that describe the WTO and the subjects on the Doha negotiating agenda. Donna Lee and Rorden Wilkinson (eds), *The WTO after Hong Kong: Progress In, and Prospects For, the Doha Development Agenda* (London: Routledge, 2007) provides an international relations perspective on the Doha negotiations, including discussions of coalitions and negotiating strategies. The UN Millennium Project Taskforce on trade report, 'Trade for Development' (New York: United Nations Development Programme, 2005) is a comprehensive analysis of the issues on the Doha agenda and their importance from an economic development perspective.

CHAPTER 2

THE WORLD TRADE ORGANIZATION

THE World Trade Organization was established on 1 January 1995. The WTO builds on the organizational structure of the GATT and its secretariat—to a significant extent it formalizes and extends the structure that had gradually evolved over a period of some 50 years. The Punta del Este Ministerial Declaration launching the Uruguay Round did not call for the creation of a WTO. In principle, it was not necessary to create an international organization to implement the outcome of the negotiations. The Canadian suggestion to establish a Multilateral Trade Organization in 1990—subsequently supported by the EU—was therefore something of a surprise.[1] The proposal was motivated by a desire to create a single institutional framework for world trade (Croome, 1999). This would encompass the modified GATT, the new agreements on services (GATS) and intellectual property (TRIPS), as well as all other agreements and arrangements concluded under the auspices of the Uruguay Round. The US initially opposed the idea, but after negotiations on the substance of the new organization, agreed to the framework that currently exists, including the name change.[2]

At Punta del Este it had been agreed that the negotiations were to be a 'single undertaking'. With the proposal to create the WTO, the concept of a single undertaking was redefined to mean that all GATT contracting parties had to become a WTO member. There was no alternative—remaining a member of GATT 1947 would have

[1] For convenience, in this book we use the acronym EU to denote both the European Union and the European Communities. The latter is formally the correct appellation in WTO contexts.

[2] The choice of name was somewhat ironic given the attention that was being given to intellectual property rights, as the acronym WTO was already in use by the World Tourism Organization, a Madrid-based special agency of the UN.