
THE POLITICAL
ECONOMY OF THE
WORLD TRADING
SYSTEM

THE WTO AND
BEYOND

Third Edition

BERNARD M. HOEKMAN
and
MICHEL M. KOSTECKI

OXFORD
UNIVERSITY PRESS

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.

1.2. 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. Tumlin, 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, 1987a). 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, 2006a).

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, 1994a, 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, 1994a: 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

CHAPTER 10

PREFERENTIAL TRADE AGREEMENTS AND REGIONAL INTEGRATION

BOTH the GATT and the GATS make explicit allowance for preferential trade agreements among a subset of members. Such agreements can be of two types: reciprocal and nonreciprocal. This chapter deals with the former; the latter are discussed in Chapter 12 as they arise in trade relations between industrialized and developing countries. Both types of preferential trade are inconsistent with the MFN principle and are therefore subject to multilateral disciplines that define minimum conditions that must be met for an agreement. The WTO also provides for multilateral scrutiny of such agreements. This chapter discusses the rationales for preferential trade agreements (PTAs) between WTO members, the WTO rules and their application in practice, and the economic literature exploring the relationship between PTAs and multilateralism (the trading system)—both theoretical and empirical. Given the steadily expanding number of PTAs, a critical question for the WTO is whether the network of PTAs create incentives to lower trade barriers on a MFN basis and thus help achieve a major objective of the drafters of the GATT.

Reciprocal trade agreements among subsets of the WTO membership have become a prominent part of the trade landscape. As of late 2007, 380 PTAs had been notified to the GATT/WTO. Of that number, 300 agreements were notified under Article XXIV of the GATT, 22 agreements involving developing countries

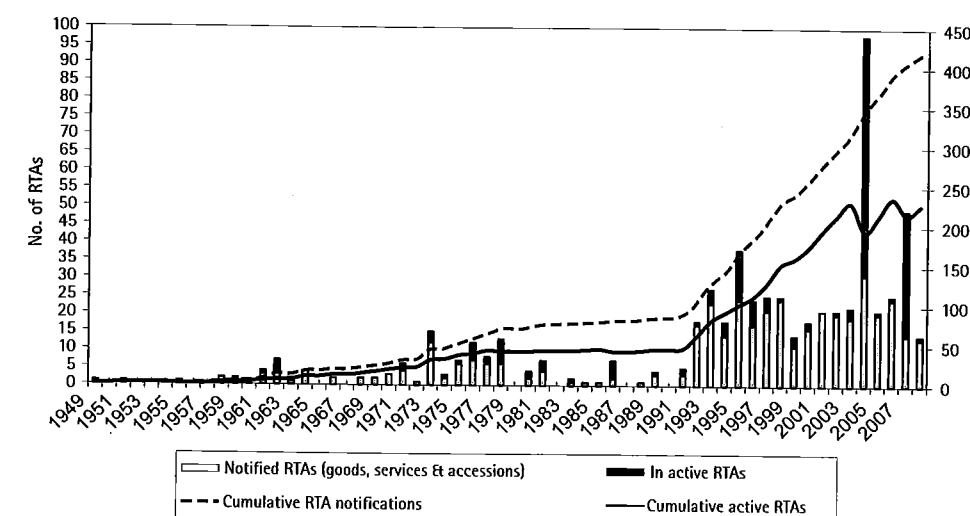


Fig. 10.1. Number of preferential trade agreements (1948–2007)

Source: WTO.

were notified under the Enabling Clause, and 58 under Article V of the GATS. Although these numbers are suggestive of a proliferation of PTAs, it is important to recognize that a large number of the PTAs notified to the GATT and WTO have involved prospective members of the EU and became irrelevant once the countries acceded to the EU.¹ About 200 PTAs were in force at the end of 2007 (Figure 10.1). Of these PTAs, customs unions account for less than 10 per cent. Many of the PTAs involve contiguous countries but many do not. In this chapter we reserve the term ‘regional integration’ for PTAs limited to neighbouring countries.

Since the late 1950s, the EC has been the market leader in the PTA business. European countries account for more than half of all PTAs notified to the WTO and that were still in force in 2008. The major regional grouping in Europe is the European Union, with 27 members in 2008. Other European PTAs include the European Free Trade Association (EFTA, Iceland, Lichtenstein, Norway and Switzerland).² and the Central European Free Trade Agreement. The EU has concluded

¹ As noted by Pomfret (2007), after the expansion of the EU by ten new members in 2004, some 65 PTAs between these countries and the EU became redundant. Note also that the numbers overstate the prevalence of PTAs because separate notifications are required under the GATT and the GATS for agreements that cover both goods and services—as many PTAs now do.

² At varying points in time EFTA also included Austria, Denmark, Finland, Portugal, Sweden and the UK, all of which left to become members of the EU. The European Free Trade Association has a very close economic relationship with the EU, governed by the European Economic Area Agreement.

Experience to date suggests that multilateral surveillance will only have a limited impact on the design and content of PTAs. An implication is that the payoff to efforts to strengthen specific WTO disciplines on PTAs is likely to be low, even if agreement could be attained. More fundamentally, it has been argued that efforts to devise a realistic rule that will ensure the trade policy stance of a PTA will be welfare-improving for members *and* the rest of the world are doomed to failure (Winters, 1999). There is simply no way to square the circle. As has been noted by many observers, the proliferation of PTAs extant clearly illustrates that WTO members regard PTAs as being in their interest. Thus, what matters is the economics—the incentives that are created by the proliferation of PTAs as regards the average level of MFN protection. That said, the rules are not irrelevant. In particular, they have had a major impact on the trade relations of a significant number of developing countries with the EU. As discussed in Chapter 12, an important reason why the EU concluded EPAs with ACP countries was to replace a system of unilateral preferences that violated GATT rules (and thus required a waiver) with a set of reciprocal trade agreements that satisfied Article XXIV. Interestingly, given the long history of nonenforcement of WTO rules in this area, it was the EU itself that was the ‘enforcer’ of Article XXIV, in that the desired trade coverage ratio of the EPAs was determined by the EU’s view of what is the minimum required by Article XXIV.

10.3. TRADING BLOCS AND THE TRADING SYSTEM

A key factor determining the importance of the effective absence of multilateral disciplines is the extent to which PTAs have detrimental effects on nonmembers and how they react. For the trading system what matters are the dynamic forces that are created by PTAs—do they create incentives that lead to a reduction in the external barriers of PTAs and nonmembers? As is the case for trade policy more generally, the most powerful pressures for reform are almost invariably domestic, not external, although external forces can help support domestic constituencies that favour a more liberal trade regime. That suggests a focus on the economic impacts of PTAs on interest groups.

What follows briefly considers three relevant questions in this regard. First, what is the impact of PTAs on trade and welfare of members? Second, what are the consequences for nonmembers? Third, what are the incentives created by PTAs once they have been formed for both members and nonmembers in terms of their trade policy strategies?

Impacts of PTAs on members

The effects of a PTA are determined by their coverage and design and on whether and how they are implemented. If a PTA is not implemented it cannot have an effect. Many of the PTAs negotiated since the 1960s were only partially implemented, if at all, or else excluded so many industries and tariff lines that their trade effects could only be minimal. The more recent vintage PTAs are generally implemented, and as mentioned tend to have more substantial coverage of merchandise trade flows.

Empirical research assessing the magnitude of the trade impacts shows, not surprisingly, that PTAs increase trade between members. That is, after the PTA is implemented, one observes greater trade flows between members. The difficulty for researchers, however, is to establish whether there is a causal effect. In a world where countries and thus trade is growing one would expect more intra-PTA trade without a PTA. Matters are compounded by countries also undertaking unilateral liberalization at the same time or before they engage in PTAs—what then is driving increased intra-PTA trade?

This suggests that empirical evaluations of PTA impacts must compare outcomes to what would have happened absent the PTA (the counterfactual). This is very difficult if not impossible as the PTA exists after all. What can be done, however, is to control for other factors and variables that affect trade flows. The basic workhorse tool that tends to be used to assess the effects of PTAs is the gravity model. This is a model that has been shown to be very effective at explaining trade volumes between country pairs, and that is consistent with what economic theory predicts are the determinants of trade. In a nutshell it postulates that trade between two countries is a function of their size, their wealth, their distance from each other, whether they are contiguous and speak the same language, and policy variables. The latter include the existence of a PTA.

Much of the literature on this subject is summarized by Schiff and Winters (2003). Surprisingly, there is no agreement on whether PTAs lead to more intra-PTA trade—indeed some studies find a negative effect. More recent research using the gravity model by Baier and Bergstrand (2007) argues that the findings of much of this literature greatly understate the trade effects of PTAs because they ignore the political economy of trade policy, i.e. why the PTA was negotiated in the first place. Technically, what researchers often assume is that the formation of a PTA is exogenous. In practice it is not likely to be—instead the level of trade can be expected to determine whether or not to join a PTA. If account is taken of this endogeneity, the impact of PTAs on trade volumes with partner countries rises significantly. On average PTAs do have a significant effect on intra-PTA trade: according to Baier and Bergstrand (2007) on average a FTA doubles trade between two members after ten years.

Of course, more trade is not necessarily good from a welfare perspective. What matters is how much of what is observed is trade creation and how much is diversion. Not surprisingly, empirical assessments of the impact of PTAs do not come to uniform conclusions. Much depends on the structure of trade before and after the formation of the PTA, on the pattern of comparative advantage, the size and composition of the PTA, etc. That said, many PTAs have been found to generate trade diversion. For example, in a study of eight major PTAs over the 1970–92 period, Frankel, Wei and Stein (1997) found that increases in intra-bloc trade were accompanied by reductions in trade with nonmembers, i.e. generated trade diversion. Similarly, Soloaga and Winters (2001), focusing on nine major PTAs over the period 1980–96, found trade diversion in European PTAs (explained by the fact that little external liberalization occurred during this period relative to intra-PTA liberalization), whereas PTAs among developing countries saw trade increase with both members and nonmembers. The explanation for this is that trade policies were reformed by the developing countries more generally, i.e. reforms were not limited to preferential liberalization. Controlling for developments in general (MFN) trade policies, the non-European PTAs had no independent effect on intra-PTA trade flows (Schiff and Winters, 2003: 42–3).

Very detailed analysis at the HS six-digit level of disaggregation (some 5,000 products) of the impact of the FTA between Canada and the US and the subsequent NAFTA by Romalis (2005) provides clear evidence of trade diversion. He shows that the greatest increases in US imports from Mexico occurred in items on which the US imposes the highest MFN tariffs, i.e. those goods where NAFTA provides Mexico with the highest preferential tariff margins. A similar result obtains for Canada. Although overall welfare effects of NAFTA for the US are small, one reason for this is the trade diversion, which results in higher prices of protected goods. Romalis (2005) also finds that volume effects are significant: NAFTA increases trade between Mexico and both Canada and the US by almost 25 per cent. Thus, studies suggest that there may well be significant market access and terms of trade benefits for countries joining a PTA, as well as distributional effects—with consumers paying the costs of any trade diversion.

Clearly, a narrow focus on merchandise trade is inadequate to assess the effects of PTAs. As, if not more, important, are the impacts on investment and FDI, and the associated potential for the acquisition and diffusion of technology, and the extent and implications of the 'deeper integration' dimensions of PTAs. Many studies have found that 'serious' PTAs may encourage FDI inflows and that these in turn can generate positive productivity spillovers (Schiff and Winters, 2003). There is nothing automatic about such investment and spillover effects, however. The experience of some 20 developing countries between 1980 and 2000 illustrates that many PTAs have not led to significant new FDI inflows (World Bank, 2005).

Impacts of PTAs on nonmembers

From the perspective of the trading system, the impacts of PTAs on nonmembers is the relevant question. Economists sometimes argue that a necessary condition for PTAs not to be detrimental to nonmembers is that the volume of imports by member countries from the rest of the world not decline on a product-by-product basis after the implementation of the agreement (McMillan, 1993). The empirical literature suggests that the trade volume test has been met in the past. Although the intensity of intra-regional trade increased in the second half of the twentieth century, the propensity of regions to trade with the rest of the world, expressed as a percentage of their GDP, has also expanded (Anderson and Norheim, 1993). Global integration—as measured by trade flows and capital flows—does not appear to have been affected negatively by PTAs.

As pointed out by Winters (1997), the 'trade volume test' is a flawed one in that it does not guarantee that nonmembers are not hurt by a PTA. For the welfare of nonmembers what matters is the impact of a PTA on trade flows and the associated change in prices. Even if the Article XXIV conditions are met, and even if the net aggregate imports of PTA members do not contract, imports of particular products by the region may decline *ex post*, or prices received by exporters may fall, harming producers in the rest of the world.

The converse of the trade diversion discussed above is that it implies a decline in exports for nonmember countries to a PTA, and perhaps an overall decline in aggregate exports if the diverted trade cannot be redirected to other markets and sold at the same price. Schiff and Winters (2003) discuss much of the literature, which finds that nonmembers have at times experienced significant reductions in exports. In the case of NAFTA, Romalis (2005) concludes that every 1 per cent reduction in intra-NAFTA tariffs causes a decline in exports to NAFTA from the rest of the world ranging from 1.3 to 3.9 per cent. Although such findings are suggestive, a more appropriate measure of the welfare impact of a PTA on nonmembers is to focus on what happens to their export prices in PTA markets after the agreement is formed. Chang and Winters (2002) show that Brazil's membership of Mercosur was accompanied by a improvement in Brazil's external terms of trade. Exporters based in the US, EU, Japan and Korea all saw the relative prices of many of their goods on the Brazilian market fall. There is also some evidence of negative investment effects (Baldwin, Forslid and Haaland, 1996).

Limão (2006) and Karacaovali and Limão (2008) have shown that in the case of both the EU and the US, PTAs may be a force working against nonmembers: they find that both the EU and the US made fewer (shallower) multilateral (MFN) liberalization commitments in the Uruguay Round on tariff lines where there were significant preference margins for imports from their preferential trading partners. Limão (2007) hypothesizes that this may reflect the use of market access as

'payment' for concessions by PTA partners in nontrade areas. Whatever the rationale, this is evidence that PTAs can have stumbling block effects.

Thus, notwithstanding the fact that there is little evidence of large-scale negative effects of the spread of PTAs—as reflected in the steady increase in world trade and openness discussed in Chapter 1—the economic literature suggests that there is no justification for complacency regarding the effects of PTAs. Preferential trade agreements impose costs on nonmembers even if they do not raise external levels of protection. Nonmember suppliers become less competitive because they continue to pay tariffs, whereas competing producers from member countries do not. Where there are economies of scale, PTAs may help lower member country firms' costs by expanding their home market. Conversely, they may restrain the ability of firms in nonmembers from realizing economies of scale.

There are various ways through which PTAs may constrain national interest groups and thus foster a more liberal external trade policy (De Melo, Panagariya and Rodrik, 1993). A first can be called the preference-dilution effect: because the region implies a larger political community, each of the politically important interest groups in member countries will have less influence on the design of common policies. The second is a preference-asymmetry effect: because preferences on specific issues are likely to differ across member countries, the resulting need for compromises may increase the probability of more efficient outcomes. The creation of PTAs may also disrupt the formation of rent-seeking interest groups, as these have to reorganize at the regional level, establishing an institutional structure that allows them to agree on a common position. But, PTAs may also facilitate the adoption of less liberal policies. Consumer interests may be harder to defend in a PTA than at the national level, whereas producer interests are more likely to be strengthened than weakened (Tumlir, 1983). Each national producer group may face less opposition when seeking price-increasing policies, and may indeed find support from other producer groups in other countries that pursue their own interests. The need for striking compromises may then result in a less liberal regulatory regime. Moreover, it may be in the interest of national politicians to let a regional organization satisfy national pressure groups as this is less transparent for domestic voters and can be justified as being necessary to maintain the agreement.

Much will generally depend on the type of PTA that is involved, FTA, customs union or hub-and-spoke system. The first two types differ from the last in that they imply nondiscrimination between the members of the agreement: any benefit granted to member country B by member country A is also available to member country C. Under a hub-and-spoke system this is not necessarily the case: each country negotiates a separate agreement with the hub country, and perhaps with other spoke partner countries as well. A major difference between a FTA and a customs union or common market is that the latter implies a common external trade policy. Whatever the extent of internal liberalization of trade and competition, implementation of a common external trade policy can give rise to an upward bias in the level of external protection over time if import-competing industries pursue instruments of contingent protection

such as antidumping actions. Thus, there may be no net increase in external trade barriers at the formation of a customs union, but there can easily be an upward trend if contingent protection is maintained. In contrast, FTAs have a different dynamic, as members in some sense compete in their external trade policies. Although the political economy of FTAs versus customs unions is complex, on balance, FTAs are likely to be more liberal than customs unions (Box 10.3).

Box 10.3. Pressures for protection: FTAs and customs unions

Under a customs union or common market the potential returns to protection-seeking will be higher than under a FTA: the expected payoff for a unit of lobbying effort increases because the size of the protected market is bigger. Moreover, liberal-minded governments that join a customs union may find it impossible to prevent domestic industries from seeking protection or to block the imposition of protection. For example, it may be the case that certain countries did not use (or make available) contingent protection before joining a customs union. However, once a member country, any domestic firm has access to the central trade policy authority and will be able to petition for AD. Indeed, the welfare gains to liberal countries from joining a customs union that employs contingent protection are reduced, as consumers are faced with higher expected levels of protection without knowing which industries will be affected (Hoekman and Leidy, 1993).

More generally, once a common external trade policy applies, decision-making structures may be biased toward more rather than less protection because of the so-called restaurant bill problem. If a group goes to a restaurant and shares the cost of the bill, each has an incentive to order more expensive dishes than they would if they ate on their own, as to some extent the others are expected to pick up part of the cost. The same is true in the EU (Winters, 1994b). The costs of protection are borne by all EU consumers, and are roughly proportional to each country's GDP. Benefits accruing to producers are proportional to the share of each country in total EU production of the good concerned. This establishes an incentive for each government to pursue protection for those products where their share of total EU production exceeds their country's share of EU GDP. Thus, the Netherlands may not like the EU-wide protection for cars sought by France and Italy, but may accept it if other policies are adopted for products in which it is relatively specialized (such as agriculture). Indeed, if larger countries are able to get the Commission to propose protectionist policies in specific areas, all EU member states have an incentive to ensure that some of their producers also obtain protection.

The external trade policy bias towards protection that may arise under a customs union will be weaker in a FTA. Because there is no common external trade policy, member countries compete in their external trade policies. Industries cannot lobby for area-wide protection. Although import-competing firms in member countries may have an incentive to obtain such protection, each industry will have to approach its own government. The required coordination and cooperation may be more difficult to sustain than in a customs union where the centralization of trade policy requires firms to present a common front. In any particular instance, some member country governments will award protection, whereas others will not. If industries in member countries are all

(cont.)

Box 10.3. (Continued)

competing against third suppliers, protection by one member may benefit industries in other member states. Such free riding can result in less protection than in the absence of the FTA (Deardorff, 1994). This benefit may be offset by other aspects of FTAs, in particular the need for rules of origin, which may allow industries to limit the extent of intra-area liberalization and can be detrimental to nonmembers (see Box 10.2 above).

Some evidence is beginning to emerge that supports these theoretical considerations on the likely dynamics of FTAs versus customs unions. Rigorous empirical research on the relationship between preferential and MFN tariffs over time is sparse as a result of data constraints—information on the implementation of PTAs and the sequencing over time between unilateral and preferential tariff reduction is not available for many PTAs. In the case of Latin America, however, a study by Estevadeordal, Freund and Ornelas (2008) concludes that the preferential tariff reduction following PTA formation in Latin America promotes subsequent external tariff reduction for those PTAs that are not customs unions. Bohara, Gawande and Sanguinetti (2004), focusing on the impact of preferential trade flows from Brazil to Argentina, find that greater imports from Brazil led to lower MFN tariffs in Argentina, especially in sectors where trade diversion occurred as a result of Mercosur. As the potential for trade diversion is especially great for South-South PTAs—because developing countries tend to have relatively high external trade barriers—the associated costs provide a powerful force for multilateralization: lowering external barriers to trade will reduce such costs.

Responses by nonmembers to PTA proliferation

As stressed by Bhagwati (1991), from the perspective of the WTO a key question is whether PTAs are a stepping stone or a stumbling block for multilateral liberalization. There is no consensus on the answer. Indeed, given that PTAs differ so much, there is no reason to expect a single, simple answer, especially as this is inherently a dynamic question—the answer depends critically on how PTAs affect the incentives of pro- and antitrade forces in both PTAs and excluded countries.

The most obvious reaction of third countries to the formation of a PTA is to seek a reduction in the bloc's external trade barriers through a MTN. As noted, this arguably has been a key role of the WTO in practice, with regional integration in Europe becoming a recurrent reason for MTNs under GATT auspices. Much of the Dillon Round (1960–1; see Chapter 4), was devoted to renegotiating a balance of concessions subsequent to the implementation of the EEC's common external tariff.

The same type of objectives played a role in the Kennedy and Tokyo Rounds. At the time of the Kennedy Round, the margins of preference for EEC members had

increased substantially, as most of the internal elimination of tariffs had been achieved. 'The record leaves no doubt that a compelling factor in the decision of Congress to pass legislation authorizing a 50 per cent linear cut in tariffs [in the Kennedy Round, see Chapter 4] ... was the belief that the Common Market posed a potentially serious threat to the growth, and perhaps even maintenance of American exports' (Patterson, 1966: 176). Thus, 'the task of the Kennedy Round ... was to attempt to mitigate [the] disruptive trade effects of European economic integration' (Preeg, 1970: 29). Some success was achieved, as the Kennedy Round reportedly prevented one-third to one-half of the trade diversion that might have occurred from European integration (*ibid.*: 220).

The first enlargement of the EEC in 1973—to include Denmark, Ireland and the UK—was a factor behind the launching of the Tokyo Round. The CAP also played a role. A major objective of the US was to improve its market access for agricultural products and to curb the EU's use of export subsidies. Links between regional integration and the Uruguay Round included the adoption of the Single European Act (the 1992 programme), the implementation of the Canada-US FTA, the negotiations on the NAFTA, and the continuing distortions of world agricultural trade induced by the CAP. The foregoing is not to say that PTAs are good because they give countries an incentive to pursue concurrent MTN-based liberalization. Without the EEC, much more progress might have been made towards multilateral liberalization (Winters, 1994b).

Another policy option is to seek to join existing PTAs. The primary example here is again the EU, which expanded from six to 27 member states between 1957 and 2007. In North America, Mexico was induced to seek accession to a Canada-US FTA, with the result being a re-negotiated trilateral FTA, the NAFTA. Other nations have negotiated FTAs with each of the NAFTA members in turn. One motivation for this is market access 'insurance'. The goal is not necessarily so much to obtain duty-free access to the regional market, as average MFN tariffs are relatively low for most products, and many potential members tend to be treated preferentially in any event. More important is the elimination of uncertainty, including the threat of contingent protection. This may be complemented by a desire to enhance the credibility of recently undertaken unilateral liberalization and structural reform efforts. However, particularly important are likely to be the firms in nonmember states that see their competitors get access to an ever larger internal market, allowing them to realize economies of scale and benefit from a reduction in real trade costs. This may well give rise to the 'domino effects' that have been observed in the case of the EU and NAFTA (Baldwin, 1995). Examples of such domino effects abound, especially in the European context. As mentioned at the beginning of this chapter, the EU has numerous PTAs with third countries.

The creation of a PTA may also generate incentives for third countries to pursue PTAs in turn. This may be a defensive rationale, reflecting a desire to strengthen their bargaining position vis-à-vis major trading partners and allow them to 'better

defend themselves against discriminatory effects of other regional groups' (Patterson 1966: 147). The formation of EFTA is an example. It was established in 1960 in reaction to the formation of the EEC. Its membership consisted of European countries that did not want to join the EEC because of concerns relating to the supranational aspects of the EEC and the likely level of the common external tariff (most EFTA countries tended to be relatively liberal). The EFTA reaction to the formation of the EEC was not unique. Japan informally proposed a Pacific Free Trade Area with the US, Canada, Australia and New Zealand in the mid-1960s for the same reason (De Melo and Panagariya, 1993). More recently, Pacific nations agreed to pursue regional free trade under auspices of the Asian-Pacific Economic Cooperation (APEC) framework. This is an example of so-called open regionalism, where PTAs are used as a focal point for concerted liberalization. Essentially this involves the formation of a privileged group (see Chapter 4). Free riding problems can be expected to be important in such efforts—it is unlikely that APEC will realize the stated goal of free trade by 2020 given the shift by many of the East Asian economies towards the negotiation of formal PTAs. However, as discussed by Baldwin (2006*b*), the resulting 'noodle bowl' may generate incentives for firms in the region to push more actively for regional, if not global, free trade.

Arguments suggesting PTAs may be detrimental to the trading system often revolve around some variant of the optimal tariff argument. As trade blocs expand, so does their market power and, at least in principle, their ability to influence the terms of trade in their favour. If successful, this is detrimental to the rest of the world. Although possibly true in some cases, it is not a well-founded generalization that PTAs will have an incentive to increase their tariffs against the rest of the world. For one thing, there are big differences between FTAs and customs unions. As already discussed, members of FTAs may have good reasons for lowering tariffs on nonmembers, as this reduces trade diversion.

Baldwin (2006*a, b*) argues that there may well be positive incentive dynamics resulting from hub-and-spoke PTAs. The domino effects noted previously may move more countries to lower trade barriers, as over time the 'balance of power' between export- and import-competing interests shifts in favour of those benefiting from a more open trade regime. Baldwin's theory of how this may play out starts from the premise that at a given point in time export interests see benefit in expanding access to locations where they can undertake parts of their product process and can get their government to negotiate a PTA. Another important part of the story is that the major players are the big markets—such as the EU—so that one result of the process is a hub-and-spoke system of PTAs. This essentially consists of a set of bilateral trade agreements. Because a hub-and-spoke system involves separate agreements between the hub (e.g. the EU) and the spoke countries, there is much scope to exclude 'sensitive sectors' from the coverage of each bilateral agreement (Snape, Adams and Morgan, 1993). Each spoke is likely to have comparative advantage in a somewhat different set of such sectors. If each country

maintains contingent protection options (AD, safeguards) against member countries, powerful import-competing industries in the hub country will have an interest in including wide-ranging safeguard clauses and relatively stringent rules of origin.

This was the case in the Association Agreements negotiated between the EU and various Eastern neighbours in 1992 (Winters, 1995). By allowing bilateral deals regarding sectoral coverage and the depth of the agreement, vested interests could be assuaged through specific rules of origin as well as safeguard provisions. The nature of these types of PTAs were such that they allowed for significant internalization of benefits by producers who wanted to restrict as much as possible the (new) regional market for themselves. Such groups also opposed broader multilateral reform. As summarized by Bhagwati (1993), such groups took the view 'the region is our market', and that 'our markets are large enough'.

In the event, however, the political economy equilibrium that underpinned the hub-and-spoke model of major 'systems' of PTAs began to break down. In part as a result of continuous technological change and in part as a result from increasing competition by China firms in Europe began to see an interest in further reducing the cost of production. One way this could be achieved was through reduction of the administrative costs of the hub-and-spoke system, in particular the associated rules of origin. One result was the adoption of the pan-European Cumulation System in 1997, under which any inputs sourced from *any* of the spokes or the EU member states counts for purposes of determining origin, and thus eligibility for duty-free treatment. This is an example of how regionalism may generate forces through which the objective of a multilateral, nondiscriminatory, trade regime might emerge endogenously.

The Baldwin (2006*a*) story is comforting for those who have been worrying about the systemic implications of PTAs. Saggi and Yildiz (2008) offer another, more theoretical, argument for the positive systemic effects of PTAs. They note that the voluminous literature on PTAs and regional integration ignores a key question: would the WTO serve the cause of global free trade more effectively if it did not include the exception to MFN provided by Article XXIV? Would global free trade be easier to achieve if all WTO members were somehow constrained to pursue trade liberalization on only a multilateral basis? The relationship between preferential and multilateral liberalization has been the subject of much theoretical analysis, but Saggi and Yildiz are among the first to treat both bilateral and multilateral liberalization as endogenous and to allow for the fact that countries are not symmetric in size. A central result is that bilateralism can provide an impetus to multilateral trade liberalization. The insight underlying this result is that a country that is choosing whether or not to participate in global free trade must consider how it would fare under the agreement that would emerge in the absence of its participation. Their model has the reasonable feature that a non-participating country is worse off under a bilateral trade agreement than under a

multilateral agreement—this is because a preferential agreement discriminates against the outsider whereas a multilateral agreement does not. As a result, a country's incentive to opt for free trade is stronger when the alternative to free trade is a bilateral agreement between the other two countries as opposed to a multilateral one. It is noteworthy that this result obtains in their model despite the fact that the formation of a PTA induces its members to impose lower tariffs on the nonmember relative to their Nash equilibrium tariffs, a result referred to as the tariff complementarity effect in the literature. In fact, even though a preferential agreement leads to more trade liberalization than a multilateral agreement in which all countries do not participate, it harms the outsider relatively more precisely due to the discrimination that is inherent to it.

The analysis of Saggi and Yildiz makes two additional points. First, they demonstrate that the debate regarding preferential versus multilateral liberalization is moot in the absence of some type of asymmetry across countries, which in their model implies that the gains generated by a shift from the status quo of noncooperative tariffs to free trade are unequally split across countries.⁵ Indeed, they show that under sufficient symmetry, both the preferential and the multilateral route lead to global free trade. A second important insight provided by their analysis is that to properly address the issue of how preferential trade liberalization interacts with multilateral trade liberalization, we need to better understand when and why countries choose to pursue the preferential route given that the multilateral route is available. Only a model in which both types of liberalization are fully endogenous can shed real light on this question.

Deeper integration and PTAs

So-called deep integration has become an increasingly important feature of PTAs over the last decade and a half as border barriers decline. This spans many aspects of product and market regulation, including standards, government procurement, services, investment, competition, labour and environmental policies, as well as IPRs and protection of other intangible and tangible assets.

As is illustrated by Figures 10.2 and 10.3, there is enormous variance across recent PTAs in the scope and depth of these policy areas, with PTAs that involve the United States generally having the broadest coverage. There is evidence that PTAs include more service sectors than countries have scheduled at the WTO, but the available research also suggests that their clauses do not move much beyond those in the WTO (Fink and Molinuevo, 2007; Roy, Marchetti and Lim, 2007). This

⁵ In a model with repeated interaction, Saggi (2006) had shown that when countries are asymmetric, the exogenous formation of a preferential trade agreement may facilitate multilateral tariff cooperation whereas such a result does not obtain under symmetry.

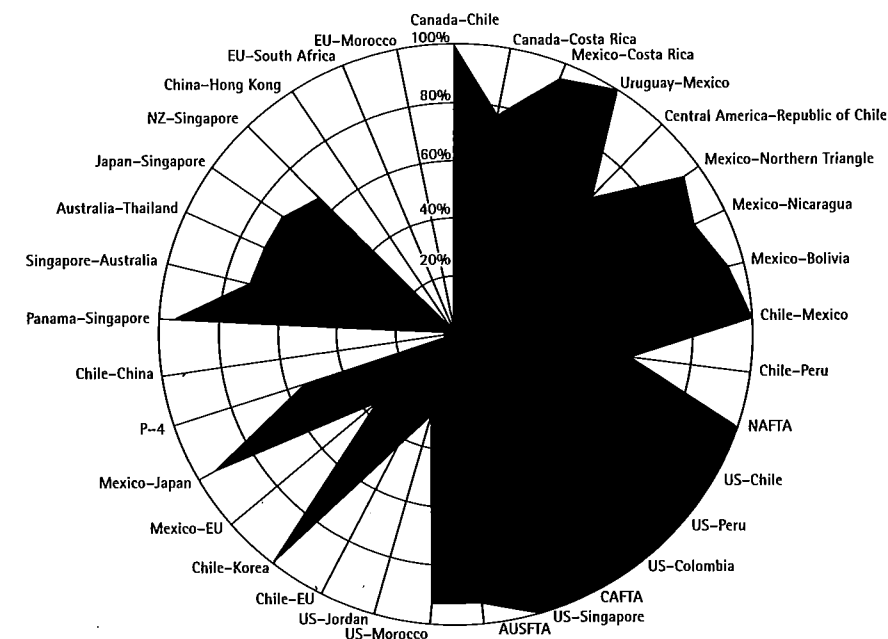


Fig. 10.2. Coverage of 17 investment provisions in selected PTAs

Source: Estavadeordal, Shearer and Suominen 2007.

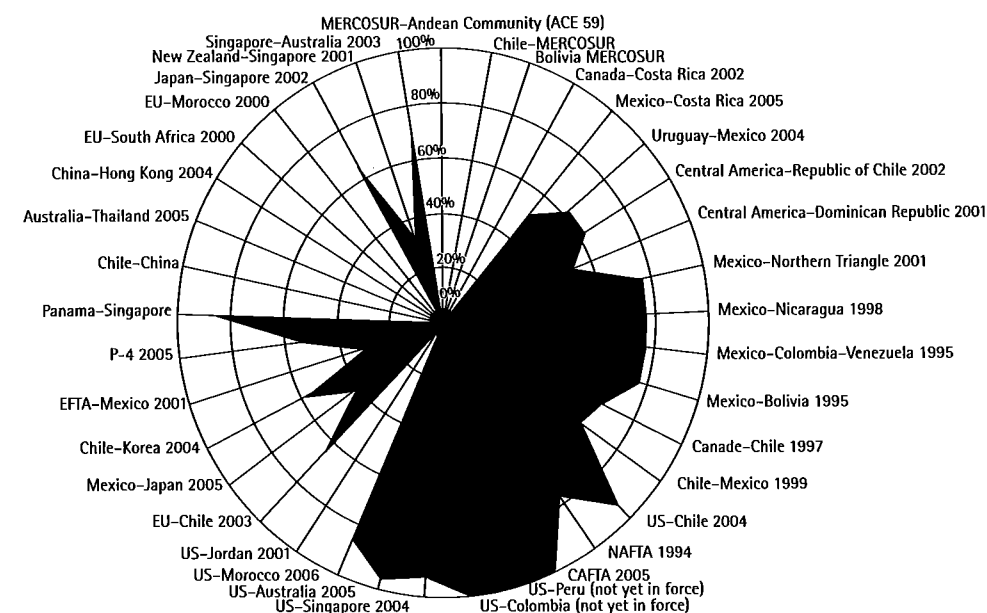


Fig. 10.3. Coverage of 29 services provisions in selected PTAs

Source: Estavadeordal, Shearer and Suominen (2007).

suggests that PTAs may broaden the coverage of commitments to lock-in service liberalization, most of which has been implemented autonomously, but they do not deepen it. It is also the case that many PTAs now cover FDI, whereas the WTO does not. But here again one can question the extent to which disciplines are additional. For example, there are already over 3,000 bilateral investment treaties in place. Moreover, in areas of key importance to the developing countries such as mobility of labour or constraints on the ability of OECD partner countries to offer incentives to investors the bilateral routes have not achieved more than what was possible in the WTO.

How much further than the WTO do PTAs go? Horn, Mavroidis and Sapir (2008) assess the coverage of 14 US and 14 EU PTAs, focusing on the prevalence of WTO+ commitments (provisions that address matters covered by the WTO but that go beyond the WTO in terms of extent of disciplines imposed) and what they call WTO-X commitments: provisions on matters that are not dealt with in the WTO. They conclude that the EU and the US both negotiate WTO+ disciplines in many areas, with the US taking the lead. With the exception of the (early) agreement with Israel and the PTA with Jordan, US PTAs have additional disciplines in almost all the categories distinguished by the authors: industrial market access; agriculture; customs; export taxes; SPS; TBT; STEs: AD; CVDs: subsidies; procurement; TRIMs; services and IPRs. Moreover, these WTO+ disciplines are legally binding (enforceable). In the case of the EU PTAs, the areas with WTO+ disciplines are fewer, in particular for export taxes, SPS, TRIMs and services.

The EU PTAs in contrast have much greater coverage of WTO-X provisions, with some PTAs covering over 30 policy areas that are not included in the WTO. Examples include competition policy, environmental laws, investment, IPRs, capital movement, consumer and data protection, cultural cooperation, education, energy, health, human rights, illegal immigration, illicit drugs, money laundering, R&D, SMEs, social matters, statistics, taxation, and visa and asylum policies. The US PTAs are much less focused on WTO-X policy areas, but at the same time there is also much less variance in the subjects that are covered. They are limited to anticorruption (not found in any EU PTA), competition policy, environmental laws, investment, IPRs, labour market regulation and capital movement. Although the EU is clearly the 'market leader' when it comes to WTO-X provisions, about three-quarters of the relevant articles in the EU PTAs do not impose binding disciplines. Instead, they tend to constitute soft law—technical assistance and cooperation. In contrast, the disciplines in WTO-X areas found in US agreements are generally legally binding (that is, enforceable). This reflects a distinct difference in the strategies and approaches pursued by the EU and the US. The US relies more on binding agreements and legal enforcement, the EU has tended to put the emphasis on embedding technical assistance and other forms of cooperation in its PTAs, and supplementing this with financial aid and policy/political dialogue.

This 'softer' approach is a basic premise of the EU's European Neighbourhood Policy, for example (Hoekman, 2007). Which approach will prove to be more influential in setting standards remains to be seen.

From the perspective of members a key question is whether 'deeper integration' provisions are beneficial in and of themselves, or simply a cost that must be paid to obtain preferential access to major markets. We return to that question in Chapter 13 as it is one that obtains in the WTO context as well. From the perspective of nonmembers and the trading system the basic question is the same as discussed above: what is the impact on them of deeper integration among subsets of countries in PTAs? This depends both on the extent of discrimination that is implied and on the scope for cooperation in a PTA context to increase the likelihood that the norms set by PTAs become the *de facto* standard or the focal point for subsequent multilateral cooperation in the WTO (or elsewhere).

The prospects for the 'multilateralization' of PTA commitments in these areas may be significant. In many cases regulation is quite naturally applied in a nondiscriminatory fashion, treating domestic and all overseas suppliers or firms equally—where 'domesticity' is defined more frequently in terms of location of production than ownership. This is quite different from tariffs and NTBs affecting trade in goods, where domestic/foreign and intra-foreign discrimination is the objective. From the perspective of achieving regulatory objectives, nationality often will (and should) not matter. But even if regulation applies to all sources of supply, it can still have the effect of segmenting markets and reducing competition.

If liberalization—defined as taking actions to enhance the contestability of a market—is more likely to be nondiscriminatory for regulations than for merchandise trade barriers, it is, equally, less likely to come about at all. This is because it is inherently more far-reaching and because it is simultaneously necessary and very difficult to distinguish between regulations that are genuinely needed for the achievement of domestic objectives and those that are oriented towards segmenting markets and protecting domestic incumbents. In practice it is certainly not inevitable that regulations are applied on a nationality-blind basis—insofar as protectionism is an objective of policymakers, regulation can be (and is) used to achieve this. One reason is that the legitimate, nonprotectionist class of regulation frequently requires the acquiescence of domestic firms if it is to be implemented effectively and almost always entails consulting those firms about any reforms. With the complex and subtle nature of many regulations, incumbents (and national regulators) will have a great deal of influence over regulatory structures and details, and may well have veto power over policymakers.

For cooperation on product market regulation and domestic policies in PTAs one can envisage three different processes of multilateralization (Hoekman and Winters, 2009). First, hegemonic multilateralization: a hegemonic economic power is essentially able to impose its own model on its partners, not necessarily

coercively but by the force of its market size. As different partners adopt the hegemon's approach over their own local one, a degree of multilateralism is achieved. And it is possible that as the partners enter further bilateral or regional arrangements with other partners the model is extended. As Schiff and Winters (2003) observe, the accretion of two different groups of supporters around two different models—say a US and a EU model—could make the final multilateral step (harmonization or recognition of equivalence) less rather than more likely. But, if a high degree of similarity or consistency is achieved, goods and services designed for one market can be sold elsewhere, greatly increasing the contestability of markets. Examples of the hegemonic model abound in 'deep integration'. The US requires partners in Bilateral Investment Treaties (BITs) to conform to an identical template and imposes its own intellectual property right (IPR) protection provisions in its PTAs—World Bank (2005). Another example is the EU interest in extending its system of geographical indications through its PTAs (see Chapter 8).

The second route to multilateralism is a convergence route. This operates within a PTA where the erosion of barriers to trade increases the pressure to harmonize regulations because they start to have greater impact on trade patterns, competitiveness and profitability. This is essentially the 'competition between rules' that featured in the EU's Single Market programme, which applies equally to goods and services. It depended, in the case of goods, not only on the removal of traditional barriers to intra-EU trade (tariffs, quotas, etc.) but on the aggressive policy of the European Commission and European Court of Justice towards other limitations on the freedom of movement of goods such as product standards (where the principle of mutual recognition was applied). In services the political sensitivity of the convergence route is evident in the constrained liberalization of cross-border services espoused by the recent Services Directive in the EU and the difficulties that have affected efforts by the EU and the US to make progress in moving towards accepting each other's regulatory norms for specific services as being effectively 'equivalent'. Note that the convergence route also spans a frequently mentioned rationale for PTA-based cooperation in 'non-WTO' areas: PTAs may be a useful forum for experimentation and learning. Successful examples of cooperation between members of a PTA may be adopted in other PTAs (or unilaterally), thus promoting multilateralization over time.

In general, the larger the region or the more important it is as a trading partner, the greater the incentives for a country to adopt the regulatory standards of the PTA. There will often be a link, implicit or explicit, between harmonization of regulatory regimes and the threat of contingent protection. One factor driving harmonization is to reduce the possibility of being confronted with contingent protection. As PTAs increasingly are instruments for such regulatory harmonization, or for the adoption of mutual recognition procedures, the potential cause for concern on the part of nonmembers is obvious. As discussed in Chapter 13, one size fits all is not necessarily optimal.

The third route to multilateralism is the one identified by Baldwin (2006a) for trade in goods—what could be called a political evolution route. Here, changes in the political weight of different parties or in the relative importance of different costs change the political economy so that groups which once sought to segment markets now seek to integrate them. This route is, of course, premised on policies being applied in a discriminatory manner vis-à-vis nonmembers of a PTA and thus will not apply (and will not be needed) if policies are applied on an MFN basis. One difference for deep integration is that, compared with restrictions on goods trade, regulations are complex and require greater complicity from the relevant industry. The strong position of incumbents may make liberalization more difficult; in particular, it is difficult to envisage incumbents in a sector seeking the liberalization of that sector. However, offsetting this is that downstream (using) sectors may have stronger incentives to oppose policies that raise the costs of services than is the case with goods.

Hoekman and Winters (2009) argue that when it comes to 'deeper integration' in PTAs, to date most reform is unilateral. There is very little direct evidence that PTAs do a lot to drive reform. One problem is to determine the direction of causality. One cannot infer from the spread of specific PTA disciplines ('templates') that PTAs are driving reforms beyond what governments had already decided was beneficial autonomously. Most research in this area focuses on legal texts, not on extent to which PTAs imply/require or result in changes in national legislation. It may well be that the source for reform has primarily been knowledge and information—the demonstration effects of successful countries or the general focus of academia and the international community on the benefits of deregulation, competition, etc. Maybe the IMF, World Bank, OECD, APEC, etc., which have been advocating for better policies and more transparency for years are the key: the World Bank's *Doing Business* report may well have been a more potent driver of reforms than the PTAs of which countries were members.

Even if developing countries are adopting disciplines in PTAs and applying them on a MFN basis, this does not imply that the norms concerned are beneficial for them. The content of the norms that are included in PTAs obviously matter. Whether these are autonomously decided or externally imposed, they need to benefit the countries that adopt them and the countries affected by them (the nonmembers). From this perspective, another important priority is the establishment of institutions or other means to help developing countries take an informed view of what they are asked to do in PTA negotiations and how neighbours' PTAs impact upon them. The pressure by high-income countries for developing country PTA partners to adopt TRIPS-plus disciplines is an example.

Finally, the opportunity costs of PTAs in terms of taking up scarce administrative capacity of developing country governments needs to be recognized. One can question whether the negotiation of PTAs is the best use of the limited policy attention and human resources that are available in many low-income countries (Schiff and Winters, 2003). Much depends on the content of a PTA and the strategic

use that is made of it. Preferential trade agreements that are not implemented or not used as instruments to realize substantial economic benefits can have a significant opportunity cost in terms of diversion of policy attention and capacity.

10.4. CONCLUSION

Always a central element of the trade policy strategy of European countries, 'regionalism' has become an important form of international cooperation on trade policy for virtually all the members of the WTO, developed and developing. Although subject to conditions contained in Articles XXIV GATT and V GATS, multilateral disciplines are not enforced. On a number of dimensions they are also weak. An example is the absence of any disciplines with respect to preferential rules of origin in the WTO (Box 10.3). Another is the absence of a requirement that PTAs be open to new members (Bhagwati, 1993). Multilateral surveillance is limited—even if the CRTA were effective, the focus is on WTO tests and not on the economic effects of PTAs. The WTO Secretariat has no mandate to monitor the trade value or terms-of-trade effects of PTAs. Developing countries can opt out of WTO disciplines on PTAs altogether by invoking the Enabling Clause.

Most of the economic literature on PTAs has been theoretical or policy-focused. Rigorous empirical research on the effects of PTAs has been limited until recently. The weight of the empirical analyses of PTAs suggests that if one abstracts from the many PTAs that were never implemented or were designed to have no effect in opening up economies, overall the benefits outweigh the costs, especially if the focus of attention is on the dynamic effects over time. Preferential trade agreements may lead or retard nondiscriminatory reductions in tariffs, but the evidence that is now emerging suggests that PTAs have complemented multilateralism in the sense of promoting lower barriers against the rest of the world.

The proliferation of PTAs has been accompanied by steadily declining barriers to trade generally and high growth rates in world trade. The uniform tariff equivalent of all applied most-favoured-nation tariffs of high-income OECD countries in 2005 was 4.8 per cent. Excluding agricultural products, the figure drops to 2.7 per cent (Kee, Nicita and Olarreaga, 2008). For the developing countries, applied MFN tariffs have also fallen substantially; Kee and colleagues estimate that the median average overall trade restrictiveness index was 7.5 per cent for the 57 countries for which data are available in 2005, compared with 12.3 per cent ten years earlier. Much of this liberalization trend is not due to PTAs. The fact that MFN tariffs have fallen significantly in almost all countries, whatever their participation in PTAs, suggests that unilateral decisions to liberalize have been paramount (Martin and Messerlin, 2007).

Whatever the overall impact of individual PTAs on trade in goods, prices and welfare, to date they do not appear to have had a serious negative effect on the overall trend towards greater openness of economies around the globe. Indeed, a good case can be made that the spread of PTAs has done little to substantially reduce the extent to which trade flows are distorted by preferences, as most of the recent PTAs are (a) between small countries; (b) between countries that have (very) low external protection and many zero-rated MFN duties; and/or (c) exclude those sectors where there is significant protection, e.g. agriculture (Pomfret, 2007).

Restrictive rules of origin and cumulation criteria further reduce the impact of many PTAs. The combination of the exclusion of agriculture and restrictive rules of origin is particularly important in EU PTAs with countries that are not accession candidates. The large export supply response following the implementation of preferential access agreements between the US and some African and Middle Eastern countries—e.g. Lesotho and Jordan—illustrates the importance of including agriculture and using liberal rules of origin for labour-intensive manufactures such as apparel. Both Lesotho and Jordan have preferential access agreements with the EU but have not seen exports to the EU expand because of more restrictive rules of origin than are applied by the US to these two countries.

There is a good case to be made that when it comes to tariffs the evidence is beginning to support those who argue that on net the recent vintages of PTAs have been building blocks. But tariffs increasingly are not at the core of the PTA action, given that average MFN duties have fallen to low levels in major markets. Preferential trade agreements may embody many good practices and some go far beyond the WTO in terms of liberalizing markets. Thus, in the EU there are no tariffs, no safeguard mechanisms and full binding of policies. To a large extent the current benchmark for good practice in trade policy is the set of policies and rules that apply to movement of goods, services, labour and capital inside the EU. The challenge is to pursue multilaterally what the serious PTAs are implementing internally. This has been the trend. Indeed, it appears that developments in PTAs are frequently reflected in analogous developments on the multilateral front. Differences between the PTAs and the multilateral trading system at any point in time have been limited in part because efforts to negotiate PTAs have stimulated concurrent, and largely successful, efforts to achieve further multilateral trade liberalization. Indeed, the multilateral system often leads (Hoekman and Leidy, 1993; Pomfret, 2007).

Preferential trade agreements represent a challenge and an opportunity for the multilateral trading system. The opportunity is to use them as experimental laboratories for cooperation on issues that have not (yet) been addressed multilaterally, especially issues where the outcome is applied on a MFN basis. The challenge is to control the discrimination that is inherent in any PTA. The inability of the CRTA to come to decisions on whether PTAs satisfy Articles XXIV and V is a problem in this regard. Absent such determinations, transparency—through

multilateral surveillance—is important, as this helps mobilize domestic and regional groups who are negatively affected by PTA policies.

Both the GATT and GATS contain provisions relating to transparency and multilateral surveillance. Countries intending to form, join or modify a PTA must notify this, and make available relevant information requested by WTO members. Although CRTA efforts to determine the consistency of the agreement with multilateral rules are not effective, they do generate information, especially since 2006 when a new transparency mechanism was established by the WTO General Council. The key need may not be more multilateral disciplines, but greater internal scrutiny by stakeholders in member countries of regional trade policy to ensure that the interests of all groups in society are considered. Although multilateral surveillance can be helpful as an objective source of information and analysis, ultimately, domestic transparency requires domestic political will. Multilateral trade negotiations can and should play a complementary role. At the end of the day, the more successful the WTO is in reducing external barriers of members through MTNs, the less problematical PTAs will be from a systemic and non-member perspective. Attempts to impose stricter rules on PTAs, or to use the dispute settlement system, are unlikely to be fruitful strategies.

10.5. FURTHER READING

The classic original treatment of the economics of regional integration is Jacob Viner, *The Customs Union Issue* (New York: Carnegie Endowment for World Peace, 1953). For an excellent historical discussion of the issue of regionalism and preferential liberalization in the GATT context, see Gardner Patterson, *Discrimination in International Trade: The Policy Issues, 1945–1965* (Princeton: Princeton University Press, 1966).

The literature on the economics of regional integration is surveyed by Richard Baldwin and Anthony Venables, 'International Economic Integration', in Gene Grossman and Kenneth Rogoff (eds), *Handbook of International Economics*, vol.3 (Amsterdam: North Holland, 1997) and by Arvind Panagariya, 'Preferential Trade Liberalisation: The Traditional Theory and New Developments', *Journal of Economic Literature*, 37 (June 2000): 287–331. WTO, *Changing Landscape of Regional Trade Agreements* (Geneva: WTO, 2007) surveys the content and coverage of PTAs and provides a number of relevant quantitative and qualitative indicators.

The relationship between regionalism and multilateralism is the subject of L. Alan Winters, 'Regionalism Versus Multilateralism', in Richard Baldwin, Daniel Cohen, André Sapir and Anthony Venables (eds), *Market Integration, Regionalism*

and the Global Economy (London: CEPR, 1999). Maurice Schiff and L. Alan Winters, *Regional Integration and Development* (Washington, DC: World Bank, 2003) offer a thoughtful and in-depth assessment of the impact of PTAs on developing countries and the policy options confronting developing country governments.

For a characteristically insightful analysis of the implications of the increasing emphasis on PTAs by governments, see Jagdish Bhagwati, *Termites in the Trade System: How PTAs Undermine Free Trade* (Oxford: Oxford University Press, 2008).