4 Trading Blocs and the Incentives to Protect: Implications for Japan and East Asia

Kenneth A. Froot and David B. Yoffie

4.1 Introduction

The Single European Act of 1985, the Canada-U.S. Free Trade Agreement, and the current talks among the United States, Canada, and Mexico may signal the beginning of a sea change in international trade policies. Government decisions to form larger trading blocs have been construed by many as a precursor to the return of trade warfare that dominated international trade in the 1930s. Just as the absence of international leadership in the interwar period produced trade rivalry among nations and intense competition for markets, so too have there been fears that the tripolar world of the 1990s will be too unstable to promote freer trade. In this context, the momentum created by Europe's 1992 program and North America's free trade zone could have severe consequences for Japan and East Asia.

There is a political as well as an economic logic for believing that the 1990s will produce fortresses in Europe and North America. On the political front, the declining competitiveness of the United States and many European nations has produced increasing pressure to protect industry and employment. In the absence of a hegemon, the countervailing forces for free trade may be difficult to find. These political arguments become reinforced by the economics of bloc formation. Medium and small countries have never made a pretense of becoming self-sufficient; it makes little sense for Luxembourg or Canada to produce

many industries. Therefore, it is often argued that, as trading blocs grow larger, the potential benefits of protectionism rise.

This paper explores the underlying incentives for protectionism in a world with economic blocs, and the consequences for trade, investment, and competition with Japan and East Asia. We argue that the incentives for protection vary by the type of industry: in traditional sectors not characterized by increasing returns, protectionism may not be appealing in large trading blocs. We suggest that, over time, trade barriers may even fall. Our logic is that factors of production, especially capital, are increasingly mobile. In industries without strong increasing returns, efficient production is possible in a variety of locations. As a consequence, any effort by one region to raise taxes will lead firms to establish operations abroad. As long as barriers to exit for capital are low, firms will seek to exploit lower-cost locations that take advantage of other large markets. Exit by domestic firms (as well as the possibility of inward investment by foreign firms) weakens the political case and the political coalition for protection in the long run. Much of the outward investment should benefit low wage but high productivity countries in East Asia, such as Korea, Taiwan, Singapore, Malaysia, and so forth.

High-technology industries, with increasing returns, offer a different picture. We argue that the emergence of regional trading blocs is more likely to produce an increase in trade restrictions in sectors characterized by large fixed costs in R&D, manufacturing scale economies, and/or steep learning curves. For these industries, there is not only a coherent case for import protection (to help promote exports as well as preserve the domestic market), but that case becomes even stronger as trading blocs emerge and grow.

Finally, we draw some implications from the model. One of the strongest findings is that there is a disjuncture between trade policies and investment policies, in both academic models and the real world. The mobility of capital undermines many of our precious assumptions about how trade policy and trade politics are supposed to work. Moreover, if our arguments are robust, they suggest some significant dangers for the future. Cross-investments between North America and Europe will undermine many of the forces for isolationism and protectionism in the long run only if cross-investment is symmetrical, that is, firms from all major regions invest in each others' territory. To date, however, cross-investment has been asymmetrical: European and Japanese firms have invested heavily in America; American and Japanese firms have invested heavily in Europe; but no significant American and European investment has gone to Japan.

If Japan remains reserved for the Japanese, incentives for strategic trade policy will continue in Japan, creating further trade tensions. Moreover, asymmetric access to Japan's markets could alter the structure of competition in the rest of Asia. Using the logic of our model, there are incentives for Japan to create a de facto trade bloc, at least for increasing returns sectors, that extends beyond

the borders of Japan to the rest of East Asia. If European and North American firms lack (or even believe that they lack) trade and investment access to a larger East Asian trading bloc, the asymmetry between North American and Japanese firms in increasing returns industries could grow. The rational response for American and European firms might then be to prevent the creation of an even larger East Asian trading bloc by actively countering Japanese expansion in East Asia.

The paper is in four sections. Following this introduction, we construct a simple model of trade in industries that do not exhibit global economies of scale. In section 4.2, we show how static and dynamic forces for and against protectionism can interact. We conclude that the long-run outlook is optimistic for free trade in these sectors, even in a world with one or two trading blocs. Next, we extend the model to "battlefield" sectors like semiconductors, where increasing returns are critical. Here we argue that the temptation to use strategic trade policies will grow along with trading bloc size. Because the critical factors of production are not as mobile in the short run as industries with increasing returns, direct investment will not have the same effects, at least in the medium term. Only over the long run will direct investment produce similar results. We use brief illustrations of trade in semiconductors to illuminate the increasing returns part of the model.

In drawing implications from the model, we present a brief case study of competition in telecommunications equipment in East Asia. The combination of fear and frustration—fear of the long-term competitive strength of Japanese companies, and frustration over lack of access to the Japanese telecommunications market—have led firms such as AT&T to fight aggressively Japanese firms in East Asia to prevent a de facto trading bloc from emerging. Although the evidence is still anecdotal, it could suggest that East Asia—not Japan—may become the next battleground outside of North America and Europe for East-West competition in high-technology products.

4.2 A Model of Protection among Trading Blocs

Here we study how the formation of trading blocs affects strategic trade incentives for goods not subject to increasing returns in production. For these purposes, we adapt a simple model used first by Gros (1987) and developed in more detail by Krugman (1991b).

Imagine that the world is composed of N countries or distinct economic regions. Each of the N regions has its own variety of indigenous good, which is produced locally (and potentially abroad) and which may be sold to other regions. These regions are divided among B trading blocs. We assume that each bloc represents a "common" market within which goods and factors move freely. For simplicity we assume that all B blocs are symmetric, that is, that they each comprise N/B regions.

The consumers of all regions are exactly alike, in that they share the same preferences for goods produced locally as well as those produced in other regions. Again, for simplicity, we assume that their utility is of the form

$$U = \left(\sum_{i=1}^{N} (C_i^{\theta})\right)^{1/\theta}, \tag{1}$$

where C_i is an individual's consumption of region i's good. The symmetry of the model implies that the elasticity of substitution between any two goods is given by $\sigma = 1/(1 - \theta)$, with $0 \leq \theta < 1$. The higher is σ the greater is the substitutability of goods in consumption.

We also assume that, while goods move freely within the confines of each trading bloc, the domestic (intrabloc) market may be protected by levying import tariffs or export taxes. As long as goods are not perfect substitutes in consumption (which would be the case were $\theta = 1$), blocs will favor some type of protection at their common border. In our model, this protection takes the form of an optimal export tax. From the optimal tariff literature, we know that the optimal (ad valorem) export tax is given by

$$\tau^* = \frac{1}{\varepsilon - 1} \tag{2}$$

where ε is the elasticity of the rest of the world's demand for a bloc's exports. We assume that each bloc sets its own external taxes or tariffs in isolation, treating other blocs' tax rates as fixed; that is, we assume that tariffs are set in a Nash bargaining process.

There is a great deal of literature on the desirability of trading blocs—typically called "customs unions" in the parlance of international trade. Much of this literature is concerned with the question of whether such blocs could ever be in participating countries' interests. By eliminating trade restrictions with one set of countries but maintaining restrictions with others, some of the newly created intrabloc trade is welfare-improving because it involves the expansion of efficient producers, but some of the new trade is welfare-reducing because it expands the production of firms that are inefficient by international standards. However, the positive effects of "trade creation" must dominate the negative effects of "trade diversion" as long as the trading bloc as a whole sets its external barriers optimally.[1] Thus, trading blocs naturally emerge in the above model, since bloc formation is in the individual interest of participating countries (although countries would be better off if all external barriers were completely removed).

There are two empirical difficulties with the theoretical proposition that protection improves a customs union's welfare. The first (and easiest to dismiss) is that protection rarely takes the form of an export tax, the type of trade barrier described by equation (2). However, it should be noted that export taxes are used here for convenience only. In our model—as in many general equilibrium models with balanced trade—export taxes are equivalent to import tariffs. (Tariffs and export taxes discourage both imports and exports because both raise domestic prices of goods and factors relative to those on the world market.) This means that all of our results below hold when tariffs are used instead of export taxes.

1. See Kemp and Wan (1972) for a rigorous derivation.

A second, more telling objection to this model concerns the motivation for imposing protection in the first place. Here the assumption is that protection can improve a country's terms of trade: tariffs lower the world price of its imports, and export taxes raise the world price of its exports. Yet in the real world, these kinds of optimal-tariff arguments are often not the motivation behind the erection of trade barriers. Occasionally, countries impose trade barriers to capture gains from increasing returns to scale—the subject of the model in section 4.2.1 (see also Milner and Yoffie 1989). In most cases, however, governments use protection as a means of transferring resources to factors that are inefficient by international standards (Aggarwal, Keohane, and Yoffie 1987). Often these transfers are not the result of some kind of market failure in which there is a wedge between social and private returns that domestic protection altogether removes.

Despite these caveats, we believe that the Gros-Krugman model has relevance to the political economy of protection when economic blocs form. One of the great fears of "fortress Europe," for example, is that some export-oriented producers, like Italian footwear firms, makers of Belgian chocolate, and British banks, would prefer to have a larger Europe to themselves. Whether or not protection for such sectors can be justified on the basis of increasing returns is unclear. But it is clear that much of the 1992 debate is about making the benefits of the customs union available exclusively to local factors (and not to foreign factors). And this is just another way of phrasing the optimal tariff argument.

The absence of globally increasing returns is important in the model of this section because the model implies that protection *hurts* competitiveness, that is, that protection leads to a diminution, not an expansion, of interbloc trade. The way that the domestic bloc benefits from imposing common external trade barriers is to limit its sales on international markets. This need no longer be true if the trade barriers are put in place to protect firms with globally increasing returns to scale. If, for example, the marginal costs of an import-competing firm fall fast enough as output increases, then tariffs may actually enhance international competitiveness: by protecting domestic production and assuring a domestic market base, the domestic producer may end up with lower costs in terms of world prices, so that its exports become more competitive.[2] The

2. For a series of models exhibiting these features, see Krugman (1991a).

implicit guarantee of domestic market share may lower total costs even though import protection tends to raise factor prices. We investigate the implications of increasing returns in section 4.2.1.

Thus far, we have not addressed the issue of whether government taxation of trade creates an incentive to locate production elsewhere. Consider, for example, the case of an import tariff. By raising domestic production of import-competing goods, tariffs tend to siphon factors of production out of export sectors and to drive up their costs. The corresponding erosion in international competitiveness may be offset through relocating production of exports abroad. The argument for producing abroad is even more direct in the case of an explicit export tax. Either way, protection may lead firms to reconsider their production-location decisions.

However, firm location decisions are usually moot in standard models of trade policy. Domestic "firms" produce local goods only locally. Firms do not consider relocation, and as a result, governments do not need to take relocation decisions into account in determining desired trade barriers. These assumptions are probably not very accurate—few Fortune 500 firms produce exclusively in the United States, and many have more than half of their labor force employed outside that country. Economists usually do not worry so much about the accuracy of the domestic-production assumption, but here we might expect firm location decisions to interact with the level of protection, especially when the world becomes dominated by a few trading blocs.

When blocs levy taxes on their own exports, exporting firms may find that they have a greater incentive to locate certain activities abroad. To be more precise, note that, on units to be sold abroad, each unit that is also produced abroad does not have to pay the export tax. Firms therefore receive $(1 + \tau)$ times as much on foreign sales produced abroad as they do on exported sales, where τ is the ad valorem export tax. As a consequence, domestic firms have an incentive to shift toward foreign-based production as long as the marginal cost of production abroad is less than $(1 + \tau)$ times as great as the marginal cost of domestic production. This leads to an equilibrium condition:

$$\frac{MC_a(q_a)}{MC_h(q_h)} = 1 + \tau, \tag{3}$$

where MC_a and MC_h are marginal cost of production abroad and at home, and q_a and q_h are the quantities produced abroad and at home, respectively. We assume that q_a is less than the quantity consumed by foreign residents, that is, that some domestic production for export always takes place. As long as we have such an interior solution, equation (3) holds, and it yields a condition on the share of firm production that is done overseas.

Equation (3) holds only for "interior" levels of production. That is, it may be that marginal costs of production at home are much lower than those abroad, in which case all production will take place domestically, and equation (3) will not be satisfied. It could also be that there are fixed costs to starting up a foreign

productive facility, so that even if marginal costs make overseas production attractive, firms may not produce abroad. Nevertheless, in what follows we presume that equation (3) holds, that we are not at a "corner" equilibrium.

Equation (3) should be thought of as applying to each productive location abroad. To simplify matters, suppose that firms treat production within each foreign bloc as an alternative to exports to that bloc. (This rules out locating in one foreign bloc as a means of exporting to others, which in any case would be economically inefficient in the equilibrium of the model. Because such foreign-produced exports would be subject to the foreign bloc's export tax, there would be no incentive for firms to locate export production overseas in the first place.)

In order for equation (3) to be operationally useful, we need to make some assumptions about how the ratio of marginal costs on the left-hand side behaves. In order to satisfy equation (3), the left-hand side must be locally increasing in q_a or decreasing in q_h; we cannot have an equilibrium in location of production if by exporting one unit fewer and producing it abroad, a company could lower its total costs of production and increase its incentive to produce even more units abroad. Perhaps the simplest assumption is that marginal costs of home production are a constant (represented by $\alpha > 0$), and that foreign marginal costs increase above the home level as output rises (represented by $\alpha + \beta q_a$, with $\beta > 0$). Equation (3) then has the form

$$\frac{MC_a(q_a)}{MC_h(q_h)} = \frac{\alpha + \beta q_a}{\alpha} = 1 + \tau, \tag{4}$$

which implies that production in each of the $(B - 1)$ foreign blocs is $q_a = \tau/b$, where $b = \beta/\alpha$. It follows that total *production abroad* is given by $Q_a = (B - 1)q_a = (B - 1)\tau/b$.

This expression for production abroad is useful in several ways. First, it assumes that the marginal cost of production at home is always less than that abroad. This means that firms locate abroad only to avoid domestic taxation (relaxing this assumption is likely to strengthen the results below), and otherwise have a preference for domestic production. Second, by letting the parameter b vary with the number of blocs, we have an easy way of incorporating scale effects into the model, even while retaining local decreasing returns. For example, when the world trading system is fragmented into many blocs of small size, blocs may be too small to merit firms establishing separate operations in each. In such a case, we might expect relatively little (and perhaps no) production in each small foreign bloc (i.e., b is large). Alternatively, when there are few blocs, each of large size, marginal costs for large foreign operations might be expected to be close to those for home production (i.e., b is small).

To incorporate this latter notion simply, we let $b = B - 1$, so that overseas production of each region's product is simply $Q_a = \tau$. (If the incentives to produce abroad are stronger as bloc size grows, then we could let $b = (B - 1)^2$,

so that overseas production is increasing in the size of blocs, $Q_a = \tau/(B - 1)$.) Since there are N/B regions in each bloc, the total amount of a given bloc's product that is produced locally in the rest of the world is

$$L^{row} = \frac{NQ_a}{B} = \frac{\tau N}{B}. \tag{5}$$

Next we need to determine the optimal tax for each bloc. Following Krugman (1991b), we normalize each region's volume of output to equal 1. This implies that a representative bloc's output is $Y = N/B$ and that output in the rest of the world is $Y^{row} = N(1 - B^{-1})$. If trade is balanced, then rest-of-world demand must equal rest-of-world output. Rest-of-world demand is spent on goods produced domestically, D^{row}, goods exported from our bloc, M^{row}, and overseas production of our goods, L^{row}. Therefore,

$$Y^{row} = D^{row} + p(M^{row} + L^{row}), \tag{6}$$

where p is the relative price (in rest-of-world prices) of goods from our bloc.

In this setting, unlike in standard models, the "optimal" export tax is, in a sense, a question of political economy. Usually it is assumed that both production and ownership of the domestic firm are entirely domestic. This leads to the presumption that an improvement in the terms of trade will be reaped only by domestic residents, workers, and capital providers. Even if this presumption is not realistic, it is consistent with the structure of the traditional model.

Once the foreign firm employs foreign factors of production, however, it is no longer immediately clear that domestic factors will receive all of the benefits of protection. There is likely to be some leakage to foreign factors. That is, these foreign factors may be able to extract some of the benefits of the domestic good's higher price on world markets. Clearly, domestic residents cannot benefit from a tax or tariff to the extent that its proceeds are transferred to foreigners. And if the government is concerned only with domestic residents' welfare, then leakage to foreigners will affect its choice of an optimal tax or tariff. In the present model, the portion of production that is located abroad avoids the export tax. Therefore the firm, not the domestic government, must distribute some of the tax revenues, both in the form of higher marginal costs and in the form of profits.

To keep things simple, we consider two types of revenue distribution by the firm. The first is the traditional case in which all tax-generated revenues (i.e., the additional firm revenues earned by moving production abroad) are returned to *domestic* factors. This assumption is probably not very reasonable in a world in which firms have international work forces and equity holders. But it is useful because it parallels the assumption in the standard model that all revenue gains accrue to domestic residents. This case might also be thought of as a kind of "short-run" optimal tax. When first moving abroad, domestic firms may be able to keep most of the excess revenues for domestic residents. But over time, as the firm becomes more international in character, foreign factors

of production may become more able to extract excess revenues from their employers.[3] Thus, our short-run optimal tariff treats the leakage to foreigners as unimportant.

The other case—which we will call the "long-run" optimal tariff—is where the government considers only those revenues that are actually collected at the border as benefiting domestic residents. This would occur if the lost tax revenues accrue entirely to foreigners, which as mentioned above is more likely to occur over time. The distinction we are drawing here between short- and long-run is obviously extreme; neither is very realistic. But our goal is to strike a balance between positive and normative theories of commercial policy. Thus, while governments' actual commercial-policy objectives may remain unclear, the optimal tariff may nevertheless be changing over time, as the benefits of protection are increasingly lost to foreign factors of production.

Once we accept this distinction between short- and long-run taxes, it is straightforward to derive their optimal levels. Since in the short run we assume that governments ignore the distinction between L^{row} and M^{row}, we can simply take logs and then derivatives of the terms in equation (6):

$$(1 - f)\hat{D}^{row} + f(\hat{p} + \hat{F}^{row}) = \hat{Y} = 0, \tag{7}$$

where $F^{row} = L^{row} + M^{row}$ is total sales to foreigners, circumflexes over the variables denote log derivatives, $\hat{D}^{row} = d\ln(D^{row}) = dD^{row}/D^{row}$, and $f = (L^{row} + M^{row})/Y^{row}$ is the share of our bloc's goods in rest-of-world consumption. Equation (7) tells us that the elasticity of foreign demand for our bloc's goods is

$$\frac{\hat{F}^{row}}{\hat{p}^{row}} = -\Big(f + (1 - f)\sigma\Big). \tag{8}$$

Using equations (8) and (2), the optimal short-run tax is given by

$$\tau^{sr} = \frac{1}{(1 - f)(\sigma - 1)}. \tag{9}$$

Equation (9) says that the optimal tariff is a function of the substitutability of domestic and foreign goods, and of the share of domestic goods in rest-of-world expenditure. The more substitutable are the goods (the higher is σ), the less there is room to extract rents, and the lower is the optimal tariff. Also, the tariff becomes smaller as share of domestic goods in foreigners' consumption falls. Note, however, that even if the domestic bloc is "small" (i.e., if $f = 0$), the optimal tariff is positive: there is still some monopoly power created by the imperfect substitutability among goods.

In the longer run, the government's perceived elasticity of substitution between domestic and foreign goods is not given by equation (9). The govern-

3. Porter (1990) suggests that workers ultimately are able to extract compensation gains from successful companies. That is, even if a firm can succeed in raising the price at which its product sells, over time it may not be able to raise its markup over costs.

ment recognizes that a tax increase stimulates additional overseas production, eroding the export-tax base. Thus equation (6) becomes

$$(1 - f)\hat{D}^{row} + f(\hat{p} + l\hat{L}^{row} + (1 - l)\hat{M}^{row}) = \hat{Y} = 0, \tag{10}$$

where $l = L^{row}/(L^{row} + M^{row})$ is the share of overseas production in rest-of-world consumption of our bloc's goods.

Next we need to know how overseas production is affected by a change in relative prices. First, note that, since $p = 1 + \tau$, it follows that percentage changes in prices and tariffs are related by

$$\hat{\tau} = \frac{(1 + \tau)\hat{p}}{\tau}. \tag{11}$$

Second, from equations (5) and (11) the percentage change in overseas production for a given percentage change in relative prices is given by

$$\hat{L}^{row} = \hat{\tau} = \hat{p}\left(\frac{1 + \tau}{\tau}\right). \tag{12}$$

Combining equations (10) and (12), we have that the long-run elasticity of substitution is

$$\frac{\hat{M}^{row}}{\hat{P}^{row}} = -\left(\frac{f + (1 - f)\sigma + l(1 + \tau)/\tau}{1 - l}\right) = \varepsilon^{lr}. \tag{13}$$

A little algebra yields that the optimal long-run tax is given by

$$\tau^{lr} = \frac{1 - l}{(1 - f)(\sigma - 1) + 2l}. \tag{14}$$

Equation (14) is similar to (9), except that (14) is a decreasing function of l. This says that, as the foreign-produced share of the domestic good rises, the optimal long-run tariff falls. If l reaches one, so that all of the domestic good is produced abroad, the optimal tariff falls to zero.

In order to understand how these taxes move in equilibrium, we must first determine the consumption and production shares, f and l. Following Krugman (1991b) we note that, at world prices, a representative bloc's expenditure must equal its output,

$$D + M + L = Y, \tag{15}$$

and the representative bloc's output is in turn

$$Y = \frac{N}{B}. \tag{16}$$

Each of the other $(B - 1)$ blocs sells a total volume of $(M + L)/(B - 1)$ (expressed in world prices) to the representative bloc; the ratio of these expenditures to the representative bloc's expenditures on its own good is

$(M + L)/D(B - 1)$. The constant elasticity of substitution (CES) utility function then implies that this ratio is equal to the relative price of foreign to domestic goods, adjusted for the elasticity of substitution, $(M + L)/D(B - 1) = p^{-\sigma}$. Substituting, this yields

$$\frac{M + L}{D} = (1 + \tau)^{-\sigma}(B - 1). \tag{17}$$

Using equation (17) and the definition of f, we have that the share of rest-of-world expenditure that falls on domestic goods is

$$f = \frac{1}{(1 + \tau)^{\sigma} + B - 1}. \tag{18}$$

To determine the foreign-produced share of domestic goods consumed by foreigners, note that l can be written

$$l = \frac{L^{row}/Y^{row}}{(M^{row} + L^{row})/Y^{row}} = \frac{\tau}{f(B - 1)}, \tag{19}$$

where we have used equations (5) and (16) to get the last expression on the right-hand side.

Equations (18) and (19) together with an expression for the optimal tariff (either equation [9] or [14]) allow us to understand how optimal taxes are affected by changes in the number of trading blocs. Let us begin with the short-run tax, τ^{sr}. Here equations (9) and (18) are all that matter (the fraction of output produced abroad, l, has no effect on either equation). Figure 4.1 shows the equilibrium. On the horizontal axis is the level of the tax, τ^{sr}, and on the vertical axis is f, the fraction of the domestic bloc's goods in rest-of-world expenditure. The curve marked TT shows the trade-off between f and τ given by equation (9). The curve is an increasing function of f: as the expenditure

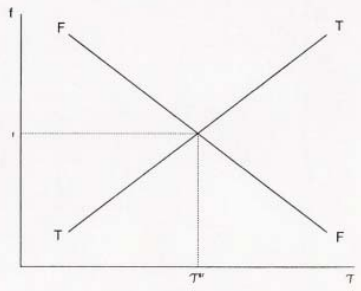

Fig. 4.1 The optimal short-run export tax

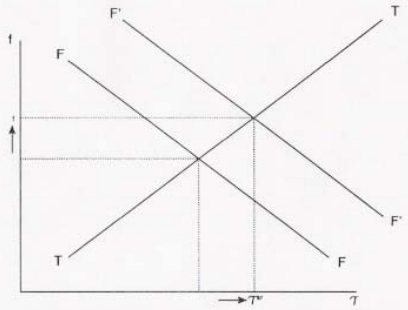

Fig. 4.2 The effect of an increase in bloc size on the optimal short-run export tax

share of domestic goods increases, the domestic bloc's monopoly power also increases. The other curve, marked *FF,* is given by equation (18). It shows that, all else equal, an increase in the tax encourages foreigners to substitute consumption away from domestic products, leading to a decline in *f.*

Figure 4.2 shows what happens to the optimal tariff as the number of blocs falls. The *T T* curve does not shift, since the optimal tax is a function only of a bloc's importance in foreigners' consumption, *f,* and not of the number of blocks, *B.* However, the *FF* curve in equation (18) shifts outward as the number of blocs falls: with fewer blocs, each bloc has a greater share in others' consumption at the preexisting tax rate. In equilibrium, the optimal short-run tax increases to reflect this higher degree of monopoly power. This simple model therefore suggests that protectionism rises as trading blocs become larger.

To see what is driving this result, ask why it is that, for any given number of trading blocs, governments are unwilling to raise taxes to even higher levels. The model's answer is that foreigners shift their consumption away from the domestic bloc's goods, reducing exports. That is, substitution in foreign consumption disciplines a bloc's ability to tax its own industries.

Now let us turn to the long-run tariff. The equilibrium here is described by equations (14), (18), and (19). These three equations are graphed in figure 4.3. The top panel shows the trade-off between *f,* the share of domestic goods in rest-of-world expenditure, and the long-run tax, τ. These two curves are similar to those shown in figures 4.1 and 4.2. In the bottom panel of figure 4.3, the relationship between *l,* the share of sales to the rest-of-world that is produced abroad, and τ is depicted. Note that the *T T* curve here is *downward* sloping: an increase in the share of overseas production reduces the domestic government's tax base, and limits the effectiveness of an export tax. At the margin this makes export taxes less worthwhile. On the other hand, the *l* schedule is

upward sloping: an increase in the domestic tax induces domestic producers to locate more of their production abroad. The optimal long-run tax is determined both by the short-run substitutability of consumption by foreigners and by the long-run substitutability of where production is located.

What happens to the optimal long-run tax as the size of the representative trading bloc increases? To clarify the effect that production-location decisions have on optimal taxes, consider the case in which goods are not very close substitutes, $\sigma = 1$. (This implies that preferences are Cobb-Douglas, so that given *B,* a fixed share of income is spent on each region's good.) Figure 4.4 demonstrates what happens. With $\sigma = 1$, equation (14) becomes

$$\tau^{lr} = \frac{1 - l}{2l}, \tag{14'}$$

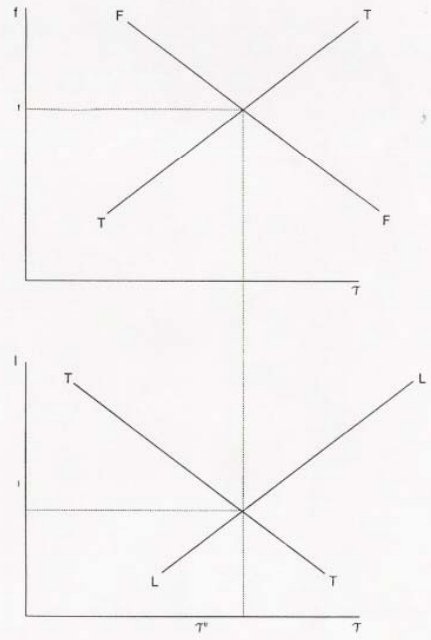

Fig. 4.3 The optimal long-run export tax

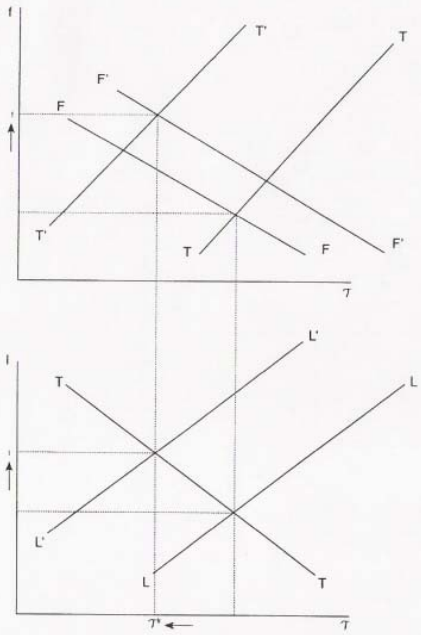

Fig. 4.4 The effect of an increase in bloc size on the optimal long-run export tax

which is a function only of l. By setting $\sigma = 1$, the sole cost of raising export taxes is that production moves abroad. Thus, when the number of trading blocs falls, the TT curve in the bottom panel of figure 4.4 does not shift.

When $\sigma = 1$, equation (18) becomes

$$f = \frac{1}{\tau + B}, \tag{18'}$$

which implies that equation (19) can be written

$$l = \frac{\tau(\tau + B)}{B - 1}. \tag{19'}$$

Equation (19′) says that a decrease in the number of blocs makes firms more willing to establish production abroad, which in turn makes the domestic government less willing to levy export taxes. This effect is captured by a shift outward in the LL curve in the bottom panel of figure 4.4. Thus, as the figure shows, the optimal long-run tax *falls* as the representative bloc increases in size.

Why is it that bigger bloc size implies a lower optimal long-run tax but a higher optimal short-run tax? Recall that the short-run tax increases because, when domestic goods are a larger share of foreign consumption, it becomes possible for the domestic government to extract more monopoly rents from foreign consumers. In the case of the long-run tax, we can for the moment suppress this effect: setting $\sigma = 1$ neutralizes the effect of bloc size on the monopoly power that a bloc's government has. Once we have suppressed the effects of substitution in consumption, the sole long-run effect is that created by substitution in production location. In the long run, firms have a greater incentive to locate abroad when foreign blocs are big, since bigger bloc size permits overseas production on a more efficient scale. The greater is the elasticity of substitution in production location, the lower is the tax that the government is willing to levy.

Of course, if we allow goods to be better substitutes for one another (by setting $\sigma > 1$), then the long-run tax will be determined by both forces: substitutability in consumption as well as substitutability in production. If bigger blocs lead to greater monopoly power in consumption, but to smaller monopoly power in firm-location choice, then the ultimate effect of bigger blocs on protection is ambiguous. However, the long-run tax will consistently be lower than the short-run tax.

This model therefore suggests that there are conflicting forces at work when trading blocs form or increase their size. On the one hand, there is a temptation to protect domestic producers from charging "too low" a price for their exports. This tends to keep external taxes high. On the other hand, the possibility of foreign direct investment helps to minimize how much the government gives in to protectionist temptations. Government interference in the best interest of the country (or of the export industry) is not necessarily in the best interest of each firm. Firms have a private incentive to avoid direct and indirect costs of protection. When firms can respond to this incentive, effective taxation remains low as trading bloc governments compete with one another to attract domestic production.

As always, it is best to think of the results from this model as suggestive. The forces that determine the optimal levels of protection discussed above do not provide a satisfactory description of the motivation behind many commercial policies. In practice, governments that impose protection (as well as industries that lobby for it) often do so in the name of *promoting* international competitiveness, not *discouraging* competitiveness, as standard optimal tariff arguments would have it. Regardless of the motivation, the general equilibrium effects of protection are clear: it raises the relative costs of domestic factors

and therefore makes local production less competitive internationally. This creates the incentive to produce local goods with less expensive foreign factors. Countries that want to keep the cost of domestic factors down and domestic production for exports will have to avoid protectionist policies.

The economics of industries with increasing returns suggests a different rationale for protection, one that may represent a closer parallel to the real world. Through protection, firms may realize lower costs of production and therefore become more internationally competitive. We therefore turn to the effects of tripolarity on protection of increasing returns sectors.

4.3 Increasing Returns and Trading Blocs

Politicians and businessmen have long argued that a protected domestic market enhances international competitiveness. Traditionally, they based their arguments on the "infant industry" notion, which says that domestic market imperfections lower private (but not social) returns in new industries, and that these imperfections are best dealt with through trade restrictions. Among economists, however, the infant industry argument receives little support. While many economists accept the existence of market imperfections (incomplete capital markets, lack of complete appropriability of R&D, externalities in production, etc.), nearly all reject the idea that trade restrictions can be a first-best means of correction.

More recently, strategic trade theory has offered a better rationale for using protection as a means of helping domestic industries. With imperfect competition among firms, protectionist policies can alter foreign competitors' beliefs about the domestic firm's strategic behavior. Sometimes (though not always) it is possible to use government policies—trade restrictions in particular—to tip the equilibrium outcome not only in favor of domestic firms, but also in favor of the domestic economy as a whole. Trade policies may be a device for conveying credibly the future aggressiveness of domestic firms, which in turn may make foreign firms less aggressive.

For the whole economy to benefit, trade restrictions must create sufficient improvements in the efficiency of the productive sector to offset what would otherwise be an increase in the price paid by domestic consumers (as well as any tendency to "crowd out" other types of domestic production). Thus, it is necessary that some kind of economies of scale, either static or dynamic, be present. We show below that larger domestic markets help leverage the effects of increasing returns. That is, a larger domestic market can enhance the domestic government's ability to capitalize on the benefits from import protection. These forces suggest that protectionism should be even greater in these sectors when blocs are large compared to when the national markets are small.

Imagine that there are N firms that share the world market for a product, say RAM chips. The demand for the product is given by

$$p = a - \sum_{i=1}^{N} q_i = a - Q, \tag{20}$$

where p is the price of chips and q_i is the output of the ith firm. Suppose that each firm chooses its output in order to maximize profits, setting marginal revenues equal to marginal costs. This implies the standard equilibrium condition for profit maximization:

$$p\left(1 - \frac{s_i}{\varepsilon}\right) = MC_i, \tag{21}$$

where $s_i = q_i/Q$, the ith firm's share of the total market for chips; $\varepsilon = \frac{pdQ}{Qdp}$, the elasticity of demand for chips; and MC_i is the ith firm's marginal costs of production. From this setup it is straightforward to show that the ith firm's output is given by

$$q_i = (N + 1)^{-1}\left(a - nMC_i + \sum_{j \neq i} MC_j\right). \tag{22}$$

What happens if the domestic market is protected so that only the domestic firm can sell there? If there are no increasing returns (so that marginal costs remain the same once the protection is put in place), then the domestic firm's foreign market share remains the same in the short run. All that changes is its share of the domestic market. Under these circumstances, protection is likely to be bad for the domestic bloc as a whole: the domestic market for chips becomes less competitive, which hurts domestic consumers more than it helps domestic producers.

If there are increasing returns to scale, however, the domestic economy can benefit from protection. Increasing returns may take several forms, including dynamic effects such as learning by doing and the proliferation of new techniques. For our purposes, however, static increasing returns (in the form of decreasing marginal costs) have the same overall impact as these more complex dynamic effects.

Suppose, then, that marginal costs decline as output increases. This implies that, as the domestic market for chips becomes more efficient, domestic firms expand their foreign market share. In this situation protection is much more likely to make the domestic bloc better off. To see the effects on output, take equation (22) as a description of the domestic firm's foreign sales. When the domestic market is protected, the domestic firm's output rises and so its marginal costs fall (i.e., MC_i declines). This has a direct, positive effect on the domestic firm's foreign sales, raising its foreign market share.

There are also several strategic effects of the protectionist policy, which may be even more powerful than the direct effects. First, other firms reduce the absolute amount of their output in foreign markets, in deference to the lower costs achieved by the home firm. To see this in equation (22), note that q_i falls

as MC_j, $j \neq i$, declines. But a foreign output reduction further spills over into *higher foreign* marginal costs, reducing foreign output even further. Finally, as equation (22) shows, higher foreign marginal costs directly raise the domestic firm's output. This then begins the cycle again, further raising domestic marginal costs and output, and lowering further foreign marginal costs and output. Once we arrive at a new equilibrium—at which point equation (22) is satisfied for all N firms—protection of the domestic market will have been translated into a competitive advantage for the domestic firm in its foreign markets as well. The greater the increasing returns, the greater is the spillover effect onto export competitiveness.

Clearly, these strategic effects are important beyond the large size of a domestic bloc. In larger blocs, domestic protectionist policies have a greater impact on the strategic outcomes abroad. Indeed, the domestic welfare consequences of protection depend importantly on how much marginal costs fall. All else equal, a larger domestic market makes it more likely that marginal costs fall substantially, and that the domestic firm gains a large strategic advantage in overseas markets. The greater the decrease in marginal costs, the greater the chance that the domestic bloc as a whole will benefit from the protection. The implication is that larger blocs have greater incentives to initiate strategic protectionism designed to take advantage of increasing returns.

Is the possibility of foreign direct investment likely to reverse this tendency toward greater protectionism, as it did in section 4.2? The answer partly lies in whether firms are willing to transfer abroad that part of the operations that is subject to increasing returns. Activities such as R&D, product development and design, and the actual production of new generation products may have the greatest increasing returns associated with them. Yet these activities may be the least likely candidates to be moved abroad—at least in the short to medium run. Production activities that are not associated with important increasing returns are probably better candidates for overseas production, to which the model of the previous section applies.

The modern semiconductor industry provides a good illustration of how this model might work in reality. Semiconductors is a relatively new industry, begun in 1959 with the invention of the integrated circuit (IC). Initially the industry had relatively low entry costs and only moderate scale economies (Intel Corporation built a state-of-the-art fabrication facility in 1972 for $3.2 million). Even R&D scale was modest: it was common for a few engineers with a good idea to design a new product. Most firms in the United States, Japan, and Europe built their manufacturing fabrication facilities in their home bases, but since transportation costs were insignificant, assembly and test operations were often moved to low-cost-labor locations.

In the mid-1970s, several changes occurred in the economics and technology of the industry. Perhaps most important was that production of chips moved from large scale integration (LSI) to very large scale integration (VLSI). A result of this change was that microelectronics became much more

capital-intensive. Estimates for building a world-class production facility varied, but most analysts concurred that the cost had risen some ten- to twenty-fold from 1975 to 1985. By 1990, every step in the production process became more capital intensive, expensive, and intricate. A high-volume plant cost approximately $400 million and would take almost two years to build and qualify the products for sale. Learning effects were also significant, with costs declining about 30–40 percent for every doubling of production. One estimate suggested that a firm had to achieve 6 percent of the world market (up from 3 percent a decade earlier) from each new plant in order to justify the capital costs. R&D expenses also rose during this period, averaging as much as 15 percent of sales in some years. As product life cycles in the industry shortened on some high-volume products (like DRAMs) from five to three years, the advantage was won by firms that introduced early and had the capacity to fill demand (Yoffie 1988).

These features of the semiconductor industry make it an ideal-typical candidate for strategic trade policy, especially in the context of growing economic blocs. The largest part of demand for semiconductors is in the United States and Japan (approximately 39 percent and 51 percent, respectively, in 1989), with Europe consuming approximately 10 percent. No individual country in Europe had adequate demand to justify new plant capacity. Once Europe becomes a larger bloc, however, the incentives for more semiconductor production are obvious. A European government could hypothetically intervene in its semiconductor industry, reduce imports, and build local scale economies. Europe might then receive a disproportionate share of the benefits from the profits or spillovers generated by the semiconductor industry. And while capital mobility allows firms to move abroad easily in industries that broadly conform with the competitive paradigm, the capital-intensive and especially the R&D-intensive nature of semiconductors makes it much harder for firms to escape from a high-cost national base in the short to medium run, or for firms to invest directly in a foreign market to avoid import tariffs.

The incentive to protect semiconductors in Europe becomes even more compelling if one looks at the history of this industry. In the early 1970s, America dominated production and consumption—controlling over 60 percent of both. To build a competitive industry, the Japanese government explicitly and implicitly restricted foreign entry until the late 1970s. Even though many studies suggested that protectionism led to initially higher costs for Japanese producers, by the end of the period, Japanese firms successfully built scale economies, moved down the learning curve, and had become the lowest-cost producers in the world of certain leading-edge chips. The temptation for any individual country in Europe to replicate the Japanese experience should be low because even Germany and France have tiny markets for chips compared to Japan and the United States. But collectively, Europe's market for chips in the 1990s is only marginally smaller than Japan's market in the early 1970s.

Not only does larger market size increase incentives for protectionism in

semiconductors, but the high fixed cost structure of manufacturing and the scale intensity of R&D make it difficult for firms to adapt to protectionism. Trade conflict between the United States and Japan in semiconductors has been intense since the mid-1980s. In other industries with comparable trade conflict, like TVs in the 1970s and autos in the 1980s, many Japanese firms invested heavily in the United States within a few years. But direct investment in semiconductors has been much slower; most firms in Japan (and the United States) have considered the cost penalties too great to move either the high value-added portion of manufacturing (i.e., wafer fabrication versus assembly and test) or large R&D facilities (many firms have small design centers in other countries where marginal changes are made in the home country designs). Even though protectionism was a reality in Japan throughout the 1970s and became a real threat in America and Europe in the mid-1980s, few plants actually moved overseas. Most companies that have announced their intentions to pursue direct investment will not be opening facilities in other countries until the mid-1990s (see table 4.1). Furthermore, most of the planned facilities are only manufacturing operations, without fully integrated R&D. Most firms continue to do the significant R&D at home and transfer designs to foreign plants.

While we do not yet know how trade, investment, and protectionism in semiconductors will evolve, experience to date is suggestive of several issues posed in our model. First, the Europeans have already showed signs of creating a fortress in semiconductors, even before the 1992 program was complete. Recent changes in antidumping laws (which had previously defined local content in chips as low value-added assembly and test, but now defines local content as "diffusion" or fabrication) have been widely interpreted within the industry as a sign that Europe wants to safeguard European chip demand for European companies. Second, the high cost of direct investment in an industry like semiconductors makes it harder for firms to adjust. The very slow pace of direct investment is evidence of this trend. But third, even if domestic firms do not like to move abroad their increasing returns activities, one should expect capital to move if it becomes a necessity for being competitive. It may take a much longer time, and not all of the increasing returns activities may relocate, but ultimately capital remains mobile. If firms penetrate each others' markets, and assuming that investment is not a perfect substitute for trade, the domestic incentives for protection could decline.

4.4 Implications

Thus far we have argued that strategic trade policies in a world with larger trading blocs will differ greatly across sectors. Incentives for protection will be greater and last longer in industries with increasing returns in production compared to industries that lack significant scale economies. For those goods that are not subject to increasing returns, the formation of large trading blocs may ultimately help lower tariffs, as firms quickly move production abroad.

Table 4.1 Overseas Semiconductor Fabrication Facilities, Memory and Microprocessors

Company	Date	Location	Product	Capacity (per month)
Mitsubishi	1989	Durham, North Carolina	1M DRAM, arrays, MCU	8,500
	1989	Taiwan	1M 4M DRAMS	
		Alsdorf, W. Germany	4M DRAM, MCU, MPU,	22,000
	1989		Arrays	
Sony	1991	San Antonio, Texas	1M SRAM	12,800
	1996	San Antonio, Texas	SRAM	N.A.
	1992	Scotland	1M SRAM	22,000
NMB	1992	New Mexico	4M DRAM	20,000
NEC	1984	Roseville, California	256K DRAM, arrays	27,900
	1987	Livingston, Scotland	1M DRAM	12,000
	1991	Roseville, California	4M DRAM	16,000
	1994	Hillsboro, Oregon	16M DRAM	16,000
	1991	Livingston, Scotland	4M DRAM	10,000
Motorola	1970s	Aizu, Japan	Logic	365,210
		Aizu, Japan	MCU, SRAM, power ICs	304,341
		E. Kilbride, Scotland	MCU, MEM, logic	
		E. Kilbride, Scotland	1M DRAM, SRAM, MPU	N.A.
		E. Kilbride, Scotland	FET, AMPS, LED	N.A.
		Toulouse, France	Bipolar, power trans	12,000
		Seremban, Malaysia	Small signal	N.A.
	1991	Aizu, Japan	Consumer ICs	N.A.
	1992	Sendai, Japan	4M DRAM, MPU, custom	25,000
National	1975	Greenock, Scotland	NMOS, XMOS, bipolar	40,000
		Greenock, Scotland	Logic	N.A.
		Greenock, Scotland	Logic Custom	7,000
		Livingston, Scotland	1M DRAM, 4M DRAM	12,000
		Ha-Emek, Israel	32-Bit MPU	6,400
Intel	1987	Jerusalem, Israel	386 MPU	21,000
	1993	Kildare, Ireland	NA	N.A.
TI	1960s and	England	PWR, discrete	20,000
	1970s	W. Germany	Logic, LIN	15,000
		Hatogaya, Japan	MCU, logic	15,000
		Hatogaya, Japan	N.A.	28,000
		Hatogaya, Japan	N.A.	18,000
		Hiji, Japan	Arrays, logic, linear	50,000
		Hiji, Japan	Arrays, LISP, MPU	20,000
		Hiji, Japan	4M DRAM	7,000
		Mijo, Japan	64K DRAM	20,000

(*continued*)

Table 4.1 (continued)

Company	Date	Location	Product	Capacity (per month)
		Mijo, Japan	256K DRAM, 1M DRAM, 256K SRAM	23,750
		Mijo, Japan	256K DRAM	20,000
	1991	Ibaragi, Japan	16M, 64M DRAM	N.A.
	1990	Italy	4M DRAM	30,000
	1990	W. Germany	Logic	3,000
	1992	Italy	16M DRAM	20,000
	1991	Taiwan	1M DRAM	30,000
	1995	Taiwan	4M 16M DRAM	N.A.

Source: Compiled from *Dataquest Newsletters.*

For those goods that are subject to increasing returns, however, firms may actually become more efficient producers by locating production exclusively within the protected domestic market. Since the presence of increasing returns can lead the domestic economy as a whole to benefit from protection, the incentives to raise trade barriers increase in a world dominated by large trading blocs.

Yet over time, capital remains mobile, even in sectors with increasing returns. And to the extent that foreign direct investment occurs, and as long as it is an imperfect substitute for trade, it should diminish the force of increasing returns–based arguments for domestic protection. Foreign firms with local production (and local employment) will advocate liberalization. Moreover, one of the most important strategic advantages of protection to domestic firms disappears—the guarantee of a large domestic market base on which efficient production can be realized. If foreign producers invest in—and ultimately share—the domestic market, trade protection may not be a fully credible guarantee of market share. Without credibility, many of the strategic advantages to protection are lost. While strategic protection may provide some local employment, it may or may not provide the type of employment (e.g., semiconductor R&D) or spillovers that would be generated by domestically headquartered firms (see Tyson 1992; Porter 1990; Reich 1990).

To reap the strategic advantages associated with increasing returns, governments would need to insulate their economies from foreign direct investment and from foreign trade. Yet many countries (or blocs) actively protect certain sectors from imports but do not discourage foreign direct investment in those sectors. This suggests that either the motivation for protection is different than the assumptions underlying our model, or that trade and investment policies in many countries are not in harmony with one another.

One could draw an optimistic conclusion about the world economy from this disjunction between direct investment and trade. On the one hand, we have argued that growing economic blocs could produce more economic conflict in the short run, but as foreign investment grows in response to protectionism,

countries will have incentives to liberalize trade. Even in increasing returns sectors, the mobility of capital will make it difficult for the European bloc or American bloc to preserve its domestic market for local firms. Over time, multinational companies will invest in each others' markets, undermining the effectiveness of strategic protectionism.

In reality, much of this process is already under way. In traditional sectors, such as autos, significant foreign investment has already taken place. A Honda produced in Ohio is difficult to distinguish from a Honda produced in Japan; one suspects the same will be true in Europe when Japanese firms bring their announced investments on stream. In the absence of restrictions on local investments, it becomes increasingly difficult for governments in Europe or North America to preserve the local market for local companies.[4] The level of cross-investment among industrial markets has reached historic proportions, partly in response to existing protectionism, and partly in anticipation of the short-run protectionism our model suggests.

If the cross-investment described above were symmetrical (i.e., each bloc invested roughly equally in each others' territory), one might predict that protectionism and strategic trade policy in a tripolar world might eventually disappear. Each bloc would have so much of each other's investment that it would be politically difficult to distinguish national origins of firms. While the outflow of investment has been fairly symmetrical across the three blocs (see figures 4.5 and 4.6), however, it has been highly asymmetrical on the inflow side. Japan appears to be the only major industrialized country whose domestic market remains effectively protected from foreign investment as well as trade in some increasing returns sectors. While there are no formal barriers to foreign direct investment into Japan (restrictions were removed in the 1970s), Japan permits far less foreign access to corporate control than it does even to its goods markets.[5] Figures 4.5 and 4.6 show the flows of foreign direct investment out of and into major countries, including Japan. Even in an era in which foreign direct investment around the world has mushroomed, inflows into Japan remain nil.

If increasing returns are important, Japan may be the only country that has pursued policies that are consistent with maximizing domestic welfare (either on purpose or by happenstance). In the presence of increasing returns, these policies also lower the rest of the world's welfare. Nevertheless, viewed in this way, Japanese policies are not hard to understand.

The possibility that Japan follows a coherent strategic trade policy has not

4. If there are going to be political consequences of direct investment, it is important that investment is not a perfect substitute for trade. In the extreme case, where trade and investments are substitutes, there is no reason to believe that the foreign firm will lobby for liberalization. Once established, the multinational might prefer to continue operating behind closed barriers. In reality, however, much of the investment among industrial countries has been to promote incremental sales without displacing all exports from the home country. Therefore, many foreign investors are likely to advocate freer trade.

5. See Froot (1991) for an analysis of Japanese foreign direct investment.

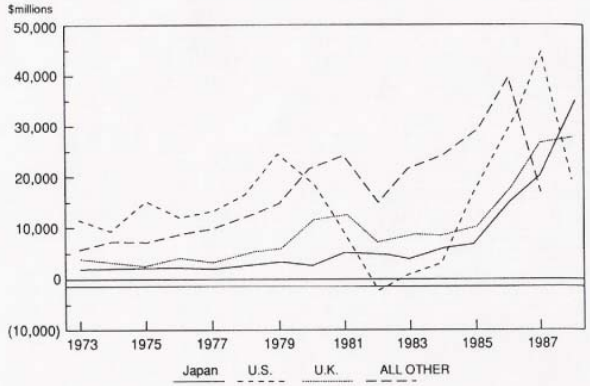

Fig. 4.5 Foreign direct investment outflows
Sources: International Monetary Fund, *Survey of Current Business,* Bank of England, Bank of Japan.

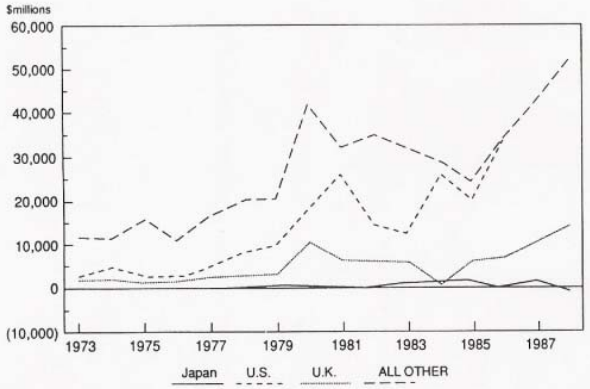

Fig. 4.6 Foreign direct investment inflows
Sources: International Monetary Fund, *Survey of Current Business,* Bank of England, Bank of Japan.

been lost on multinational corporations in increasing returns industries. In businesses like telecommunications equipment, firms based in Europe and America fear a repeat of Japan's performance in semiconductors. The strategy of companies like AT&T may be symptomatic of the coming battles for increasing return sectors in East Asia. Below we briefly describe the dynamics of trade and investment strategies in telecommunications equipment. While the information is still anecdotal, it suggests that concern over a Japanese trading bloc is real and that it is motivating changes in corporate behavior that could influence the structure of competition in the region.[6]

Central-office switches (COSs) form the heart of the public telecommunications network. For most equipment manufacturers, large digital switches are the flagship products of their entire equipment line. Physically, a digital COS consists of arrays of several hundred circuit boards, containing thousands of integrated circuits, wired together in metal cabinets of 400 to 1,000 cubic feet. Digital COSs range in size from 5,000 line units to more than 100,000 lines (connecting remote modules up to forty or fifty miles away). Switches were also highly differentiated products, which could be segmented by differences in size and degree of functionality. Ericsson switches, for instance, were originally designed for international markets and were traditionally more vanilla-like—that is, adaptable but simple, and deployable in relatively small increments. At the other extreme, AT&T switches, such as the 5ESS, were designed for a more advanced network, with larger concentrated volumes of usage, greater functionality (especially centrex), and extraordinary levels of reliability.

The R&D costs for COSs are large, and for the generic software, continuing. Only companies with considerable financial resources and technical personnel, or substantial government support, have entered this business. Fewer still survived the 1980s. In addition to initial development costs, which range from $800 million to more than $1 billion for each manufacturer's switch, there are annual expenses for software modifications of as much as $200 million per firm. To recover costs of this magnitude, most firms had to receive government subsidies (directly through transfers or indirectly through high domestic switch prices), or they had to win a significant share of their domestic markets as well as some share of the world market beyond their national borders. If a firm in a small country, such as GPT in the United Kingdom, could not sell overseas, its ratio of R&D to sales would inevitably become unsustainable without government subsidies, even with a monopoly at home.[7]

Manufacturing COSs also required very large scale plants. Although simple assembly can be done locally on a small scale, firms that manufacture their circuit boards and integrated circuits require large facilities. In the largest national market—the United States—the dominant supplier (AT&T) produces

6. This short case study is based on extensive interviews with AT&T and the other major telecommunications equipment companies worldwide. The full results are reported in Vietor and Yoffie (1993).

7. Ultimately, GPT had to sell out to Siemens (Cowley 1990).

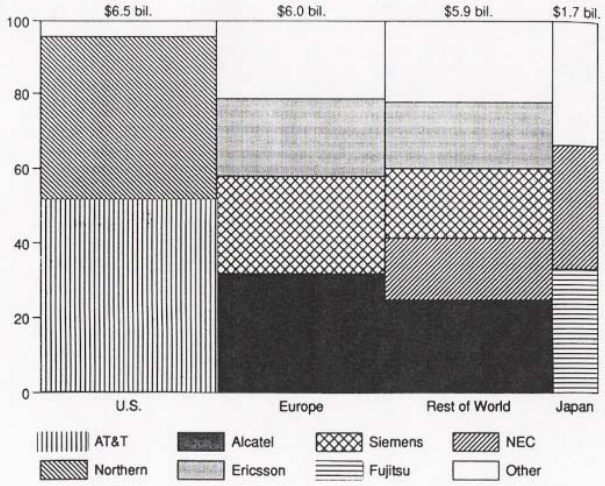

Fig. 4.7 Regional COS market shares
Note: Other in the United States consists of Siemens/Stromberg, Ericsson, and NEC; Fujitsu is also trying to enter. *Other* in Europe consists primarily of ItalTel/AT&T, Bosch, and TeDeWe. *Other* in Japan consists of Hitachi, Oki, and Northern Telecom.

all its 5ESS switches at a single plant and its 4ESS (interexchange) switches at a single plant. Northern Telecom similarly produces its DMS switch at single plants in the United States. Since all other markets are smaller, one can assume that efficient manufacturing scale for digital COSs is a single plant with capacity of as much as 6.5 million lines. Under this assumption, the 1989 world market of 40 million lines could support at most six players, if equally sized, at efficient manufacturing scale. In sum, this is a classic increasing returns business, where the advantages of strategic trade policy should be apparent to all parties.

The strategies of firms in this industry are depicted in figures 4.7 and 4.8. National firms dominate their home markets. In addition, the American firms are aggressively pursuing pieces of Europe, the European and Japanese firms are aggressively pursuing the United States, but most of the non-Japanese players have given up or forgone opportunities to sell in Japan. Real or imagined barriers, technical and political, have discouraged all of the major players (except Northern Telecom, which has less than 1 percent of the Japanese COS market) from selling or investing in the Japanese market.

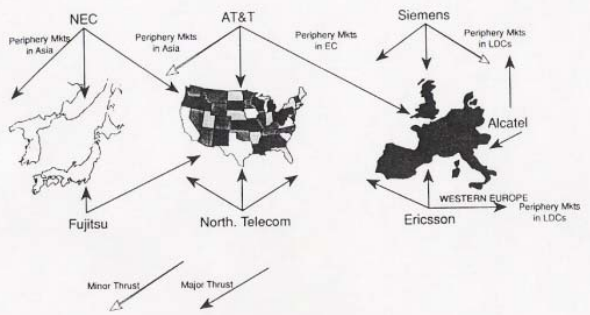

Fig. 4.8 Geographic thrusts in telecommunications

AT&T is particularly interesting in this regard. Starting in the mid-1980s, AT&T pursued a strategy of aggressive expansion into Europe and Japan. Its goal was to generate 25 percent of its revenues from non-US. sales. Recognizing its lack of international experience, AT&T launched a patchwork of "strategic alliances," one with Philips, the Dutch electronics giant, to help AT&T break into Europe's COS market; and a variety of distribution and joint venture manufacturing agreements were signed throughout Asia, including joint ventures in Korea and Taiwan for COSs, and two agreements in Japan—including a distribution arrangement with Toshiba for PBXs.

Japan and Asia were a major focus from the beginning. With NEC and Fujitsu each investing more than $500 million in manufacturing and software development for the U.S. market, AT&T management saw penetrating the Japanese market as a necessity. However, Japan's NTT, after considerable pressure from the United States, rejected the AT&T 5ESS switch. NTT decided to take on only one additional outside supplier, giving Northern Telecom a $650 million contract over five years. Toshiba also decided to develop and market its own PBX directly in competition with AT&T. At this point, AT&T's management concluded that it would be difficult, and perhaps impossible, to challenge Japanese suppliers at home. To counter the possible advantages gained by its Japanese competitors, AT&T concluded that every effort must be made to stop Japan from building a larger protected arena in East Asia. The joint ventures in Taiwan and Korea took on new importance, and AT&T decided to contest aggressively every new contract in Asia.

When Indonesia and Malaysia opened bidding for digital switches in 1990 and 1991, AT&T's responses were indicative of its strategy. Even though it would ship its initial products to Indonesia from its subsidiary in Holland rather than the United States, AT&T politicized the contract to the highest

levels of the American government. It mobilized the U.S. government and especially the U.S. embassy in Jakarta in an effort to offset any influence of the Japanese government on behalf of the other major contender—NEC. A senior Indonesian official claims that "everyone is trying to push, pull, do this, do that" (*Wall Street Journal,* June 21, 1990). Ultimately Indonesia decided to double the size of the contract and give half to NEC and half to AT&T. A similar pattern was being repeated in Malaysia in late 1991: this time it was Fujitsu and AT&T going head to head (*Wall Street Journal,* October 11, 1991).

The logic of AT&T's strategy is consistent with the logic of our paper. With Japanese and European firms investing aggressively in AT&T's home market, the possibilities of strategic trade intervention in COSs becomes increasingly problematic for the United States. While AT&T has been able to counter the challenge of a European bloc, to a limited extent, by winning big contracts in Italy and Spain, it failed at penetrating Japan. AT&T's solution has been to attack the periphery of Japan to reduce the long-run Japanese advantage. The other major non-Japanese firms—Siemens, Northern Telecom, Alcatel, and Ericsson—have followed variations on this strategy.

This analysis suggests two conclusions. First, if firms in other increasing returns industries see Japan as a closed market, then East Asia may well become the next battleground for trade and investment. Second, we have to wonder about the stability of these arrangements when Japan seems to be in a position to follow coherent strategic policies, while the other blocs are politically confounded by cross-investment. Even with growing trade and investment by European and American firms in East Asia in increasing returns sectors, the ongoing asymmetry in direct investment between Japan and the rest of the world makes the outlook uncomfortable, at best.

References

Aggarwal, Vinod, Robert O. Keohane, and David B. Yoffie. 1987. The Dynamics of Negotiated Protectionism. *American Political Science Review* 81 (2): 345–66.

Cowley, Peter. 1990. Telecommunications. In *Europe, 1992,* ed. Gary Hufbauer. Washington DC: Brookings Institution.

Froot, Kenneth A. 1991. Japanese Foreign Direct Investment. In *U.S.-Japan Economic Forum,* ed. M. Feldstein. Chicago: University of Chicago Press.

Gros, Daniel. 1987. A Note on the Optimal Tariff, Retaliation, and the Welfare Loss from Tariff Wars in a Model with Intra-industry Trade. *Journal of International Economics* 23: 357–67.

Kemp, Murray, and Henry Wan. 1972. The Gains from Free Trade. *International Economic Review* 13: 509–22.

Krugman, Paul R. 1991a. Import Protection as Export Promotion: International Competition in the Presence of Oligopoly and Economies of Scale. In *Rethinking International Trade,* Cambridge: MIT Press.

———. 1991b. Is Bilateralism Bad? In *International Trade and Trade Policy,* ed. E. Helpman and A. Razin. Cambridge: MIT Press.

Milner, Helen, and David B. Yoffie. 1989. Between Free Trade and Protectionism: Strategic Trade Policy and a Theory of Corporate Trade Demands. *International Organization* 43 (2): 239–72.

Porter, Michael. 1990. *The Competitive Advantage of Nations.* New York: Free Press.

Reich, Robert. 1990. Who Is Us? *Harvard Business Review* 68 (1): 53–64.

Tyson, Laura. 1992. *Who's Bashing Whom?* Washington, D.C.: Institute for International Economics.

Vietor, Richard, and David B. Yoffie. 1993. Telecommunications: Deregulation and Globalization. In *Beyond Free Trade: Firms, Governments, and Global Competition,* ed. David B. Yoffie, 129–92. Boston: Harvard Business School Press.

Yoffie, David B. 1988. The Global Semiconductor Industry, 1987. Harvard Business School Note no. 9-388-052. Boston: Harvard Business School.

Comment

Marcus Noland

Kenneth Froot and David Yoffie have written an interesting and provocative paper. The paper first constructs a model of trading blocs based on optimal tariff arguments. It is shown that protection will tend to rise with the size of the bloc and that foreign direct investment will act as a counterweight to protection. This raises several questions.

First, Kemp and Wan (1976) have shown that, assuming constant returns to scale but permitting a very general set of impediments to trade, any welfare-maximizing customs union will iteratively expand until it encompasses the whole world in a free trade area in the limit. It is unclear why the Froot-Yoffie bloc wouldn't be subject to the same forces.

One reason that customs unions do not expand forever is that political systems are unable to make the necessary compensating transfers. By the same token, in the real world protection appears to be undertaken more often for narrow parochial reasons, not for optimal tariff reasons. This suggests that blocs could arise as an equilibrium solution in a model of international trade emphasizing transactions costs and the political economy of protection along the lines sketched out in Peter Petri's paper (chap. 1 in this volume). This could be a fruitful avenue for investigating bloc formation.

What this paper does is point to the role of foreign direct investment (FDI) as a counterweight to protection. In the Froot-Yoffie model this occurs because the benefits of optimal trade policies are diluted by the presence of foreign firms. This consideration must enter into policymakers' calculations. However, a number of recent cases (Brazilian infomatics, European autos) indicate that

Marcus Noland is a research fellow at the Institute for International Economics and visiting assistant professor of economics at Johns Hopkins University.

government officials behave in a more sophisticated (or at least in a more complicated) way than in the Froot-Yoffie model, restricting FDI as well as trade as part of an overall program.

Nonetheless, I suspect Froot and Yoffie have the impact of FDI on trade policy right, if for somewhat different reasons. Rather than reducing the incentives to protect because of dilution of benefits in the home country, FDI acts as a counterweight to protection by generating antiprotection lobbying by firms that fear retaliation against their own operations in the foreign target country. In other words, FDI creates hostages, and the existence of hostages discourages rash actions. In either view, Japan's distinctive position with regard to inward FDI is troubling.

While the Kemp-Wan argument holds under constant returns, the existence of increasing returns sectors gives rise to opportunities for rent shifting that are taken up in the second part of the paper. Here I think there is room for a significant extension. The paper assumes that countries can pursue optimal policies without any foreign reaction. I would argue that, as the size of the blocs increases (or, in the paper's terms, as the number of blocs decreases), the likelihood of retaliation increases. This, it seems to me, is central to the bloc story. The United States does not really care if Belgium inflicts its optimal policies on the United States; the United States would care (and certainly respond) if the whole European Community did. One can envision a first pass in which this is treated as a Nash bargaining game.

Permitting foreign response is important because once one allows for retaliation, my sense is that the case for activist policies is weakened, due to the possibility of mutually destructive trade wars. Under these circumstances, it would seem to me that the optimal policy response is twofold: announce a tit-for-tat strategy as a deterrent measure (which is what the United States' section 301 and super 301 can do), then try to negotiate some sort of accord to constrain policymakers' options (which is essentially what GATT does).

This leads me to the final section of the paper, which consists of a discussion of trade in telecommunications equipment. Because GATT does its job so well, the simple tools (tariffs, quotas, subsidies) to implement optimal trade policies are unavailable, policymakers are forced to use less traditional instruments, and there is a well-developed body of anecdotal evidence on this point for Japan. I would thus interpret the Structural Impediments Initiative talks as an attempt to constrain the use of policy tools not already covered under GATT.

However, a second issue arises with regard to Japan. What if the trade impediments do not take the form of governmental policy, but of private preferences. In this case, the natural response would be for foreign firms to form strategic alliances with Japanese firms. For example, the dominant tendency among innovative U.S. producers was not to go to the U.S. government seeking the creation of a bastion market in the United States, but was rather to form strategic alliances with foreign firms.

A question for the authors, then, is whether what is occurring in Japan is more akin to the traditional bloc formation analyzed in the first part of the paper, or the private alliance game described in the final section. If it is the former, then it would be useful to analyze the interests of each East Asian country in joining each potential bloc. Presumably the developing countries have their own interests, and these could conceivably involve siding with one bloc on certain issues and with the other on others. Neither part of the paper really speaks to this issue.

If the contest is more of a private game, however, this opens up an even larger set of issues. We would want to analyze the incentives of firms to join alliances, and what is the proper role of government policy in this situation. This in turn raises issues revolving around the definition of national interest (when, for example, the interests of putatively national firms and national residents diverge, or when the country is home to firms in competing alliances), and the structure of interest-group politics.

Reference

Kemp, Murray, and Henry Wan, Jr. 1976. An elementary proposition concerning the formation of customs unions. In Jagdish N. Bhagwati, ed., *International Trade*. Cambridge: MIT Press.

Comment

Jeffry A. Frieden

This very useful essay could be expanded to cover the political economy of foreign direct investment (FDI) more comprehensively if its approach to FDI itself were more complete. The principal gap in the presentation of the authors is the implicit notion that it is primarily shareholders of a mother firm who benefit from trade barriers that protect one of their foreign affiliates. This is incomplete and leads to a somewhat misleading view of the political economy of trade policy in the context of FDI.

It is widely recognized that FDI is mostly about the transfer of technological, management, and marketing skills and assets that are highly specific to an industry or a firm. In this context, while some of the benefits of increased business for the industry or firm accrue to shareholders, some of them are also realized by workers and managers (or suppliers and customers) whose skills and experience are especially important to the firm or industry.

In fact, industry-specific trade protection is beneficial to *all* factors of production specific to the industry, whether the factors are owned by locals or by foreigners. Workers with skills specific to an industry gain when that industry is protected and can be expected to support such protection. FDI does not

Jeffry A. Frieden is professor of political science at the University of California, Los Angeles.

change this fact; it only changes, if anything, the nationality of the shareholders who also gain from protection.

However, the protectionist rents accruing to an industry can accrue to different *firms* within the industry. In this sense, if a worker has skills specific to a particular *firm* rather than to an industry as a whole, and if protection draws in foreign firms that garner some or all of the rents from protection, the worker might not be better off.

Put differently, adding FDI to a picture of trade policy formation simply adds a new set of *firms* that can benefit from protection. We probably should expect some of the benefits from protection that go to new foreign firms to go to foreigners (shareholders or headquarters personnel of the mother firm), but this is not guaranteed: if the inward investment simply substitutes for lost export markets, the foreigners might in fact be worse off. In any case, even with FDI it is very likely that many of the benefits from protection will go to local people employed by the foreign firm. So those whose factors are specific to the industry as a whole have an incentive to lobby for protection; those whose factors are further specific to an existing domestic firm have an incentive to lobby for protection *and* against inward direct investment. This is simply a variant of the more general desire of members of a cartel to reduce the possibility of new entrants into the cartelized industry.

To take an example, protection of the U.S. auto industry should help those whose skills can be used in the auto industry generally. Inward direct investment in auto manufacture (the "transplant" phenomenon) harms those whose fortunes are tied not to the auto industry generally, but to the Big Three American firms. So auto parts producers and their employees who can sell to *all* assemblers will do well and support both protection and free FDI, while managers and workers with long-term implicit or explicit contracts with domestic auto firms will do less well and will support protection but oppose inward direct investment. I believe that this comes relatively close to describing the political economy of these issues.

The introduction of FDI does not *eliminate* protectionist rents, and probably does not reduce them substantially. It may not even have too appreciable an impact on the rents accruing to local residents (or at least this is purely an empirical question). But it does change the contours of the groups aided by various trade policies in important ways.