VOUSDEN

CHAPTER 3

# Import quotas and tariffs: Some other issues

In the previous chapter we established the result that in a static competitive economy, a tariff and an import quota are equivalent in the sense that a tariff yielding the same domestic price as the quota also yields the same level of imports and output and the same deadweight loss to the community. Given such apparent equivalence, it is reasonable to conclude that it does not matter whether tariffs or quotas are used to protect output or restrict imports.

Certainly there are situations in which the equivalence of the two policies can be used to deduce interesting and relevant results. One such case is the use of trade restrictions to constrain the imports of a class of heterogeneous commodities (e.g. automobiles). This problem was considered in the previous chapter for the case of tariffs. In Section 3.1, we shall exploit the equivalence of tariffs and quotas to consider the effects of an import quota applied to such a class of commodities. Of particular interest is the way in which quotas on the volume of imports shift the composition of imports towards more expensive items.

Despite the usefulness of equivalence in deriving such results, it is now well known that tariffs and quotas are generally not equivalent in most situations beyond the static competitive model. In particular, if there is monopoly in the protected sector or if the policies are imposed by larger countries involved in a trade war, it will matter very much which policy is used. These cases are taken up in Chapters 5 and 4, respectively. Even in the simple case of a small competitive economy, equivalence breaks down in a number of important cases. In Sections 3.2–3.4, we consider these cases and explore a number of issues which relate to the comparison of tariffs and quotas.

In identifying situations where tariffs and quotas are not equivalent, we start with the simplest departure from the basic model: allowing the parameters of the economy to shift over time. Once the demand or supply curves for a good, or its world price, are subject to change, quotas and tariffs affect the market for the good in quite different ways (this case is often referred to as *dynamic non-equivalence* and is considered in Section 3.2). This, in turn, suggests that the two policies must also differ in the presence of uncertainty (where the parameters exhibit fluctuations which are unknown *ex ante*). Indeed, the case of uncertainty raises other interesting issues, in particular the so-called insurance argument for protection. The case of uncertainty is dealt with in Section 3.3.





Finally, in Section 3.4, we incorporate rent seeking and its costs into the basic model. Like most government policies, tariffs and quotas create rents which are the subject of competition among economic agents. Rent seekers use economic resources to lobby for a share of the revenue from a tariff or for the rents embodied in the scarce import licences under a quota. Such use of resources is largely wasteful in the sense that it merely redistributes existing wealth without changing aggregate welfare. It is thus possible to identify and estimate rent-seeking costs of a tariff or quota, which may be added to the Harberger deadweight loss triangles of the previous chapter. These costs are thought to be empirically quite large and are therefore of considerable interest to the policy maker. However, it is again notable that the outcome under rent–revenue seeking differs depending on whether a tariff or a quota is used, yet another example of the non-equivalence of the two policies.

## 3.1. Quotas and heterogeneous product categories

We now consider the implications of applying an import quota to a class of heterogeneous or differentiated commodities. For ease of exposition, we shall work in terms of the following example. Suppose a country wishes to restrict imports of cars and that there are only two types of car, cheap cars (A) and expensive cars (B). Assume that type A and B cars are gross substitutes and that equiproportionate increases in the prices of both types of cars reduce total excess demand (=import demand) for cars but do not change the ratio of excess demands for the two types. In other words, we are assuming that both car types have similar income effects and similar cross-substitution effects with other commodities. Such an assumption is reasonable if the two car types are close substitutes for each other. Its role here is to make the composition of car imports depend only on the relative prices of the two car types. If one type becomes relatively more expensive, the proportion of that type of car in total car imports will fall (see Falvey, 1979). Finally, we assume that the market for import quota licences is competitive; thus, in equilibrium, there are no unexploited profit opportunities.

Now consider the effects of a volume quota, that is, a limit on the total number of cars imported, so that imports of A and B type cars ($M_A$ and $M_B$, respectively) satisfy the constraint

$$M_A + M_B = \bar{M}.$$

Let $p_A$ and $p_B$ be the domestic prices of cheap and expensive cars, and let $p_A^*$ and $p_B^*$ be their respective world prices. Given that the market for quota licences is competitive, importers allocate their imports between cheap and expensive cars so that the premium per car (the excess of the quota-distorted domestic price over the world price) is the same for both types of car. If this is not the case (say, $p_A - p_A^*$ is greater than $p_B - p_B^*$), then the profit per car will be greater from importing additional cheap cars and fewer expensive cars. Such reallocation will continue until either only one type of car is imported or until the premia for both types of car are driven to equality; that is, until

$$p_A - p_A^* = p_B - p_B^* = T. \quad (1)$$

Thus (assuming both types of car are imported in equilibrium), a volume quota is equivalent to a uniform specific tariff of $T$ per unit applied to the goods subject to import restriction. In the previous chapter we saw that a uniform specific tariff is an optimal instrument for restricting the volume of imports of a class of goods. It is not surprising that a volume quota (which achieves such a noneconomic objective directly) should be equivalent to such a policy. The interesting question here is how a volume quota affects the composition of imports within the restricted category relative to free trade. To answer this question, given our earlier assumptions, we merely have to check how the volume quota affects relative prices. Under the quota, the difference between domestic relative prices and free-trade relative prices is

$$\frac{p_A}{p_B} - \frac{p_A^*}{p_B^*} = \frac{p_A^* + T}{p_B^* + T} - \frac{p_A^*}{p_B^*} = \frac{T(p_B^* - p_A^*)}{p_B^*(p_B^* + T)} > 0 \quad (2)$$

since A is, by definition, the cheaper type of car ($p_A^* < p_B^*$). Thus a volume quota increases the relative price of cheaper cars, leading to a shift towards more expensive cars in the composition of imports. This tendency for volume quotas to bias the import mix towards more expensive items has been noted in relation to a number of markets to which quantitative import restrictions have been applied. Falvey (1979) cites the cases of textiles and steel imports into the United States, where restrictions expressed in volume terms have led to upgrading of the quality of goods imported. Similar effects have also been noted in relation to the voluntary export restraints imposed in 1981 on Japanese auto exporters to the United States (Collyns and Dunaway, 1987; Feenstra, 1984) and the orderly marketing agreements that protect the United States footwear industry (Aw and Roberts, 1986). All of these examples relate to voluntary export restraints (VERs) or orderly marketing agreements (OMAs) rather than explicit import quotas, but the basic restriction is of the same general kind. It is therefore not surprising that an upgrading effect should apply (for a discussion of some different features of the VER case, see Chapter 4). The same type of effect has also been observed in the composition of Australian car imports which were restricted by a volume import quota until 1988.

In contrast to a volume quota, a value quota does not distort the import mix at all. If the government wishes to restrict the value of car imports (at world prices) to some value (say, $V$ dollars), we have

$$p_A^* M_A + p_B^* M_B^* = V.$$



Importers facing a limit on the dollar value of car imports will now import type A and B cars up to the point at which the premium from a dollar's worth of additional imports is the same for both car types (any divergence between these premia for types A and B implies higher profits from importing more of the car type with the higher premium per dollar). Thus, in an equilibrium where both types of car are imported,

$$\frac{p_A - p_A^*}{p_A^*} = \frac{p_B - p_B^*}{p_B^*} = t. \quad (3)$$

The value quota is seen to be equivalent to a *uniform ad valorem tariff* at rate $t$ applied to all goods in the category. Again, this is not surprising, because such a tariff was seen in Chapter 2 to be an optimal means of restricting the total value of a class of imports. Given this equivalence, it is immediately clear that the relative price of type A and B cars is the same under the value quota as under free trade since

$$\frac{p_A}{p_B} = \frac{(1 + t)p_A^*}{(1 + t)p_B^*} = \frac{p_A^*}{p_B^*}. \quad (4)$$

Given the assumptions which make import composition depend only on relative price, the import mix under a value quota is the same as under free trade. The implication is that if a country wishes to restrict imports of a class of closely substitutable goods, then it can minimize the distortion associated with such a policy by restricting the value rather than the volume of imports of goods in the class.

The difficulty is that value quotas are difficult to implement, primarily because they are subject to manipulation. If, for example, foreign exporters can use transfer pricing to artificially reduce the c.i.f. import price ($p_A^*$ or $p_B^*$) of some or all of the restricted goods, they can recoup their profits via rents from the import licences (they may either hold the licences themselves or they may collude in some way with domestic importers). Uniform ad valorem tariffs are an acceptable alternative but may be unavailable if their use runs counter to GATT-negotiated tariff reductions. In any case, political factors also seem to favour the use of quotas rather than tariffs (see Chapter 8).

Thus it may be that a government (for political and administrative reasons) chooses volume quotas as its instrument of import restriction. We must then pose the question: is there a way of allocating these quotas which makes them less distortionary? One proposal which has been canvassed is the allocation of some or all of the quota for a particular class of goods by means of ad valorem tender bids; that is, importing firms submit bids for import licences expressed in ad valorem terms, and the bidders above the quota-clearing bid receive licences. Such an allocation scheme (which has been tried in the Australian car industry; see Industries Assistance Commission, 1980) is seen as replicating a uniform ad valorem tariff (or, equivalently, a value quota) and thus avoiding the distortions in the import mix associated with a volume quota. Unfortunately, unless such an allocation mechanism can somehow remove the underlying constraint on the volume of imports which is implicit in the use of a volume quota, it will make matters worse. We know that a uniform specific tariff is the optimal means of restricting the volume of a class of imports. It follows that using the equivalent of a uniform ad valorem tariff to achieve that same volume of imports involves a higher distortion. In particular, under the system of ad valorem tender bids, the marginal deadweight loss on foregone units of imports is higher for expensive items than for cheaper items, implying that the proportion of cheaper goods in total imports is too high under such a policy. Of course, allocation by ad valorem tender bids would be optimal if somehow the volume constraint could be (surreptitiously) replaced by a value constraint, but it is not clear how this could be implemented in practice. In short, there does not appear to be any way of making a volume quota "change its spots". These issues and other aspects of the Australian experiment are analysed in detail in Gibbs and Konovalov (1984).

We now turn to a range of cases in which tariffs and quotas are not equivalent.

## 3.2. Dynamic non-equivalence of tariffs and quotas

The simplest case of non-equivalence arises when the parameters of the economy are shifting over time. Here we consider one example of this so-called *dynamic non-equivalence*, the case of a shift in an industry's supply curve (Figure 3.1). As in the previous chapter, *DD* represents the total demand by residents for the importable good (food), and *D'D'* is the demand curve for domestically produced food. $p^*$ is the c.i.f. import price and $p_Q$ the domestic price under the quota Q. In the initial quota equilibrium (at *E*), the quota is equivalent to the implicit tariff of $p^*p_Q$ per unit. Output under both policies is $X_1$, consumption is $C_1$ and imports are equal to the quota.

Now let the supply curve *SS* shift upwards to *S'S'*. It is straightforward to show that the two policies are no longer equivalent. Under the quota, a new domestic equilibrium is established at *T*, where *D'D'* intersects the new supply curve; with imports fixed, equilibrium domestic price must rise to induce higher domestic output to meet domestic demand. Output falls to $X_2$, consumption falls to $C_2$ and domestic consumers pay a higher price than before ($p'_Q$). In contrast, under the tariff which is equivalent to the quota in the initial equilibrium ($p^*p_Q$ per unit), domestic price cannot rise above $p_Q$ (because there is a perfectly elastic supply of imports at the world price plus the tariff); hence output under the tariff falls further than under the quota (to $X_3$, where $p_Q$ intersects *S'S'*). Thus, from the point of view of the protected sector, the quota



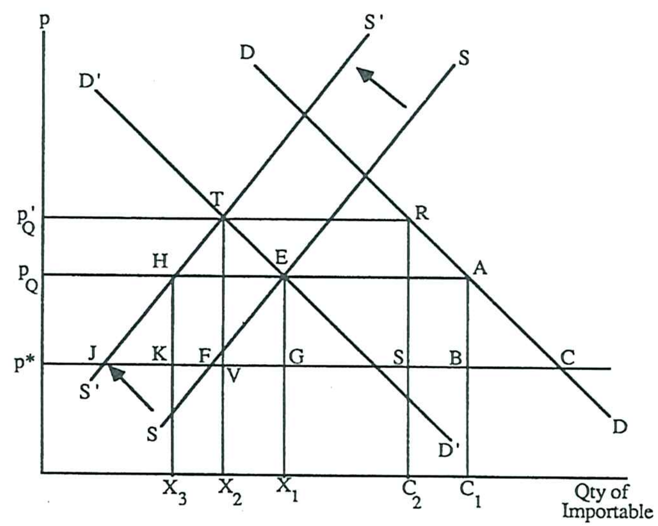


Figure 3.1

is the preferred policy because it results in a smaller reduction in output and a higher price than under the tariff. Rents in the sector are higher under the quota by the area $p'_QTHp_Q$. On the other hand, the quota imposes a higher deadweight loss on the community than the tariff, and the excess cost of the quota is given by the area (*HTVK* + *RABS*).

Clearly two policies which were equivalent before the shift in the supply curve are no longer equivalent after the shift. One aspect of this non-equivalence which is particularly worth noting is the manner in which protection under the quota adjusts upwards (from $p^*p_Q$ per unit to $p^*p'_Q$ per unit) in response to an increase in the industry's costs of production, whereas protection under the tariff remains fixed at $p^*p_Q$ per unit. It is in this sense that quota protection is often described as "open ended". Such open endedness is often perceived as a particular disadvantage of quotas; in particular, policy makers have the concern that an industry protected by quotas may, in the presence of other pressures (e.g. powerful unions), have less incentive to resist cost increases because part of the burden of higher costs can be passed on to domestic consumers. A tariff, on the other hand, imposes a discipline on firms because the firm must bear the full burden of any upwards shift in its cost curve.

We have now seen one example of dynamic non-equivalence of tariffs and quotas. A similar analysis could have been carried out for shifts in the demand curve or the c.i.f. import price. In each case, it is easy to verify that tariffs and quotas imply very different responses to such shifts.

There is a further important implication of the preceding analysis: If the effects of tariffs and quotas differ in response to shifts in parameters, then we might expect them also to differ in the presence of uncertainty (when parameters are shifting in a random way). In the next section we present a simple uncertainty model in which the two policies are clearly seen to differ in the presence of random fluctuations in world price and the demand curve for imports.

## 3.3. Tariffs and quotas under uncertainty

### (a) *Uncertain world price*

Suppose the small country we have been considering faces a randomly fluctuating world price for its importable good. Such price uncertainty is common in the real world, particularly with respect to raw materials. If, as the analysis of the previous section suggests, a tariff and a quota cannot be fully equivalent in this situation, then on what basis should they be compared? A popular basis of comparison (or non-economic objective) is the achievement of a common expected level of imports (see Pelcovits, 1976; Young and Anderson, 1980). Before we proceed, it is important to distinguish between the effects of specific and ad valorem tariffs when the world price is fluctuating. Under an ad valorem tariff, the world price movements are reflected in magnified movements in the domestic price; for example, a rise of $\varepsilon$ in world price implies a rise of $(1 + t)\varepsilon$ in the domestic price, where $t$ is the ad valorem tariff rate. However, under a specific tariff, the movements in the domestic price exactly match the movements in the world price. We shall begin by showing that a specific tariff yields a higher value of expected consumer surplus than a quota when the world price is fluctuating randomly.

In Figure 3.2, $D_m$ is the economy's demand curve for imports of the good, and $p^*$ is the mean value of its c.i.f. import price. Suppose that this price can take on values of $p^* + \varepsilon$ and $p^* - \varepsilon$, each with probability $\frac{1}{2}$.

We now compare the expected losses of consumer surplus (relative to free trade) under the two policies for a given expected quantity of imports. In Figure 3.2, $Q$ denotes both the level of the quota and the expected level of imports under the tariff. $p_Q$ is the domestic price under the quota regardless of the level of world price. In other words, the quota acts as a price-stabilizing tool, insulating the domestic economy completely from foreign price fluctuations. In this particular case (where the only disturbance is a fluctuating world price), the quota is formally identical to a variable import levy of the kind employed by the European Economic Community (EEC) to support the domestic prices of its agricultural products (variable import levies are discussed in detail in the next chapter). Under a specific tariff of $T$ dollars per unit, the domestic price is always $T$ dollars above the world price, with imports adjusting accordingly. In Figure 3.2, when the world price is $p^* - \varepsilon$, imports under the tariff are $Q_1$, and



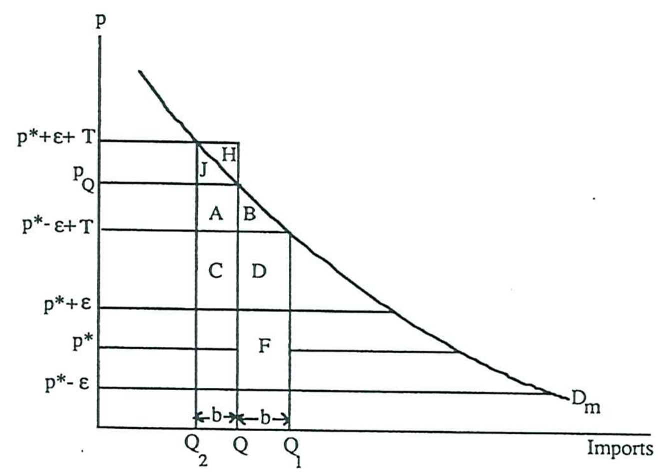


Figure 3.2

when the world price is $p^* + \varepsilon$, imports under the tariff are $Q_2$. Because $Q$ is, by definition, the mean of $Q_1$ and $Q_2$, the distances $Q_1Q$ and $QQ_2$ are equal. We call their common value $b$.

Now, using consumer surplus measures, consider the excess deadweight loss of the quota relative to the tariff for the two alternative values of the world price. When the world price is $p^* + \varepsilon$ the tariff has an excess DWL, which may be written in terms of the areas marked in Figure 3.2 as $(J + A + C)$. This is clearly less than the rectangular area $(J + H + A + C)$, which equals $bT$. When the world price is $p^* - \varepsilon$, an excess DWL is associated with the quota of $(B + D + F)$, which is greater than the rectangular area $(D + F)$, which also equals $bT$. Thus the expected excess DWL of the quota relative to the specific tariff is greater than

$$\frac{1}{2}(-bT) + \frac{1}{2}bT = 0.$$

In other words, the specific tariff is superior to the quota. It can also be shown to be superior to an ad valorem tariff. As explained earlier, an ad valorem tariff results in the world price fluctuations being magnified in the domestic price. In terms of Figure 3.2, under an ad valorem tariff when world price is high, the domestic price is above $(p^* + \varepsilon + T)$; and when world price is low, the domestic price is *below* $(p^* - \varepsilon + T)$. This means that if a specific tariff is replaced by one expressed in ad valorem terms, $(J + A + C)$ is widened to the left while area $(B + D + F)$ is widened by an equal amount to the right (since $Q$ is still the arithmetic mean of the two alternative import levels under the tariff). However, the height of the addition to $(J + A + C)$ exceeds the height of the addition to the area $(B + D + F)$. It follows that the excess DWL of the tariff when the price is high is increased relative to the excess DWL of the quota when the price is low, implying that the expected excess DWL of a quota over a tariff is lower when the tariff is ad valorem rather than specific. We conclude that the specific tariff dominates both policies.

Whether an ad valorem tariff can also be superior to a quota in this case is unclear from the preceding analysis. However, Pelcovits (1976) was able to show that if the demand curve for imports is linear, then an ad valorem tariff is superior to a quota provided the tariff rate (and the associated magnification of world price movements) is not too high.

The result that a specific tariff dominates an ad valorem tariff and an import quota as a means of attaining a given expected level of imports, should not be at all surprising to the reader who followed our earlier analysis of import restriction for heterogeneous product categories. The constraint on imports is just like a volume quota on a heterogeneous class of goods with the two "types" of goods subject to restriction being the importable in the two states of the world (high price and low price). This parallel can be seen immediately if we write the constraint on expected imports as

$$\frac{1}{2} M_A + \frac{1}{2} M_B = Q, \quad \text{that is,} \quad M_A + M_B = 2Q$$

where $M_A$ and $M_B$ are imports in the two states of the world. We know that a uniform specific tariff is the optimal means of achieving such an objective. A specific tariff which is uniform across states of the world is just an ordinary specific tariff. On the other hand, a fixed import quota ($Q$) implies a fixed domestic price in both states of the world and thus a higher implicit tariff per unit in the low-price state of the world, $[p_Q - (p^* - \varepsilon)]$ per unit, than in the high-price state, $[p_Q - (p^* + \varepsilon)]$ per unit. This means that under a quota, the consumer surplus foregone on the marginal unit is higher in the low-price state of the world than in the high-price state, implying that expected consumer surplus can be increased by importing less when world price is high (since the cost of foregoing a unit is lower in this state) and importing more when it is low. Under an ad valorem tariff (which implies the same percentage loss of surplus on the marginal unit in both states), the marginal loss per unit is higher in the high-price state, implying gains to society from reallocating some units of the fixed quantity of imports from the low-price state to the high-price state (i.e. a narrowing of the domestic price dispersion back towards the specific tariff).

In the preceding analysis, we were comparing a tariff and a quota which yielded the same expected quantity of imports. An alternative common objective would be a given *value* of imports (i.e. a given foreign exchange expendi-



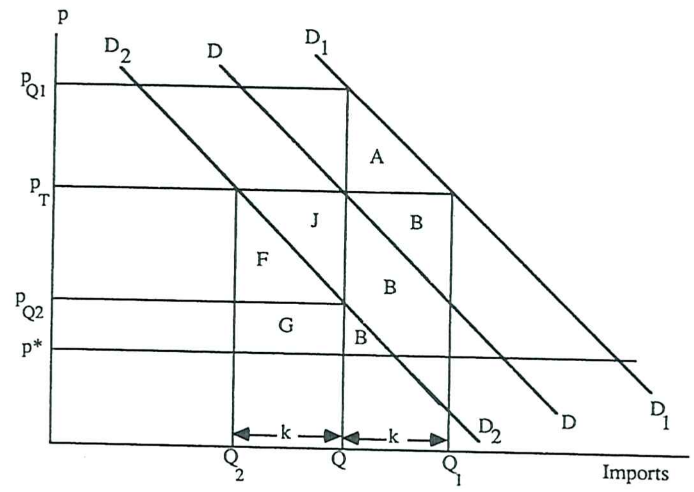


Figure 3.3

ture). The optimal outcome in this case would be one which equates the loss of consumer surplus on the marginal unit of foreign exchange (rather than the marginal physical unit) across states of the world. This outcome is achieved by having the excess of domestic over world price the same percentage of world price in both states (i.e. an ad valorem tariff). See Young and Anderson (1980) for a formal proof that an ad valorem tariff is indeed the most efficient policy for import restriction in this case.

### (b) *Uncertain demand for imports*

It is also straightforward to compare a tariff and a quota which yield the same expected level of exports when there are random fluctuations in the demand curve for imports (arising from fluctuations in the demand and/or the supply schedules for importables). Without fluctuations in the world price, the ad valorem and specific tariffs are equivalent. Moreover, when the demand curve shifts, the price changes under a quota but not under a tariff; under a tariff, imports bear the full burden of adjustment (as in the case analysed in Section 3.2). This is illustrated in Figure 3.3.

The (assumed fixed) world price of the importable is $p^*$, and $Q$ is the expected level of imports under both policies. $DD$ is the mean of the two possible demand curves $D_1D_1$ and $D_2D_2$, each of which is equally probable; thus, at any given price, the horizontal distance between $DD$ and $D_1D_1$ equals the horizontal distance between $DD$ and $D_2D_2$. Under the quota $Q$, in the high-demand state of the world $(D_1D_1)$ the domestic price is $p_{Q1}$, whereas in the low-demand state of the world ($D_2D_2$) the domestic price is $p_{Q2}$. The tariff which yields the same expected quantity of imports as the quota is $p^*p_T$ per unit. This involves a domestic price $p_T$ in both high- and low-demand states with imports $Q_1$ in the high-demand state and $Q_2$ in the low-demand state. Clearly, $Q_2Q = QQ_1$; let this common value be denoted by $k$.

When the demand curve is in position $D_1D_1$, there is an excess DWL of the quota over the tariff given by the area $(A + B)$ in Figure 3.3. When the demand curve is $D_2D_2$, there is an excess DWL of the tariff over the quota equal to the area $(F + G)$. Clearly, $(A + B) > B = k(p_T - p^*)$. It is also clear from the diagram that $(F + G) < (F + G + J) = k(p_T - p^*)$. It follows that the expected DWL of the quota over the tariff is

$$\frac{1}{2}\{A + B - (F + G)\} > \frac{1}{2}[B - (F + G + J)]$$
$$= \frac{1}{2}[k(p_T - p^*) - k(p_T - p^*)] = 0.$$

Thus the quota involves a higher expected DWL than the tariff and is therefore the inferior instrument in this case.

### (c) *Risk aversion*

The foregoing analysis suggests that in the presence of uncertainty a tariff is a better tool for restricting imports than a quota. This reinforces the common view among economists favouring tariffs over quotas. On the other hand, this conclusion has been reached by choosing the policy which (subject to the constraint on imports) maximizes expected consumer surplus. It is well known however that consumer surplus is a valid measure of welfare only under special conditions. In particular, for the case in which only one commodity price is changing, as in sub-section (a), changes in utility are correctly measured by consumer surplus methods only if the coefficient of relative risk aversion equals the income elasticity of demand for the good in question (see Turnovsky, Shalit and Schmitz, 1980). In other words, the analysis of the previous section was placing implicit limits on the degree of risk aversion of producers and consumers.

For the reader who is unfamiliar with the concept of risk aversion, we offer the following brief explanation. Consider an individual who is faced with a gamble in which the prize is $x$ dollars with probability $\pi$ and $y$ dollars with probability $(1 - \pi)$. The expected payoff for such a gamble is $\pi x + (1 - \pi)y$ dollars. If the individual would rather receive this expected payoff with certainty than participate in the gamble, he is said to be risk averse. For an expected utility function $u$, risk aversion amounts to assuming that

$$u(\pi x + (1 - \pi)y) > \pi u(x) + (1 - \pi)u(y) \qquad (5)$$



that is, the utility of the expected payoff from the gamble exceeds the expected utility of the gamble itself. Because Equation (5) is the same as the condition for the expected utility function $u$ to be concave, risk aversion is synonymous with concavity of the expected utility function. The coefficient of relative risk aversion is just a measure of this concavity and is defined as

$$\rho \equiv -\frac{Yu''(Y)}{u'(Y)}$$

where $Y$ denotes income. Note that $\varrho$ is defined so as to be independent of the units in which expected utility is measured.

If agents are risk averse, then any fluctuations in real income associated with a particular policy tend to render the policy less attractive. For example, in the case of a randomly fluctuating world price, as in sub-section (a), a tariff yields greater fluctuations in real income than the quota. This is because it restricts imports in the state of the world when the world price is high, thus imposing a cost on the economy additional to the real income loss associated with the higher world price. Similarly, when the world price is low, the country benefits from both the expanded trade under the tariff and the positive income effect of a lower world price. Although the income effects of the changes in world price are present under both policies, the point is that the changes in imports under the tariff reinforce these income effects, increasing rather than reducing the magnitude of fluctuations in real income. If individuals are risk averse, these fluctuations reduce the attractiveness of the tariff, and if the degree of risk aversion is sufficiently high, the quota is the preferred policy (see Young and Anderson, 1982, for proof). One possible implication of this result is that the prevalence of quantitative import restrictions in the real world is evidence of high degrees of risk aversion on the part of economic agents. However, given the range of other explanations for the spread of quotas and similar policies (see Deardorff, 1987), it is perhaps unwise to put too much weight on this particular interpretation of events. Finally, we emphasise that the preceding results single out a particular policy (e.g. a quota) as preferable to other policies, assuming that import restriction is the government's objective. The results should in no way be construed as offering an argument for trade restriction or protection per se.

### (d) *Protection as insurance*

So far we have considered the use of tariffs or quotas to achieve a non-economic objective such as a targeted expected value of imports. We now consider an argument that an economy may benefit from protection in the presence of uncertainty even if there is no underlying non-economic objective. Not surprisingly, the argument is based on second-best considerations, specifically the imperfect operation of insurance markets; but like all second-best arguments, interest centres on the details of the distortion or market failure which the proposed policy is to correct and the manner in which it is corrected.

Newbery and Stiglitz (1984) consider the following example of "Pareto-inferior trade". Suppose there are two crops produced in the world, one risky (e.g. because of the possibilities of drought or frosts) and one safe. If there is no international trade, the fall of the output of the risky crop in a bad season drives up its market-clearing price, reducing the fall in producer incomes. Similarly, a fall in the price when there is a good crop reduces the rise in incomes in a good season. For the special case in which the demand curve for the good has unitary price elasticity, the price changes just offset the crop variations, keeping income the same from season to season. In any case, changes in the market-clearing price work to reduce the variance of producer incomes. Essentially, the risk is borne by consumers who face highly variable prices. When the economy is opened up to trade with the rest of the world, it is possible that fluctuations in producer incomes will replace fluctuations in price. This would be the case if, for example, the rest of the world has a good season while the country in question has a bad season or vice versa. For the special case in which the country's output is perfectly negatively correlated with that of the rest of the world, the world output of the crop is constant at any given world price; thus, for a given world demand curve for the crop, the equilibrium world price is constant regardless of the season in each country. This means that the income of producers in a particular country varies in proportion to their output, being high in a good season and low in a bad season. International trade has made income from the risky crop more risky but, by stabilizing world price, has reduced the risk borne by consumers. If producers are risk averse, some resources will shift out of production of the risky crop into the safe crop, pushing up the world market-clearing price of the risky crop. If consumers are not too risk averse, they will lose more because of the higher average price than they gain from reduced price variability. On the other hand, if producers are sufficiently risk averse, they lose more from the increased variance of their income than they gain from a higher average price. It is possible that all are made worse off by opening the economy up to international trade.

If insurance markets were perfect, agents could completely insure against risk. Then the riskiness of particular activities would not matter, and international trade would necessarily be welfare-improving in the usual sense. However, it is well known that agents' ability to cover risk by means of insurance is constrained. Insurance markets for some contingencies may not exist at all, whereas for others, only incomplete insurance may be offered. This "incompleteness" of insurance markets is explained in part by the phenomena of moral



hazard and/or adverse selection (see Arrow, 1970). *Moral hazard* exists when insurers cannot observe the level of effort expended by the insured to prevent an undesirable risky event from taking place, and when the presence of insurance reduces such effort by the individual; for example, individuals who have insured the contents of their home against theft may be careless in securing their home against break-ins. *Adverse selection* arises because the insured may have better information about the riskiness of the events against which they are insuring than the insurer; for instance, chronically ill people are more likely to purchase medical insurance. The presence of either of these factors in a particular insurance market may imply that either the market is inoperative or the insurance it offers against risk is only partial. Given such constraints on agents' ability to insure against risk, risk will matter, and the Pareto-inferior trade we have just described would appear to be a distinct possibility. The implication is that individual economies can gain from restricting trade and that protection has a role in offering "insurance" which would not otherwise be provided. The role of trade policy in this context has been explicitly modelled by Eaton and Grossman (1985).

The main problem with this "tariffs as insurance" argument is that there is no reason to suppose that such government-provided insurance is immune from the problems of moral hazard and adverse selection which led to incompleteness of private insurance markets. This point has been made cogently by Dixit (1987a, 1989), who argues that the sources of a "failure" of the market should be explicitly modelled when considering government policies which are designed to correct the market failure. Consider, for example, how moral hazard and adverse selection might enter the Newbery and Stiglitz scenario. Producers insure against the contingency of a bad harvest; and if they think that the insurer cannot monitor their effort, they may reduce their own productive efforts, safe in the knowledge that their income in the low-crop state of the world is assured[1] (moral hazard). In addition, farmers who are less adept at coping with or preparing for bad seasons, whose land is unsuitably located, who are growing varieties of crops which are more susceptible to bad seasons, and so forth, will be the most likely purchasers of any form of crop or income insurance (adverse selection). The private insurance industry responds to these problems by offering incomplete coverage of risks. Under free trade, output of the risky activity is accordingly less than it would otherwise be. Now suppose the government in one country protects producers of its risky crop by imposing a tariff on imports of that crop. More factors then move into that sector, implicitly taking up the tariff insurance offered by the government. However, these factors tend to be the ones which the private insurance sector was attempting to discourage: the producers and land for whom the downside risk is higher and the producers and workers who are more likely to respond to insurance of their income by reducing their effort. Thus the tariff has merely imposed costs on the community which the private insurance sector sought to avoid. Such considerations cast considerable doubt on the validity of the insurance argument for protection.

This completes our analysis of tariffs and quotas under uncertainty. We now move to an entirely different extension of the basic model, inclusion of the rent-seeking costs associated with tariffs and quotas. Allowing for these costs is the first step in the process of identifying costs of protection beyond the "Harberger triangles" used in this and the previous chapter for evaluating different policies. Introducing rent seeking into the model also provides a further instance of the non-equivalence of tariffs and quotas.

## 3.4. The rent-seeking costs of tariffs and quotas

The term *rent seeking* was used by Anne Krueger (Krueger, 1974) to describe the competition for the rents that accompany most forms of pre-existing government intervention or regulation. The real resources used up in such activity are generally regarded as being wasted because their use does not create wealth but merely transfers existing wealth between groups or individuals[2] (see Tullock, 1967). In this section we consider how measures of the social cost of protection should be modified to allow for additional waste arising from rent seeking. Before we consider the different forms that this rent seeking can take, it is worth noting that we are not considering the lobbying for or against a particular policy which precedes the introduction of that policy. Such lobbying (which is also frequently called rent seeking) can be viewed as something which would have occurred anyway (whether or not the policy was adopted) and so should not be counted as a cost of the policy.[3] For the remainder of this section we turn our attention to the rent seeking which is induced by pre-existing tariffs and import quotas. Lobbying aimed at persuading government to introduce a distortion will be analysed in Chapter 8 in which we consider why a particular protective instrument is adopted and what determines its level.

In her original (1974) article, Krueger considers the particular case in which the rents being contested arise from quantitative trade restrictions (e.g. import quotas). In the case of an import quota, these rents occur if import licences are handed out free of charge. The rent seeking which ensues is a contest to capture a share of the scarce import licences, and it may take various forms, the most obvious of which are direct lobbying of politicians (involving such costs as trips to the capital city and establishing premises there), overinvestment in physical plant (to quality for licences allocated in proportion to firms' capacities), bribery[4] and other forms of corruption. There is reason to believe that in many countries these quota rents (and the associated rent-seeking costs) can be very large indeed. For example, Krueger calculates that in India in 1964, total rents amounted to 7.3 per cent of GNP. Of this figure, rents associated with import



licences represented over two-thirds. In 1968 in Turkey, quota rents were calculated to be about 15 per cent of GNP (Krueger, 1974 p. 294). Moreover, there is reason to suppose that the rent seeking associated with quota licences is just the tip of the iceberg. For example, by including other distortions such as price controls and rationing, Mohammed and Whalley (1984) estimate total contestable rents associated with pre-existing government policies in India to be somewhere between 30 and 45 per cent of GNP.

When the instrument of import restriction is a tariff, we would similarly expect to find some resources directed at capturing a share of the recycled tariff revenue; the frequent, highly visible campaigns by interest groups for government funding of particular projects constitute an obvious example of such revenue seeking (Bhagwati and Srinivasan, 1980). The contest for the tariff revenue may however be less clearly defined than the contest for the import quota rents, with the latter usually being conducted among a small number of participants, each facing a clearly defined expected return (a certain number of quota licences). In contrast, the expected return from the tariff revenue (which is traditionally spread over a large number of people) may be so small that it is not profitable for any individual to incur the cost of entering the contest. It seems more likely that the revenue-seeking contest takes place in a broader sphere, the object being a share of total government revenue. If this is the case, then we are faced with the problem that the lobbying for revenue may occur whether the tariff is there or not and so cannot be counted as a specific cost of the tariff. For this reason, quota rent seeking appeals as a more interesting and relevant notion than tariff revenue seeking. Nevertheless, there are many instances, particularly in developing countries, in which import duties generate a sizable proportion of total revenue, and the notion of tariff revenue seeking does seem appropriate in such cases.

We now consider both the effects and the social costs of rent seeking. In evaluating the latter, the economist must consider two distinct questions:

i. What is the value of resources dissipated by the rent seekers in their quest for the rent or revenue? In particular, what relationship does this resource dissipation bear to the value of the rent contested? A standard working assumption here is that the value of resources dissipated equals the value of the rent or revenue contested. A partial equilibrium view (partial because the rent-seeking sector itself is excluded) combined with the assumption that rent seekers compete for the entire tariff revenue or quota rents implies that the social cost of a quota or its tariff-equivalent is equal to the sum of areas $A$ and $B$ (the Harberger triangles) in Figure 3.4 and the rectangle of quota rents or tariff revenue, $R$ (referred to as the *Krueger rectangle* for obvious reasons). This would certainly be a convenient way of evaluating the costs of protec-

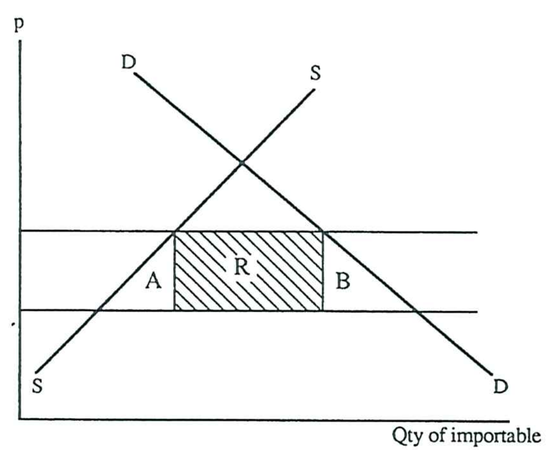


Figure 3.4

tion; in the absence of a theory of the rent-seeking sector, the usual practice has been to use $R$ as an estimate of rent-seeking costs.

ii. Accepting for the moment that rent seekers waste resources equal in value to the rent–revenue rectangle, does that rectangle represent the cost of rent seeking to society when general equilibrium considerations are taken into account?

Question (i) has been the subject of much research for the case in which a single rent seeker captures all of an indivisible rent, such as the granting of a government contract or the securing of a monopoly position as a result of government regulation (see Corcoran, 1984; Hillman and Katz, 1984; Tullock, 1980). The relevant case here, in which rents are shared among the rent seekers (as is usually true for tariff revenue or quota licences) has been analysed by Long and Vousden (1987). They show that the Krueger rectangle is a good *long-run* measure of the resources dissipated by the rent seekers if (i) there is free entry (exit) into (out of) the rent-seeking activity; and if either (ii) individuals are risk-neutral or (iii) individuals are risk-averse and the rent contested is a small proportion of the rent seekers' total wealth. Certainly, risk aversion (together with the riskiness of the rent-seeking activity) does reduce the proportion of rents dissipated below unity. However, this effect is negligible if the rent in question is a small proportion of the individuals' total wealth.

We now turn to question (ii). Assuming that the value of rents sought *is* a good measure of the value of resources dissipated in seeking those rents, does that resource cost equal the actual social cost



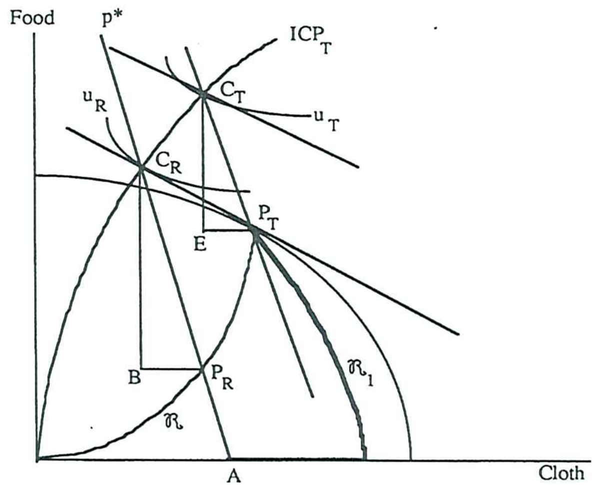


Figure 3.5

of the rent-seeking activity when general equilibrium effects are taken into account? To make the main points in the simplest possible way, we start with the case in which all of the tariff revenue or quota rents are sought. We shall then proceed to the more realistic case in which only part of the revenue or rents is sought.

## (a) *Entire rent sought*

We begin by considering the case of a tariff, illustrated in Figure 3.5. In the absence of revenue seeking or lobbying, tariff-distorted equilibrium involves production at $P_T$, consumption at $C_T$ and utility $u_T$. Now suppose that all of the tariff revenue is sought by competitive rent seekers. Given our assumption that the value of resources dissipated in this process equals the revenue sought, the tariff revenue effectively "disappears", and domestic factor income equals expenditure valued at domestic prices; that is, the economy's consumption point is moved down the income consumption path (ICP) for the tariff-distorted price ($ICP_T$) to position $C_R$. Because of the balance-of-trade constraint, the economy can only consume a bundle equal in value to what it produces, so the new production point must lie on the world price line $p^*$ through $C_R$.

The production point $P_R$ lies inside the production frontier, reflecting the loss of productive resources to revenue seeking. To fix $P_R$, we construct the so-called Rybczynski curve ($\mathfrak{R}$) through $P_T$. This represents the locus of outputs of food and cloth as capital and labour are withdrawn in the ratio required by revenue- or rent-seeking activity at the given (tariff- or quota-distorted) domestic price ratio. Whereas the Rybczynski theorem and its extensions relate to the effects of various patterns of factor growth on sectoral outputs, this Rybczynski curve illustrates that process in reverse. Because the economy's production point must be on both the Rybczynski curve $\mathcal{R}$ and the world price line $p^*$, it will be at their intersection point $P_R$, as shown in Figure 3.5.

A number of features of the equilibrium in Figure 3.5 are worth noting:

i. Imports ($BC_R$) probably differ from their value ($EC_T$) in the absence of revenue seeking. Clearly, the level of food imports depends on the form of the curve $\mathcal{R}$, which in turn depends on the lobbying technology. If the quantity of imports is changed by revenue seeking, then the equilibrium tariff revenue also changes as does the value of resources dissipated in revenue seeking.

ii. The equilibrium consumption point and utility are independent of the lobbying technology (they are always $C_R$ and $u_R$, respectively).

iii. Although the Rybczynski curve is drawn with a positive slope in Figure 3.5, let us note that it could easily be negatively sloped depending on the relationship between the capital intensities in food, cloth and lobbying. To take a simple case, suppose lobbying uses only labour and that food is the labour-intensive good. Then the lobbying activity amounts to a contraction of the economy's labour supply. Rybczynski's theorem applies (in reverse), and the capital-intensive sector expands while the labour-intensive sector contracts, giving a negatively sloped Rybczynski curve like $\mathcal{R}_1$ in Figure 3.5. This is an interesting case (discussed in detail by Bhagwati and Srinivasan, 1980),[5] which opens up the possibility that the economy's production point might move back towards the free-trade point (along $\mathcal{R}_1$), a potentially welfare-improving change. However, in the present case with full revenue seeking, consumption is anchored at $C_R$, and $p^*$ cannot intersect $\mathcal{R}_1$ in its negatively sloped region. Equilibrium production must be on the price line $p^*$. In the case illustrated, it is attained at point $A$: The lobbying activity expands, moving the economy down $\mathcal{R}_1$ until it hits the cloth axis; then, as further resources are drawn into the lobbying effort, the economy's only remaining output (cloth) contracts further along the horizontal axis until point $A$ on price line $p^*$ is reached. Consumption is then at point $C_R$, and there are no remaining uncaptured rents to attract further revenue seekers into the activity. As noted in (ii), equilibrium consumption and utility were not affected by the particular form of the Rybczynski curve.

iv. For full revenue seeking, the social cost of the rent seeking valued at tariff-distorted prices is exactly equal to the value of the tariff revenue in the absence of revenue seeking (the horizontal distance between the domestic price lines through $C_T$ and $C_R$). In general, this differs from the actual tariff revenue in the presence of revenue seeking, which is the horizontal distance between the domestic price line through $C_R$ and a parallel line through $P_R$. Whether the actual tariff revenue overstates or understates the true cost of the revenue seeking depends on where the Rybczynski curve cuts the price line $p^*$ (in Figure 3.5, imports and tariff revenue are higher after rent seeking than before), which in turn reflects how the withdrawal of resources associated with revenue seeking changes relative outputs of food and cloth. Thus the Krueger rectangle is not an exact measure of the cost, but it may still be a reasonable approximation.



These remarks apply to the case of a tariff. Let us now reinterpret Figure 3.5 as applying to an import quota for which all of the quota licences are sought. As for the tariff, let the equilibrium without rent seeking involve production at $P_T$, consumption at $C_T$ and utility $u_T$. Then, when all of the quota rents are sought, the consumption point moves down onto the domestic price line tangent to the economy's production point. It is here that we encounter a nonequivalence between tariffs and quotas. You will recall that imports in the revenue-seeking equilibrium do not, in general, equal imports in the absence of revenue seeking. Indeed, in the diagram as drawn, they tend to be larger with revenue seeking than without it ($BC_R > EC_T$). However, under a binding quota, imports must necessarily be $EC_T$ with or without rent seeking. Thus, with rent seeking, the quota is violated unless the relative price of food rises to clear the excess demand for food. The economy's consumption point is still a point of tangency between a community indifference curve and the domestic price line tangent to the economy's production point, but the slope of that domestic price line has changed (so as to clear the domestic market for food), and the production point has rotated around the production frontier (with more food being produced). The quota is clearly not equivalent to its implicit tariff in this case. In fact, it is inferior to the tariff because it involves a higher equilibrium domestic relative price of food.[6] Furthermore, the equilibrium price under the quota depends on the lobbying technology as embodied in the Rybczynski curve (the position of which determines the level of imports $BC_R$ before the price adjusts to clear the market). This non-equivalence of tariffs and quotas in the presence of rent or revenue seeking can be reinforced if the rent-seeking technology differs from the revenue-seeking technology (different Rybczynski curves).

In general, the entire tariff revenue or quota rents are not sought by the lobbyists. Prior expenditure commitments by governments tie up considerable amounts of revenue, which lobbyists realize is not worth competing for. Similarly, import licences may sometimes be allocated (at least partially) by mechanisms which are immune from rent seeking (e.g. by lottery), and such licences

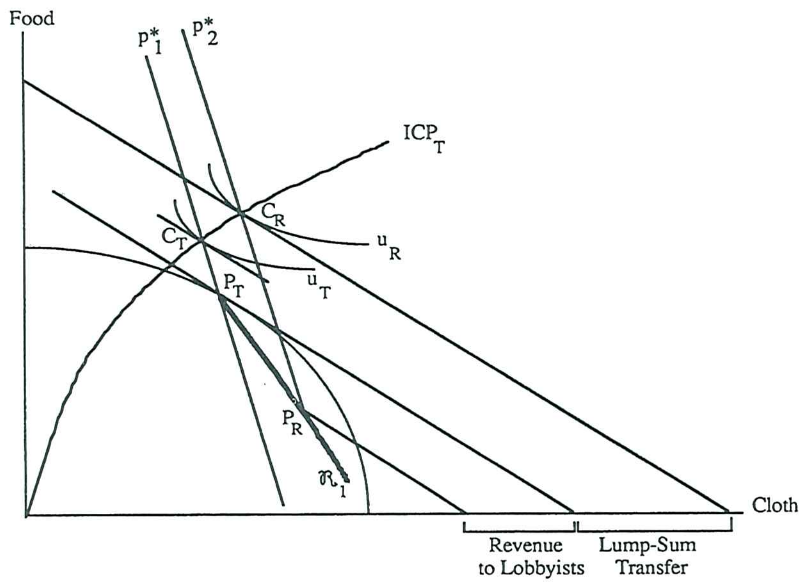


Figure 3.6

obviously do not give rise to any lobbying. The following sub-section considers what happens in such cases.

### (b) *Partial rent seeking*

Once again, we consider tariffs first. Suppose lobbyists seek only part of the tariff revenue, while the remainder is distributed to the community as a lump-sum transfer. Suppose also that the Rybczynski line is of the negatively sloped form $\mathcal{R}_1$ and lies to the right of the world price line $p_1^*$ through $P_T$. Then it can be seen from Figure 3.6 that a tariff-ridden economy may be better off with revenue seeking than without it.

As before, the tariff-distorted equilibrium without revenue seeking is at $P_T$, $C_T$. With partial revenue seeking, production is at $P_R$ on $\mathcal{R}_1$ and the consumption point is at $C_R$ on the world price line $p_2^*$ through $P_R$. Utility is $u_R$, a welfare improvement over the no-lobbying, tariff-distorted equilibrium ($u_T$) so that, as suggested, revenue seeking actually benefits the community. This result is made possible by the recycling of part of the revenue to the community, which allows a positive income effect to work with the negatively sloped Rybczynski curve to move the economy's consumption point out along the income consumption path $ICP_T$. The details of the process are as follows. Lobbying



withdraws resources in a way which causes production of the exportable to expand at the expense of the importable, thereby moving the world price line out from $p_1^*$ to $p_2^*$. Imports rise, increasing the tariff revenue, including the lump-sum transfer component of that revenue (and also, presumably, the revenue captured by lobbyists).[7] The resulting increase in income (at domestic prices) provides the means of sustaining the higher consumption point $C_R$. Although this result may seem counter-intuitive, it has to be remembered that lobbying (which is equivalent to contraction of factor supplies) is being introduced into a system which is already characterized by a distortion (the tariff). The economy is thus in a second-best situation, and the introduction of lobbying may lead to a welfare improvement. The loss due to dissipation of resources in lobbying activity is more than offset by the benefits of a pattern of production which is closer to free trade.

We now turn our attention to quotas, assuming that only part of the quota licences are actively sought. Once again, the crucial difference between a tariff and a quota is that imports can change under a tariff but are fixed under a quota. This constancy of imports under the quota prevents any welfare-improving income effects through increased recycled quota rents. Thus the gains from increased trade which were potentially attainable under a tariff with partial revenue seeking are unattainable under an import quota (i.e. $u_R$ necessarily lies below $u_T$). We conclude that rent seeking cannot increase welfare in the presence of an import quota (see Anam, 1982). Nevertheless, if quota rents are only partially sought, there is no reason why $P_R$ should not lie on the negatively sloped portion of $\mathcal{R}_1$, implying some reduction of the cost of rent seeking below the value of rents sought because of an improved allocation of resources between food and cloth.

This completes our analysis of tariffs and quotas in the presence of distortion-triggered lobbying. We now return to the question of how to measure the cost of protection when resources are dissipated through rent seeking.

### (c) *The cost of rent seeking*

As explained earlier, there may be some justification for assuming that the Krueger rectangle is an accurate measure of the resources used by rent seekers. The question remains: does the value of resources dissipated necessarily equal the social cost of rent seeking in a general equilibrium context? On the basis of the foregoing analysis, the answer to this question is clearly no. We have seen, for example, that the withdrawal of factors for lobbying may take the economy to a more efficient production point and thereby reduce the cost of rent seeking below the rent rectangle (the case of quotas) or even make the cost negative (the welfare improvement in the case of partial tariff revenue seeking). It is obviously safer to work with a full general equilibrium model.

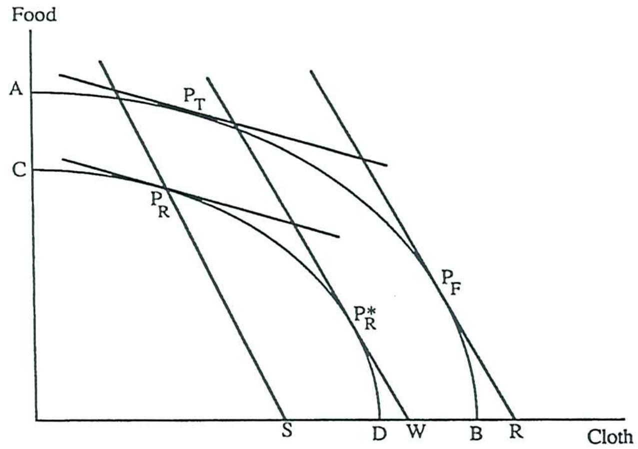


Figure 3.7

Bhagwati (1980) considers the possibilities that may arise if the two-sector general equilibrium model is used to estimate the deadweight loss of a tariff or quota. Because the consumption cost is largely unaffected by the presence of rent seeking, we consider only the estimation of the production cost. In Figure 3.7, *AB* represents the economy's production frontier with $P_F$ the free-trade production point and $P_T$ the production point under a tariff in the absence of lobbying. $P_R$ denotes the actual production point of the economy with the tariff and the revenue seeking that accompanies it. It may be thought of as a point of tangency of a tariff-distorted price line and a "shrunk" production frontier (*CD*), which represents the locus of efficient outputs of food and cloth when the capital and labour required for lobbying have been removed from productive use. If those same resources were being used for lobbying at *free-trade* prices, the economy would produce at $P_R^*$. Of the four production points we have identified, only one ($P_R$) is observable. Presumably, in estimating the standard production cost of protection, we would choose a procedure which minimizes the number of hypothetical production points we have to calculate. We can observe $P_R$ and estimate either $P_F$ or $P_R^*$.[8] If we estimate $P_F$, then we conclude that the production cost of the tariff is *SR*. This will in fact be the entire production cost including any costs or benefits arising from rent seeking. It would therefore be unnecessary in this case to derive a separate estimate of the cost of rent seeking per se. On the other hand, we may form our estimate of the free-trade production point on the assumption that the economy's supplies of capital and labour are the same as at $P_R$ (i.e. we are not aware of the leakage of resources to distortion-triggered lobbying). This would lead us to point $P_R^*$ on the inner production frontier *CD*, and we would estimate the production cost of the tariff as *SW*, omitting the component *WR* of the total production cost. *WR* would then have to be estimated separately as a cost of rent seeking. It is possible that the Krueger rectangle (revalued in terms of world prices) may be a suitable approximation for *WR*, but that remains an open question.



## 3.5. Summary

In this chapter we have explored further issues relating to tariffs and import quotas in competitive markets. One of these, the quality upgrading that accompanies volume quota restriction of imports of a group of goods, was explained by noting the equivalence between a volume quota and a uniform specific tariff on the one hand and a value quota and a uniform ad valorem tariff on the other. Because a uniform specific tariff makes the more expensive items in a restricted group relatively cheaper, it encourages a shift towards imports of these items (upgrading). The same is true of a volume quota. On the other hand, a value quota does not significantly affect relative prices within the restricted group and so does not induce a shift in the composition of imports.

Although tariffs and quotas have been seen to be equivalent in the simple static competitive model, one of the main points of this chapter is that this equivalence breaks down when other elements are added to the problem. In particular, the shifting of parameters over time (in either a deterministic or a stochastic sense) was seen to imply very different outcomes under the two policies. In the case of an uncertain world price, it was found that a government wishing to constrain the expected volume of its imports achieves a higher level of expected consumer surplus by using a specific tariff than by using either an ad valorem tariff or an import quota. On the other hand, if it wishes to constrain the expected value of imports, expected consumer surplus is highest under an ad valorem tariff. These results are changed if economic agents are sufficiently risk averse, in which case expected consumer surplus is not a valid criterion for policy comparison; with sufficient risk aversion, the price (and income) stabilizing properties of an import quota make it the preferred policy. Finally, it should be emphasised that even if a particular policy is found to be the most efficient means of restricting trade, it remains a *second-best* policy for a small country. The first-best policy is not to restrict trade at all.

If rent- or revenue-seeking costs are incorporated into the model, further evidence is provided of the non-equivalence of tariffs and quotas. In addition, the costs of activity directed at capturing a share of the quota rents or tariff revenue is thought to be an empirically significant component of the overall cost of protection. The value of resources used in this non-productive activity may be reasonably approximated by the value of the rents or revenue contested, the so-called Krueger rectangle. However, general equilibrium effects associated with the waste of productive resources by the rent-seeking sector may, in certain cases, reduce the welfare cost of the rent seeking below the area of the rectangle.