DRAFT

Trade in Services: Economics and Law

Aaditya Mattoo

April 2000

Structure of the presentation:

I. Introduction

II. Services, the modes of delivery and the scope of the GATS

III. The instruments of protection

IV. Discrimination in services trade, MFN and the GATS

V. Domestic regulations

VI. A summing up

I. Introduction

Some stylized facts about services:

- Services account for an increasing share of production and employment in both developed and developing countries. Many of the fastest growing sectors are services (telecommunications, health, finance);

- The share of services in world trade and investment has been increasing. They have been among the fastest growing components of world trade over the last 15 years. Services trade, as estimated from balance of payments statistics, was greater than $2 trillion last year, representing over one-fifth of world trade in goods and services. This value is widely viewed as understated, because much "trade" in services takes place through an established presence, i.e via FDI, and hence generates local activity and value added that do not appear as exports in balance of payments statistics.

- Today, more than half of annual world FDI flows are in services, and the value of sales abroad by foreign affiliates of US service firms is estimated to be 3.5 times greater than their cross-border exports.

Why is trade in services growing?

There are two main reasons:

1. Technological progress, especially in telecommunications and information technology.

2. Broad trend towards liberalization/regulatory reform in key service industries; starting with the Thatcherite revolution in the UK in late 1970's, pursued in the US (air, rail, break-up of telecom monopolies); embracing of privatization and competition around the world in key infrastructural service sectors (telecoms, transport, finance).

What is special about trade in services?

(i) Trade in services is in some important ways different from services. First, many services require proximity between the supplier and the consumer, and hence factor mobility is necessary for international transactions. Secondly, the limited scope for "border" restrictions implies that domestic regulations have a much stronger influence on trade in services. A key difficulty is distinguishing between regulations that are necessary (and even need to be strengthened) to achieve legitimate objectives, e.g. financial stability, competitive market conditions, and universal service, and those that have a protectionist impact.

(ii) Although some interesting twists arise because of the way services are traded and regulated, the basic insights from the theory of trade in goods apply to trade in services. There are likely to be substantial gains from liberalizing trade in services, immediately and in the longer term.

(iii) The WTO rules on services trade, embodied in the General Agreement on Trade in Services (GATS), represent a significant achievement in terms of creating a basic framework, but have not so far produced significant liberalization. A new round of services negotiations has just begun and represents a valuable opportunity to go further, both in terms of liberalizing commitments and in terms of improved multilateral rules.

II. Services, the modes of delivery and the scope of the GATS

What are services?

It may be useful to begin by asking: how are services different from goods? Services are often seen as intangible, invisible and perishable, requiring simultaneous production and consumption. Goods, in contrast, are tangible, visible and storable – and hence do not require direct interaction between producers and consumers. However, there are exceptions to each of these characteristics of services: a software program on a diskette or an architect's design on paper are both tangible and storable, many artistic performances are visible, and automated cash-dispensing machines make face-to-face contact between producers and consumers unnecessary. These exceptions do not however detract from usefulness of the general definition of services presented above.

Instead of worrying about precise definitions, the GATS proceeded simply to list the entire range of services covered:

1. Business services
2. Communication services
3. Construction services
4. Distribution services
5. Educational services
6. Environmental services
7. Financial services
8. Health-related and social services
9. Tourism and travel-related services
10. Recreational, cultural and sporting services
11. Transport services
12. Other services not elsewhere included

How are services traded?

It is possible to distinguish between services that necessarily require physical proximity between the user and the provider and those that do not – though proximity may enhance the quality of the service even in the second case. The proximity requiring services include: construction, where it is necessary for the supplier to move to the location of the consumer; tourism, where it is necessary for the consumer to move to the location of the supplier; and hair cuts or surgical operations, where either the supplier and the consumer could move. But, of course, for many services – whose number is growing with the development of electronic means of delivery – proximity is not necessary. These include a variety of financial, entertainment and communication services.

Nevertheless, the proximity aspect of many services transactions has a crucial implication for the nature of services trade. The need for proximity creates the need for factor mobility; i.e. trade in many services is inextricably linked to foreign direct investment, labour movement or both.

Accordingly, the GATS defines trade in services (in Article I) to include four modes of supply:

- *Cross-border*: services supplied from the territory of one Member into the territory of another. An example is software services supplied by a supplier in one country through mail or electronic means to consumers in another country.

- *Consumption abroad*: services supplied in the territory of one Member to the consumers of another. Examples are where the consumer moves, e.g. to consume tourism or education services in another country. Also covered are activities such as ship repair abroad, where only the property of the consumer moves.

- *Commercial presence*: services supplied through any type of business or professional establishment of one Member in the territory of another. An example is an insurance company owned by citizens of one country establishing a branch in another country.

- *Presence of natural persons*: services supplied by nationals of one Member in the territory of another. This mode includes both independent service suppliers, and employees of the services supplier of another Member. Examples are a doctor of

one country supplying through his physical presence services in another country, or the foreign employees of a foreign bank.[1]

International trade in services through the movement of factors

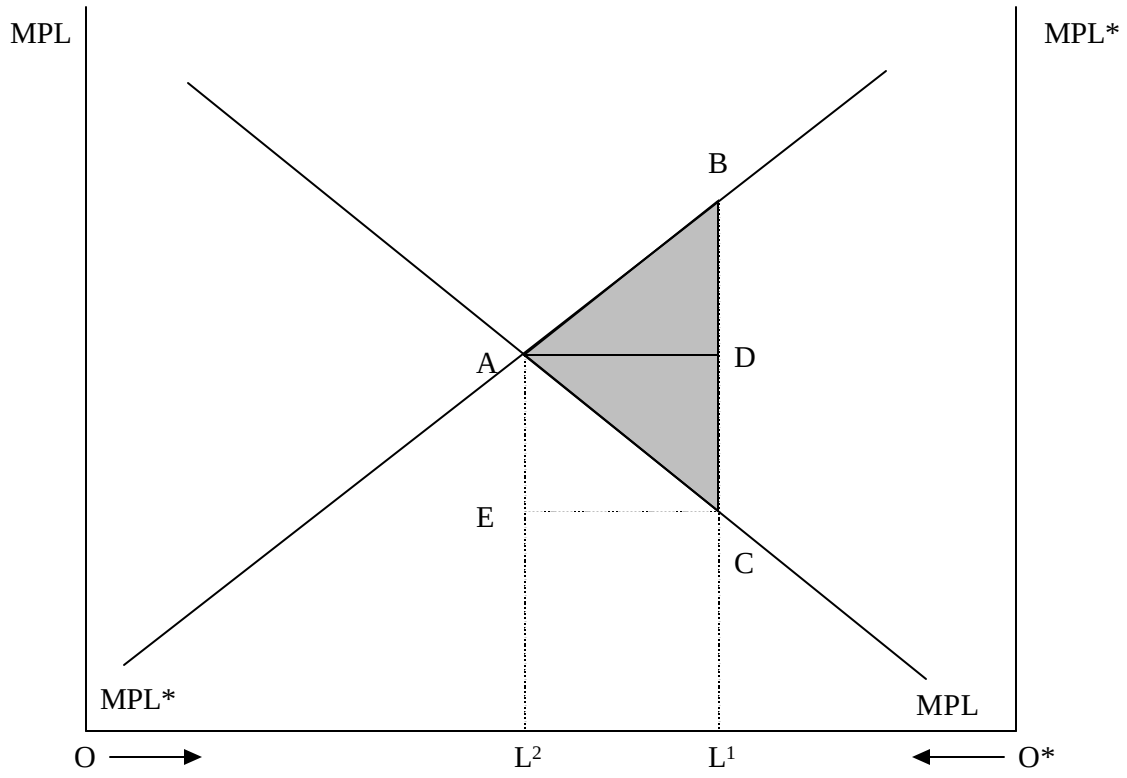
If trade must take place through the movement of factors, are the basic propositions of trade theory – based on the notion of cross border trade – put into question? It turns out that the economics is not much different, but the politics can be.

One problem does arise from the point of view of positive theory: if different modes of supply are close substitutes, it is not easy to predict whether comparative advantage will manifest itself as a trade flow, investment flow or labour flow.

However, from the point of view of normative theory there is no obvious problem: a country gains from import of services, irrespective of the choice of mode, if the terms at which international transactions take place are more favourable than those available on the domestic market.

DIAGRAM 1

[1] The GATS does not apply either to measures affecting natural persons seeking access to the employment market of a Member, or to measures regarding citizenship, residence or employment on a permanent basis.



We illustrate this with an analysis of factor mobility. Assume a world where there are two countries, "home" and "abroad" and two factors of production: capital and lawyers. Capital and lawyers together produce legal services which can only be traded by the movement of lawyers to the location of the consumers. Capital is assumed to be immobile but the same analysis applies to capital mobility.

In the diagram, the width of the box is equal to the total lawyer endowment of the two countries. Every point along the horizontal axis represents an allocation of lawyers between home and abroad. The number of lawyers at home are measured from the origin 0, and the number of lawyers abroad is measured from the origin 0*. Suppose initially there are $0L^1$ lawyers are at home, and the remaining 0^*L^1 lawyers are abroad. The two lines, MPL and MPL*, represent the marginal products of lawyers in each country, holding the amount of other factors and know-how constant. MPL and MPL* are assumed to be declining in the number of lawyers.

Now assuming that factors are paid their marginal product, lawyers will be paid L^1C at home and L^1B abroad. By adding up the marginal product of each lawyer in a particular country, we can find out the total product (that is, output or GDP) of that country. This equals the area under the MP curve for that country, up to the number of lawyers it has. For example, the output of home is everything under MPL to the left of point L^1, while the output abroad is everything under the MPL* curve, to the right of point L^1.

Note that with this initial allocation, the return to lawyers is lower at home than abroad. If lawyers can move, some will leave home and go abroad to take advantage of these higher returns. This flow from home to abroad will stop when the return to lawyers is equal in

both countries, which will occur where the MP curves cross and L^1L^2 lawyers have moved. At this point, home will have OL^2 lawyers, and abroad will have 0^*L^2 units. The equalized return will be L^2A.

Now consider the impact of the movement of lawyers on aggregate economic welfare and on income distribution, both internationally and nationally.

Global welfare

Abroad's output is everything under its MP curve up to point L^2; it has risen by area AB L^1L^2. At the same time home's output has declined by area ACL^1L^2. Total world output has risen by the area ABC. Allowing factors to move to their most productive locations has clearly increased global welfare.

Total gains and losses:

Gain of lawyers who move: ADCE
Loss of capital at home: ACE
Gain of capital abroad: BAD
World gain: ABC

International income distribution

World welfare improves. But whether welfare improves for home, the country of emigration, depends on whether the welfare of moving lawyers is included. If such movement is permanent migration, and excluded from the calculation of national welfare, then welfare declines at home. However, if such movement is temporary, or if migrants make substantial remittances, then welfare can increase at home as well.

National income distribution

The lawyers who originally worked in home receive higher returns, but those who originally worked abroad receive lower wages. Capital abroad benefits from the increased supply of lawyers but capital at home is made worse off.

Provision of services through the movement of individuals like trade is driven by international differences in resources. Furthermore, like trade it increases world production but is associated with strong income distribution effects that make liberalization difficult to accomplish.

Dynamic benefits of services trade liberalization

It is not easy to model the dynamic gains formally, but there are strong intuitive reasons to believe that well functioning service industries contribute to growth in different ways. An efficient financial sector allows resources to be deployed where they have the highest returns. Improved telecom efficiency generates economy-wide benefits as telecommunications are a vital intermediate input and are crucial to the diffusion of

knowledge. Similarly, transport services contribute to the efficient distribution of goods within a country, and greatly influence a country's ability participate in global trade. Business services such as accounting and legal services are important in reducing transaction costs; education and health services are necessary in building up the stock of human capital, a key ingredient in long run growth performance.

The growth effects of services liberalization also arise from allowing movement of factors of production. A country that liberalizes its services sector is likely to augment its stock of capital (through increased FDI) and crucially the stock of human capital and technology that is embodied in or associated with such FDI. The impact of this on long run growth is unambiguously positive. Furthermore, there is evidence that the presence of foreign factors can help enhance the productivity of domestic resources. This is as true for developing country capital importers as for developed country importers of skilled labor services. The contribution of imported skilled labor to the high-technology sectors in the US is now widely recognized.

III. The instruments of protection

Typical restrictions on services trade

In order to see how far these prescriptions relevant in the services context, let us briefly consider the typical trade-restrictive measures in services. First, *tariffs* are difficult to impose on services imports because they are often not delivered through the cross-border mode, and even when they are, it is usually in an intangible form. *Quotas*, on the other hand are pervasive. On cross-border trade, they are most evident in the transport sectors. Foreign providers are either completely shut out (i.e. a zero quota) of certain segments, such as cabotage, or only provided limited access, as in international transport. On consumption abroad, quotas are sometimes implemented through foreign exchange restrictions; e.g. the ability of citizens to consume services, such as tourism and education, abroad is limited by limits on foreign exchange entitlements. On commercial presence, quotas are imposed on the number of foreign suppliers who are allowed to establish in sectors like telecommunications and banking. Quotas on foreign participation also take the form of restrictions on foreign equity ownership in individual enterprises. Finally, quotas are perhaps most stringent in the case of movement of service-providing personnel, and affect trade not only in professional services, but also in a variety of labour-intensive services.

Apart from quotas, there are numerous *internal policies that discriminate* against foreign providers. These include measures that directly provide a cost-advantage to domestic providers, such as subsidies. And other measures that impose a cost or create a competitive disadvantage for foreign providers. Most obvious are internal (direct or indirect) tax instruments. Less visibly, foreign suppliers may be provided less favourable access to essential facilities such as ports, airports or telecommunications networks. General incentives (e.g. tax concessions) for use of local content may cause domestic producers to favour the use of locally-produced services (AUTOPACT).

Finally, there is the class of measures which are neither quotas, nor do they discriminate explicitly against foreigners but they can nevertheless impede trade. These include a variety of *domestic regulations* which we examine in the final section.

Economics of services trade policy

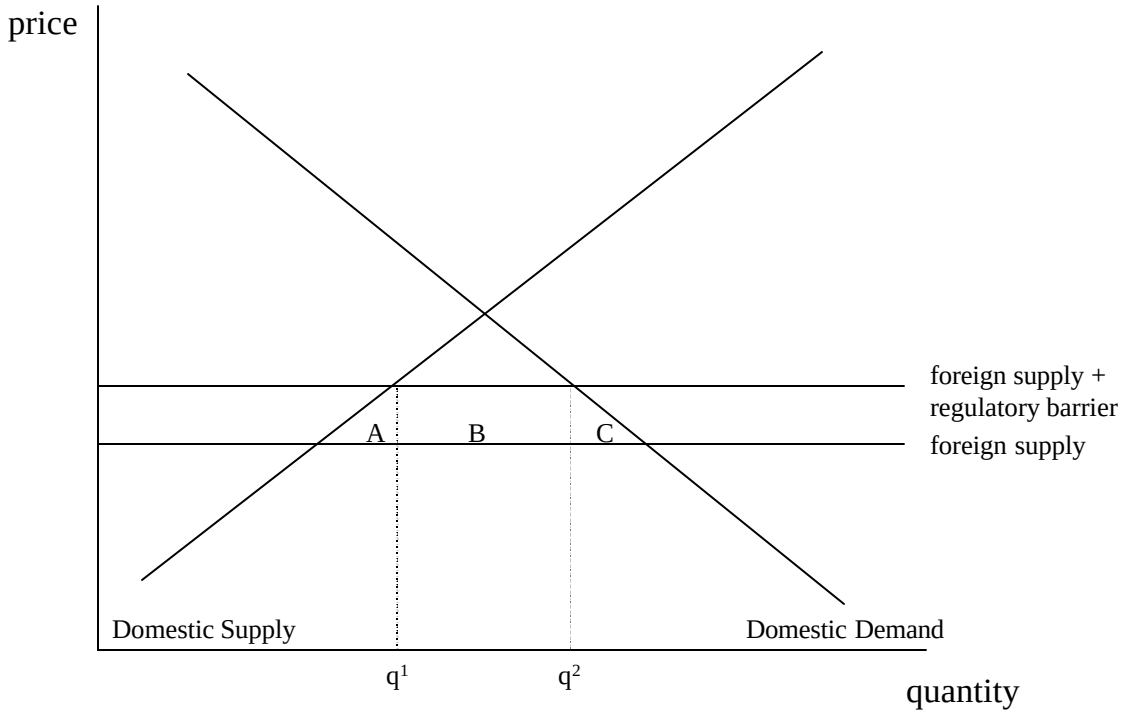
We have relatively straightforward policy rankings of instruments affecting trade in goods. First, if the object of policy is to sustain the output of an import-competing industry at a higher level than would otherwise be the case, then a subsidy to the output of the industry is superior to a trade restriction, such as a tariff. Secondly, if for some reason it is necessary to restrict imports, than a measure such as a tariff is generally to be preferred to a quantitative restriction – especially when quota rents are appropriated by foreigners, there is scope for rent-seeking behaviour or markets are imperfectly competitive.

How can the prescriptions of the standard goods theory help us to navigate through the spectrum of instruments described above? First, the superiority of subsidies over trade restrictions continues to be valid. Both encourage national production, the former by reducing the private costs of national producers and the latter by imposing a cost on foreign service providers. The latter is an inferior instrument because it leads to a deterioration in the price-quality mix that foreigners are able to provide local consumers. Secondly, in principle, tariffs are to be preferred to quotas for much the same reason as in the case of goods.

But there are at least three reasons why differences arise. The first reason is that the difficulty of imposing tariffs implies that the substitution of a more desirable policy instrument for a less desirable one is not always feasible. Secondly the instruments which have a tariff-like effect in terms of increasing costs of foreign providers, are not however tariff-like in generating revenue. Finally, the range of restrictions that are imposed on foreign direct investment and the movement of personnel directly affect the market structure.

Consider the consequences of restrictive measures that increase foreign costs but do not generate revenue. Now part of the loss in consumer surplus is not offset by increase in tariff revenue. So loss in welfare is much greater. Similarly, when quotas are imposed, their welfare consequences could be alleviated if the rents generated accrued to domestic importers rather than foreign exporters. But the difficulties of intermediation in services suggest that quota-rents are more likely to be appropriated by exporters.

DIAGRAM 2

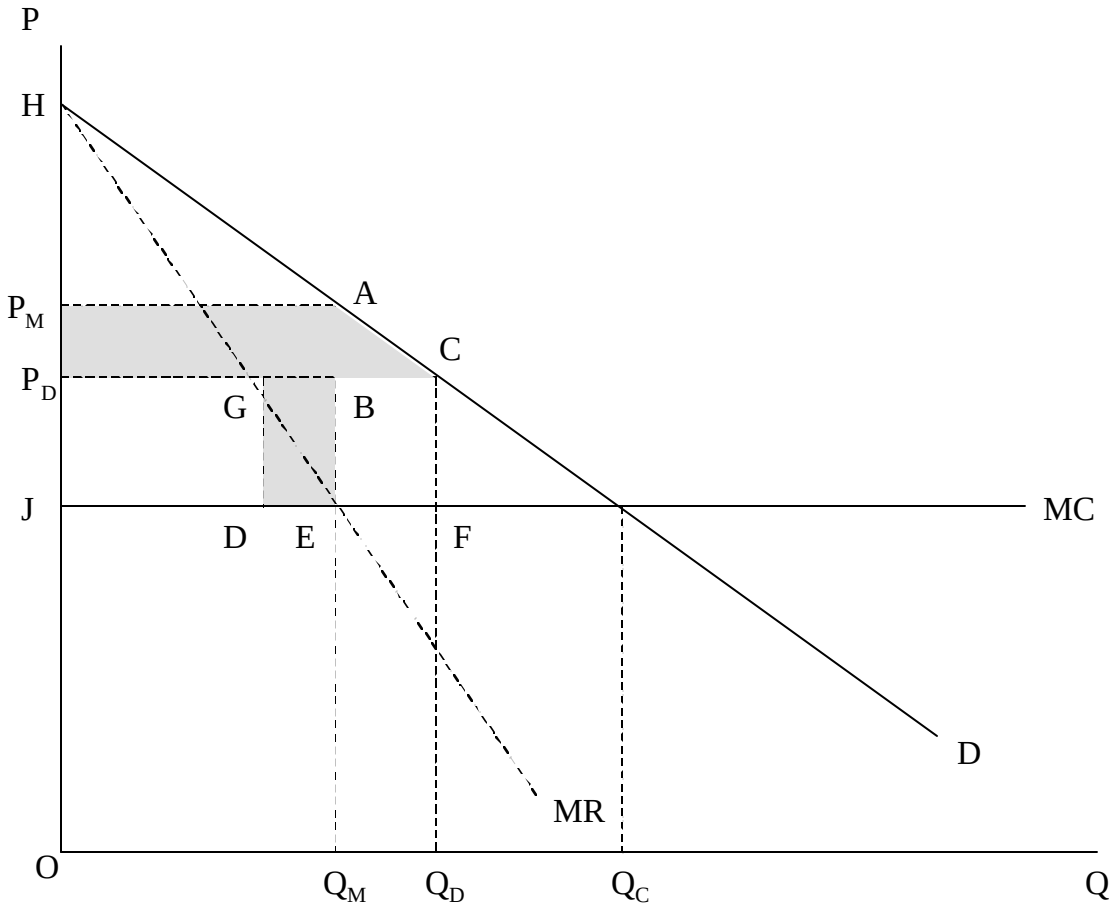


If complete liberalization is not feasible, a shift from both quotas and non-revenue generating measures to fiscal measures would lead to an increase in welfare.

The implications of alternative restrictions on FDI

Restrictions on foreign investment assume particular significance in the case of services where cross-border delivery is not possible, so that consumer prices depend completely on the domestic market structure. Restrictions on new entry and on the participation of foreign capital are the most common, particularly in communications and financial services. A basic conclusion from the literature on privatization is that larger welfare gains arise from an increase in competition than from simply a change in ownership from public to private hands. What are the implications of alternative policies vis-à-vis FDI for welfare?

DIAGRAM 3.



Foreign entry into a monopolistic market

Foreign investment clearly brings benefits even in situations where it does not lead to enhanced competition (i.e., there are entry restrictions). Foreign equity may relax a capital constraint, can help ensure that weak domestic firms are bolstered (e.g. via recapitalizing financial institutions), and serve as a vehicle for transferring technology and know-how, including improved management. However, if FDI comes simply because the returns to investment are artificially raised by restrictions on competition, the net returns to the host country may be negative (returns to the investor may exceed the true social productivity of the investment). To some extent the rent appropriation may be prevented by profit taxation or by holding competitive auctions of licenses or equity, but the static and dynamic inefficiencies from lack of competition would still exist.

Why do we observe such widespread restrictions on entry? While it is possible to construct special models of market and/or regulatory failure where entry barriers enhance welfare, restrictions generally aim to protect the incumbent suppliers (not necessarily national) from immediate competition for infant industry type reasons, to facilitate "orderly exit" or simply because of political economy pressures. Monopolistic or oligopolistic rents may also be seen as a means to allow firms to fulfil universal service obligations. Both of these arguments are considered further below. In some cases a form of "investment pessimism" exists, leading to the belief that promises of oligopoly rents are necessary to attract new investment. However, it is not clear why the market structure needs to be determined by

policy, unless there are some initial investments the benefits of which may be appropriated by rivals. Finally, governments may seek to raise revenue (or rents for politicians/bureaucrats) by auctioning monopoly or oligopoly rights. This amounts to indirect appropriation of consumers' surplus. But the static and dynamic inefficiencies consequent upon lack of competition would still exist.

Entry restrictions are becoming harder to justify in the face of growing evidence of the benefits of competition. In Latin America, for example, countries that granted monopoly privileges of six to ten years to the privatized state enterprises saw connections grow at 1.5 times the rate achieved under state monopolies but only half the rate in Chile, where the government retained the right to issue competing licenses at any time. Several other studies find a significant negative relationship between performance (measured by price and quality indicators) and the number of firms and the existence of an independent regulator, and generally a weaker relationship with the share of public and foreign ownership. These results support the view that the consumer benefits arise more from increased competition and effective regulation than from a change in ownership.

Full liberalization in terms of removal of all barriers to entry leads to an increase in social welfare. Partial liberalization, in terms of limited entry or only change of ownership, could lead to a decline in national welfare.

The GATS rules on measures affecting services trade

The GATS rules operate at two levels. First, there are a set of general rules that apply across the board to measures affecting trade in services, and then there are a set of sector-specific commitments that determine the extent of liberalisation undertaken by individual countries.

The specific commitments on market access and national treatment are the core of the GATS, and the impact of the Agreement depends to a large extent on the commitments made by Members.[2] Article XVI stipulates that measures restrictive of market access which a WTO Member cannot maintain or adopt, unless specified in its schedule, include limitations on :

(a) the number of service suppliers;
(b) the total value of services transactions or assets;
(c) the total number of services operations or the total quantity of service output;
(d) the total number of natural persons that may be employed in a particular sector;
(e) specific types of legal entity through which a service can be supplied; and
(f) foreign equity participation (e.g. maximum equity participation).

With the exception of (e), the measures covered by Article XVI all take the form of quantitative restrictions.

[2]Both the market access and national treatment provided for in the schedules must be extended to all foreign service suppliers on a non-discriminatory basis, irrespective of whether a country has listed any MFN exemptions.

Three aspects of Article XVI are important. First, the Article XVI list does not include all measures which could restrict market access. Perhaps most significantly, fiscal measures are not covered. Thus, a Member could maintain, without being obliged to schedule, a high non-discriminatory tax on a particular service which severely limits market access. Secondly, Article XVI has been interpreted to cover both discriminatory and non-discriminatory measures, i.e. measures of the type "only five new *foreign* banks will be granted licenses" and also measures such as "only ten new [*foreign and domestic*] banks will be granted licenses". Finally, the limitations must be read as "minimum guarantees" rather than "maximum quotas", i.e. a country which has promised to allow five foreign banks entry is free to grant entry to more than five. We shall return to these issues.

Article XVII:1 states the basic national treatment obligation:

> "In the sectors inscribed in its Schedule, and subject to any conditions and qualifications set out therein, each Member shall accord to services and service suppliers of any other Member, in respect of all measures affecting the supply of services, treatment no less favourable than that it accords to its own like services and service suppliers."

Unlike Article XVI, Article XVII provides no exhaustive list of measures inconsistent with national treatment. Nevertheless, Article XVII:2 makes it clear that limitations on national treatment cover cases of both *de jure* and *de facto* discrimination.

The Explanatory Note provides two examples of limitations on national treatment. If domestic suppliers of audiovisual services are given preference in the allocation of frequencies for transmission within the national territory, such a measure discriminates explicitly on the basis of origin of the service supplier and thus constitutes formal or *de jure* denial of national treatment. Alternatively, consider a measure stipulating that prior residency is required for the issuing of a license to supply a service. Although the measure does not formally distinguish service suppliers on the basis of national origin, it *de facto* offers less favourable treatment for foreign suppliers because they are less likely to be able to meet a prior residency requirement than like service suppliers of national origin.

A Member's specific commitments can be seen as the outcome of a two-step decision. Each Member first decides which service sectors will be subject to the GATS market access and national treatment disciplines. It then decides what measures will be kept in place for that sector which violate market access and/or national treatment respectively. Commitments on both market access and national treatment have been specified by modes of supply.[3]

[3]The GATS schedules of commitments are structured in the following manner. In the left hand column of the table are inscribed the service activities which are the subject of specific commitments. For each of the four modes of supply noted in column two, columns three and four state whether there are limitations on market access and national treatment respectively. The extreme right hand column provides for the additional commitments on other measures affecting trade in services.

Entries in the schedule in a given sector with respect to a particular mode of supply fall into one of four categories. (i) *Full commitment*: "none" or "no limitations", which implies that the Member does not seek in any way to limit market access or national treatment through measures inconsistent with Articles XVI or XVII. (ii) *Commitment with limitations*: the Member describes in detail the measures maintained which are inconsistent with market access or national treatment, and implicitly commits itself to take no other inconsistent measures. (iii) *No commitment*: "unbound" indicates that the Member remains free to maintain or introduce measures inconsistent with market access or national treatment. (iv) *No commitment technically feasible*: "unbound*" indicates that in the sector in question, a particular mode of supply cannot be used, for instance cross-border supply of hair-dressing services.

GATS rules do not create an hierarchy of instruments like in the GATT. Countries have often conceded increased "market access" under pressure from trading partners in the form of limited new entry or increased foreign ownership of existing domestic firms, rather than by allowing new entry. It is important to design trade rules to create a presumption against welfare-reducing restrictions on entry.

b. Precommitment to future liberalization

One reason governments may be reluctant to liberalize immediately is a perceived need to protect the incumbent suppliers from competition—either because of infant industry type arguments or to facilitate "orderly exit." One reason for the failure of infant industry policies in the past, and the innumerable examples of perpetual infancy, was the inability of a government to commit itself credibly to liberalize at some future date. The GATS offers a valuable mechanism to overcome the credibility difficulty. Governments can make binding commitments to provide market access and national treatment at a future date. Failure to honour these commitments would create an obligation to compensate those who are deprived of benefits, making the commitment more credible than a mere announcement of liberalizing intent in the national context. A precommitment to liberalize can also instill a sense of urgency in domestic reform, and in efforts to develop the necessary regulatory and supervision mechanisms.

Several governments have taken advantage of the GATS to strike a balance between their reluctance to unleash competition immediately on protected national suppliers and their desire not to be held hostage in perpetuity either to the weakness of domestic industry or to pressure from vested interests. The most striking examples are in basic telecommunications, where a number of developing countries have bound themselves to introduce competition at precise future dates. The use of the GATS as a mechanism for lending credibility to liberalization programmes has been disappointing in other sectors.

IV. Discrimination in services trade, MFN and the GATS

Article II on most-favoured-nation (MFN) treatment requires each Member to "accord immediately and unconditionally to services and service suppliers of any other Member treatment no less favourable than that it accords to like services and service suppliers of any other country". We explore the implications of this principle and departures from it.

In principle, the MFN obligation precludes discrimination between trading partners. But it proved difficult to make it applicable without exception. Two particular challenges arose: the desire to maintain or conclude new preferential arrangements – sometimes in the form of regional integration agreements; and the desire to maintain negotiating leverage.

The consequences of discrimination through taxation (or duties) are well understood. Does discrimination in services through domestic regulations and other non-tariff instruments raises new analytical issues from the legal and economic point of view?

Economic implications

To begin with, consider domestic regulations pertaining to technical standards and qualification requirements. The requirement to meet a certain standard may involve a change in variable costs, fixed costs, or both. The requirement for financial institutions to maintain local reserves is an example of a standard that may increase variable costs of operation. The requirement for foreign professionals to requalify increases the fixed costs of entry for each individual, but if we define units so that each professional is assumed to provide a single unit of service, then even qualification requirements can be seen as affecting variable costs. *The problem is that it would not usually be correct to treat all the additional costs imposed on foreign services or service suppliers as a form of protection. It is necessary to distinguish between the regulatory burden imposed on the foreign supplier which is necessary to ensure the desired quality of the service and that which is excessive.*

This issue can be illustrated by considering the qualification requirements imposed on foreign professionals. Even though professionals obtain composite degrees and qualifications in each country, it should usually be possible to distinguish between a basic "universal" part of the training, consisting of u units, which is identical between countries, and country specific training (say in local law) equal to v_i units. In some professions, like medicine and engineering, the universal component is likely to be high, whereas in other professions, like law and accountancy, the country-specific component is likely to be high. Outside of professional services, for instance in construction, financial and transport services, it is reasonable to presume that there is a high universal component to standards though there is usually also a variable country specific elements.

Let us also assume that the cost per unit of training in country i is constant at c_i. The variations in c_i are meant to capture inherent advantages that certain countries have in certain areas. It would seem obvious that if a foreign lawyer wished to practise in country i, he would necessarily have to incur an entry cost equal to $c_i v_i$. But it is possible that the first country also refuses to acknowledge the equivalence of the universal part acquired in the home country, and insists on full requalification, implying entry costs $c_i(u + v_i)$. In this case, $c_i u$ would be a measure of the excessive regulatory burden. More subtle forms of protection could involve understatement of the universal element u, and exaggeration of the country specific element v_i.

Recognition agreements may exempt certain suppliers from incurring whole or part of these costs.[4] In situations where the impact of regulation is on variable costs, as is assumed here, the analysis of discriminatory regulation can proceed in a manner analogous to tariffs.[5] Assume that there are three countries: country X is an importing country, facing an upward sloping domestic supply of professionals (say because of increasing opportunity costs) and countries Y and Z are potential exporters. Let us say that $c_x > c_y > c_z$, and that in the absence of any trade the prevailing price in country X is p*. Assume now that country X recognizes the equivalence of the universal component of training obtained in country Y, but not in country Z. There are several possibilities, but we consider only two:

No trade prior to recognition: Prior to recognition, professionals from both countries Y and Z were required to obtain country-specific training in country 1 but neither found it worthwhile to do so, i.e. $c_y(u + v_y) + c_x(u + v_x) > p^*$ and $c_z(u + v_z) + c_x(u + v_x) > p^*$.[6] But $c_y(u + v_y) + c_x v_x < p^*$, i.e. when the universal component of training in country Y is recognized as equivalent to that in country X, then professionals from Y find it worthwhile to export to country X, and the price in X will fall. Hence, if all foreign professionals had been completely deterred from practising in country X by the absence of recognition, then any recognition agreement is necessarily trade creating.

Country Z exports to country X prior to any recognition: In this case, $c_z(u + v_z) + c_x(u + v_x) < p^* < c_y(u + v_y) + c_x(u + v_x)$, i.e. when all foreign professionals were required to requalify, only those from the third country were willing to supply country 1. But $c_y(u + v_y) + c_x v_x < c_z(u + v_z) + c_x(u + v_x)$, i.e. once the second country professionals are exempted from the basic qualification requirement, they gain a competitive edge over the third country. That is, if lawyers from the third country were already present in the first country, then the recognition of second country professionals would put them at a competitive disadvantage and could lead to trade diversion.

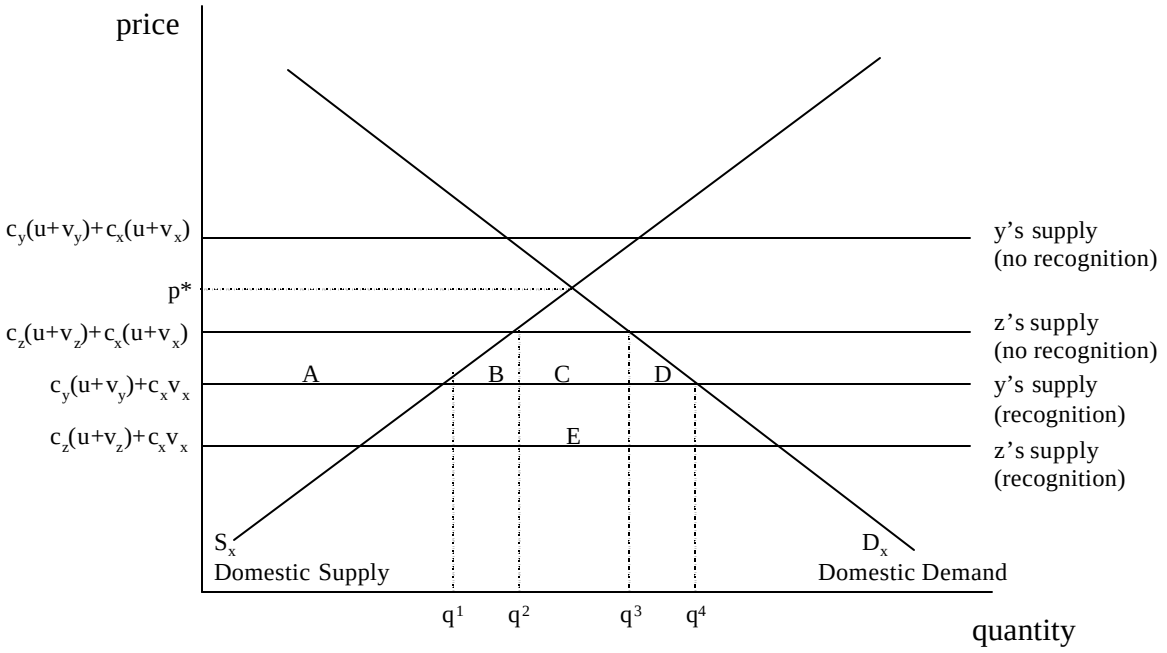
Diagram 4

---

[4] For instance, owing to the reciprocal recognition of the proof of solvency between the EU and Switzerland, foreign direct (other than life) insurance companies which have their principal place of business in the territory of one of the contracting parties are not obliged to localise funds to a significant extent. The United States agreement with Canada eliminates the need for chartered accountants trained in these countries to duplicate all steps in the licensing process, and provides for abbreviated examination requirements.

[5] Where the impact of the regulation is on fixed costs, the effects are likely to be seen in terms of entry and exit of suppliers from different countries depending on the regulatory burden.

[6] It may be asked why some individuals from countries Y and Z do not directly qualify as professionals in country X instead of first obtaining qualifications at home. One reason is that professionals are often allowed to enter foreign markets only on a temporary basis (as under the GATS), so they need to be qualified also to serve the home market. Where longer term movement of professionals is allowed, we need to assume that the individual elements of qualifications are not separable, and that there is a part, say w, which is universally recognized as equivalent. The incentive to obtain the qualification at home arises as long as $c_y w$ and $c_z w$ are sufficiently smaller than $c_x w$. The w element is suppressed here to keep the notation simple. Non-separability is indeed an aspect of many qualifications: a student can usually not switch institutions after doing only part of a course.



This situation is depicted in the Diagram. The pre-recognition situation involves domestic output q_2, consumption q_3, and imports from Z, $q_3 - q_2$. After country X recognises qualifications in country Y, domestic output declines to q_1, consumption increases to q_4, and imports from Y, $q_4 - q_1$, displace imports from Z. So we are witnessing both trade creation and trade diversion. The welfare effects are straightforward. Consumer surplus rises by A + B + C + D as consumption expands from q_3 to q_4. The area A is a gain at the expense of domestic suppliers, whose surplus falls with their output. Area D is the gain from the better allocation of consumption expenditures; area B the gain from the resources released as inefficient domestic supply contracts; and area C the gain arising from the elimination of wasteful requalification.

The area C + E is of crucial significance, and can be interpreted in several different ways. It helps to recall the analysis of preferential arrangements when tariffs are the instruments of protection. In that case, area C + E would be the loss in government revenue because preferential imports displace high tariff imports. While C is gained by consumers, E is completely lost because supply comes from the more expensive source, and is the loss due to trade diversion. The net gain to the country is only B + D – E, and could be positive or negative. In the example here, C + E were the costs of requalification for country Z professionals when they supplied country X. If these costs were completely dissipated, then they do not enter the welfare calculus of country X. That is, *there is no cost of trade diversion for the importing country* and the net gain to the importing country from the recognition agreement is B + C + D. If, however, part of these requalification costs (say a fraction α) were appropriated by country X, perhaps as the producer surplus of its training industry or as some form of regulatory rent, then they would be foregone with trade diversion and would need to be taken into account: the net gain to the importing country from the agreement would be $B + C + D - \alpha(C + E)$.

When tariffs are the instruments of protection, the costs of trade diversion for the importing country may be an important deterrent to preferential liberalization agreements. Despite the increase in consumers' surplus from any liberalization, governments may nevertheless be averse to such agreements because the displacement of high-tariff imports from third countries by low or no-tariff imports from preferential sources implies lost revenue. The same reasoning also applies to other regulations which imply a transfer from foreign suppliers to domestic interest groups. However, the situation is different when the protectionist instrument is a regulatory barrier which imposes a cost on the exporter without yielding a corresponding revenue for the importing government or other interest group. There is then no cost to granting preferential access because there is no revenue to lose. Therefore, preferential liberalization is necessarily welfare enhancing for the importing country.

In this simple model, trade diversion reduces the third country's sales to the first country. The simplest way of depicting the negative welfare effects on the third country is by assuming that its supply curve is upward sloping, so that lower sales imply lower prices and loss of producers' surplus. Interestingly, it is possible to show that recognition agreements may well increase global welfare even though suppliers who are left out of such agreements lose. In effect, a recognition agreement is like a positive cost shock to a certain class of suppliers leading to a reduction in price in the importing country. The gain to consumers from any decline in price is necessarily greater than the loss to a subset of suppliers. This makes intuitive sense: eliminating wasteful duplication of training should enhance global welfare. Though, of course, a non-preferential recognition agreement would enhance welfare even more because the service would be supplied by the most efficient locations.

V. Domestic regulations

The economic case for regulation in all services sectors arises essentially from market failure attributable primarily to three kinds of problems, natural monopoly or oligopoly, asymmetric information, and externalities. The social case for regulation is motivated by distributional concerns. No one denies that there is good reason to regulate. In fact the GATS is at pains to affirm that liberalization does not mean deregulation. Nevertheless, it is true that certain regulations (or their absence) can become impediments to trade. For instance, foreign professionals are sometimes subject to elaborate requalification requirements. How much is justified to ensure service quality, and how much is a disguised form of protection? In telecommunications, the absence of pro-competitive regulation can be a barrier to trade since entrants need to interconnect with existing networks at reasonable rates.

One of the ironies of the GATS is that among its weakest provisions are those dealing with domestic regulations, which have such an obviously powerful influence on international trade in services. The reason is not difficult to see: it is extremely difficult to develop effective multilateral disciplines in this area without seeming to encroach upon national sovereignty and unduly limiting regulatory freedom.

Dealing with domestic regulations at the multilateral and national level			
Market failures	**Services sectors**	**Multilateral approach**	**Action required at national level**
Monopoly/ oligopoly	Network services: transport (terminals and infrastructure), environmental services (sewage) and energy services (distribution networks).	Generalize key disciplines in telecom reference paper to ensure cost-based access to essential facilities, be they roads, rail tracks, terminals, sewers or pipelines.	Develop pro-competitive regulation to protect consumer interests where competitive market structures do not exist.
Asymmetric information	Intermediation and knowledge based services: financial services, professional services, etc.	Non-discrimination and generalization of the "necessity" test. Use the test to create a presumption in favour of economically efficient choice of policy in remedying market failure.	Strengthen domestic regulation to remedy market failure in an economically efficient manner.
Externalities	Transport, tourism, etc.		
Social objectives: Universal service	Transport, telecommunications, financial, education, health		Devise economically efficient means of achieving social objectives in competitive markets.

Market failure due to natural monopoly or oligopoly may create trade problems because incumbents can impede access to markets in the absence of appropriate regulation. Because of its direct impact on trade, this is the only form of market failure that needs to be addressed directly by multilateral disciplines. The relevant GATS provision, Article VIII dealing with monopolies, is limited in scope. As a consequence, in the context of the telecom negotiations, the reference paper with its competition principles was developed in order to ensure that monopolistic suppliers would not undermine market access commitments. These principles should be generalized to a variety of other network services, including transport (terminals and infrastructure), environmental services (sewage) and energy services (distribution networks), by ensuring that any major supplier of essential facilities provides access to all suppliers, national and foreign, at cost-based rates. At the same time, there is a need to strengthen Article IX to deal with international cartels (e.g. in transport services) which cannot be adequately addressed through national competition policy.

In all other cases of market failure, multilateral disciplines do not need to address the problem per se, but rather to ensure that domestic measures to deal with the problem do not serve unduly to restrict trade. (The same is true for measures designed to achieve social objectives.) Such trade-restrictive effects can arise from a variety of technical standards, prudential regulations, and qualification requirements in professional, financial

and numerous other services; as well as from the granting of monopoly rights to complement universal service obligations in services like transport and telecommunications. The trade-inhibiting effect of this entire class of regulations is best disciplined by complementing the national treatment obligation with a generalization of the so-called "necessity" test. This test leaves governments free to deal with economic and social problems provided that any measures taken are not more trade restrictive than necessary to achieve the relevant objective. This test is already part of the recently established disciplines in the accountancy sector. It is desirable to use it to create a presumption in favour of economically efficient choice of policy in remedying market failure and in pursuing non-economic objectives. For instance, in the case of professionals like doctors, a requirement to re-qualify would be judged unnecessary, since the basic problem, inadequate information about whether they possess the required skills, could be remedied by a less burdensome test of competence. In sum, the telecommunications and accountancy models, suitably developed and generalized, can together ensure that domestic regulations achieve their objectives without sacrificing economic efficiency.

The development of multilateral disciplines is in no way a substitute for strengthening domestic regulatory mechanisms and institutions. At least three areas are of considerable importance.

(i) Dealing with monopolies

The telecom Reference Paper illustrates both the strengths and the limitations of the multilateral approach. The primary concern of the paper, as of WTO rules in general, is to ensure effective market access, and hence the focus on the terms of interconnection. Wider concerns about consumer interests and how they may be affected by monopolistic behaviour are not addressed by the Paper. While there can be little doubt that price determination is ideally left to competitive markets, and regulatory price setting is fraught with difficulties, yet regulatory authorities in developing countries where competition is slow to develop need to equip themselves, legally and technically, with the ability to regulate prices.[7] This would seem particularly desirable in countries like some of those in the Caribbean, which have locked themselves into exclusive supply contracts with a single telecom provider well into the next century. Importantly, while nothing in the GATS prevents a country from pursuing any form of pro-competitive regulation provided it is not discriminatory, the capacity of most developing countries to exercise such regulation is limited.

(ii) Dealing with asymmetric information

The need for effective regulation of financial services needs no elaboration, particularly in light of the recent experiences of many countries. Again it is incumbent on the countries themselves to create adequate mechanisms for such regulation. And such regulation is clearly necessary to benefit fully from liberalization. Other areas where

[7] In many developed country markets where fully competitive conditions have not been established, such as the telecommunications sector in the United Kingdom, the final price itself has been regulated.

the inadequacy of regulatory mechanisms to deal with asymmetric information is a problem have received relatively less attention. For instance, in professional services, low standards and disparities in domestic training and examinations can become a major impediment to obtaining foreign recognition. Thus inadequacies in domestic regulation can legitimize external barriers to trade. A further twist is that domestic consumers may actually prefer cheap, low quality products. The question of how best to achieve the needs of export markets given domestic preferences for quality is clearly an area where much more research is needed.

(iii) Achieving universal service and non-economic objectives

Attaining social objective in an economically efficient manner is a major challenge for national policy-makers. The manner in which they pursue this objective can have a profound impact on trade in a variety of areas, ranging from financial, transport, telecommunications, health and education services. Interestingly, the telecom Reference Paper acknowledges the right of a country to define universal service obligations provided they are administered in a transparent, non-discriminatory and not excessively burdensome manner. But it does not prescribe the appropriate means to achieve this objective – this is left to national governments.

Historically, governments frequently relied on public monopolies to pursue (often unsuccessfully) universal services objectives, either through cross-subsidization across different segments of the market, or through transfers from the government or government-controlled banks. In addition to the inefficiencies created by monopolistic market structures, the burdens imposed by these obligations on existing national suppliers are even now a major impediment to liberalization in many countries. For instance, domestic banks saddled with bad debts because of past directed-lending programmes are not well equipped to deal with foreign competition.

Nevertheless, the current handicap of universal service obligations can in principle also be imposed on new entrants. Thus, such obligations were part of the license conditions for new entrants into fixed network telephony and transport in several countries. But as in many other cases, recourse to fiscal instruments has proved more successful than direct regulation. For instance, in Chile, government subsidies equivalent to less than 0.5 percent of total telecommunications revenue, allocated through competitive bidding in 1995, mobilized 20 times as much private investment to extend basic telephone services to rural areas.

A third instrument is to fund the consumer rather than the provider. Governments have experimented with various forms of vouchers, from education to energy services. This last instrument has at least three advantages: it can be targeted directly at those who need the service and cannot afford it; it avoids the distortions that arise from artificially low pricing of services to ensure access; and finally, it does not discriminate in any way between providers.

VI. A summing up

(i) *Substantial gains from liberalization of services trade*

(ii) *Successful domestic liberalization requires*

- Emphasis on competition
- Credibility
- Efficient domestic regulations

(iii) *Effective access to foreign markets requires*

- Elimination of explicit restrictions
- Disciplines on implicit regulatory barriers

Use the GATS as a forum for

- meaningful reciprocity-based negotiations
- binding and precommiting to liberalization
- developing economically sensible rules for domestic regulations