LORRAINE EDEN

Taxing Multinationals: Transfer Pricing and Corporate Income Taxation in North America

UNIVERSITY OF TORONTO PRESS
Toronto Buffalo London

these problems and the various solutions that have been suggested by the OECD 1994 pricing report and the recent U.S. section 482 regulations.

The simple rule – set the transfer price equal to the marginal cost of the exporting firm – changes when there are costs to using external markets such as tariffs and other trade barriers or differences in corporate tax rates. In these cases the MNE as an integrated business has more degrees of freedom in terms of arbitraging these market imperfections than do unrelated firms. The profit-maximizing and money transfer prices become determinate. The transfer price which the MNE, of its own volition, chooses on profit-maximizing grounds is not likely to be the arm's length price on the external market, nor is it likely to be the efficient transfer price chosen in the pre-tax situation. We turn to these questions in Chapter 6, where we first examine two simple cases, tariffs and a profit tax, and then develop a more realistic model of transfer pricing in which there are corporate income tax differentials.

6

Taxing Multinationals in Theory

Introduction

The topic 'taxing multinationals' must be covered from both the theoretical and empirical perspectives. We do this in two chapters: this one (Chapter 6: Theory) and the next (Chapter 7: Practice).

In this chapter, we extend the microeconomic models developed in Chapter 5 to encompass government regulation of the multinational enterprise. This provides us with a theory of the effects of taxation on transfer pricing. First, we develop a variety of models that show how various types of taxes affect MNE behaviour. We examine the impact of trade taxes on the MNE, looking at tariffs as an example. We then turn to pure profit taxes and corporate income taxes (CITs), and a model that includes both tariffs and CITs.

In all these models we are looking at the external motivations for transfer price manipulation, that is, for over- or underinvoicing intrafirm transfers within the MNE. There are at least two ways through which governments can lessen the incentive to manipulate transfer prices for tax purposes. First, the tax authority can impose penalties for transfer price over- or undervaluations, as the U.S. government has recently introduced in its section 6662 regulation (see Chapter 9 for more details). Second, governments could move to a formula apportionment (unitary taxation) system whereby MNEs are taxed on their worldwide income. This approach has been suggested by many academics (see Chapter 12 for a full discussion) for its ability to reduce the incentive to manipulate transfer prices. We investigate the implications of penalty regulations and unitary taxation for transfer price manipulation in the latter part of this chapter.

Chapter 7 tells the other half of this story: the empirical work on taxation and multinationals. Our review of the empirical work is divided into country, industry, and financial studies. Then we turn to evidence on transfer price manipulation for tax purposes in North America. We discuss the popular view in the United States that foreign MNEs pay little or no U.S. taxes, and the recent debate in Canada over the taxes paid by Canadian MNEs. Finally, we conclude with new Canadian and U.S. evidence on taxes paid by multinationals in North America.

Taxes and Transfer Pricing in Theory

While there are many types of taxes that can affect the MNE's transfer pricing policies, this chapter focuses on the three most common forms that have been reviewed in the literature: trade taxes, taxes on pure profits,[1] and corporate income taxes (including withholding taxes). We are particularly interested in the effects of corporate income taxes on MNE transfer pricing policies and resource allocation decisions.[2] If one or more governments levy taxes or tariffs on intrafirm trade, three questions arise:

1. How do taxes and/or tariffs affect the transfer price?
2. Does the MNE use the tax-and-tariff-adjusted transfer price as the relevant price for its internal decisions (i.e., the firm keeps one set of books) or does it choose a transfer price for tax and tariff purposes separate from the price for output allocation (the firm keeps two sets of books)?
3. If the tax-and-tariff-adjusted transfer price is used by the MNE for both internal and external purposes, what impact does the tax/tariff have on MNE real activity?

We address these questions in the context of tariffs, a pure profits tax, and a corporate income tax. Initially, we assume that the price the MNE sets for tax/tariff purposes (the money transfer price **w**) is the same price used for resource allocation purposes within the MNE (the profit-maximizing transfer price **p***), so that the multinational keeps one set of books. We later distinguish between these transfer prices and the efficient transfer price (λ) the MNE would set in the absence of taxes and tariffs. We also introduce a new price $\hat{\mathbf{W}}$, which represents the regulated transfer price imposed by the authorities on the MNE. (See Box 5.1 in Chapter 5.)

Transfer Pricing and Trade Taxes

In this section we focus on efficient transfer pricing in response to a tariff on intrafirm trade flows between the parent and the subsidiary.[3] There are other cases which we could investigate – for example, export tax/subsidies,[4] quotas and voluntary export restraints,[5] anti-dumping and countervailing duties, rules of origin in a free trade area, and foreign exchange constraints[6] – but these are variations on the same theme: how the MNE reacts to trade barriers. Since our purpose in this book is to focus on the corporate income tax, we examine only the simplest trade barrier (the tariff) below and leave the more sophisticated cases for later work.

The results in the tariff case depend critically on the form of the tariff, whether it is specific (i.e., per unit) or ad valorem (based on the price). In the former case, the transfer price per se has no direct impact on the tariff revenue, whereas in the latter case a lower transfer price immediately reduces the tariff revenue per unit. The opportunities for transfer price manipulation are therefore greater in the latter case than in the former. We investigate each case below.

A Specific Tariff on Intrafirm Trade

Assume the importing country levies a specific tariff V on imports so that the total tariff revenue generated is $(\mathbf{p} + V)X$, where **p** is the initial transfer price and X is the volume of intrafirm trade. As in Chapter 5, we assume firm 1 is the exporter and firm 2 the importer. We assume the MNE is horizontally integrated, so the tariff is on imports of the finished good. If π is the pre-tariff profit function of the MNE, let π^* be the after-tariff profit function. Then the net-of-tariff profit function of the MNE is:

$$\pi^* = [R_1(Y_1) + \mathbf{p}X - C_1(Q_1)] + [R_2(Y_2) - C_2(Q_2) - (\mathbf{p} + V)X] \qquad (1)$$

where $X = Q_1 - Y_1 = Y_2 - Q_2$. Adding in the constraint that all output must be sold, and rearranging (1), we have the MNE's objective function:

$$\begin{aligned}\pi^* = {} & R_1(Y_1) - C_1(Q_1) + R_2(Y_2) - C_2(Q_2) - VX \\ & + \lambda_1[Q_1 - Y_1 - X] + \lambda_2[Y_2 - Q_2 - X]\end{aligned} \qquad (2)$$

We differentiate the MNE's total profit function with respect to Y_i, Q_i, and X to determine the optimal amounts of output, sales, and intrafirm trade. This process leads to the following profit-maximizing condition:

$$MR_1 = MC_1 = MR_2 - V = MC_2 - V = \lambda_1 = -\lambda_2 - V$$

or, alternatively,

$$MR_1 + V = MC_1 + V = MR_2 = MC_2 = \lambda_1 + V = -\lambda_2 \qquad (3)$$

We can proxy the direct impact of the specific tariff on the MNE's profits by partially differentiating the objective function (equation 2) with respect to the tariff. The *envelope theorem* tells us that the direct impact of a small partial equilibrium change can be proxied by the partial differential of the objective function, holding all other variables at their equilibrium levels.[7] Partially differentiating (2) with respect to V, we have:

$$\partial\pi^*\backslash\partial V = -X < 0 \qquad (4)$$

Equation (4) shows that the specific tariff reduces MNE profits. Note that the transfer price does not appear in (4); therefore the firm cannot use under-invoicing to avoid the tariff.

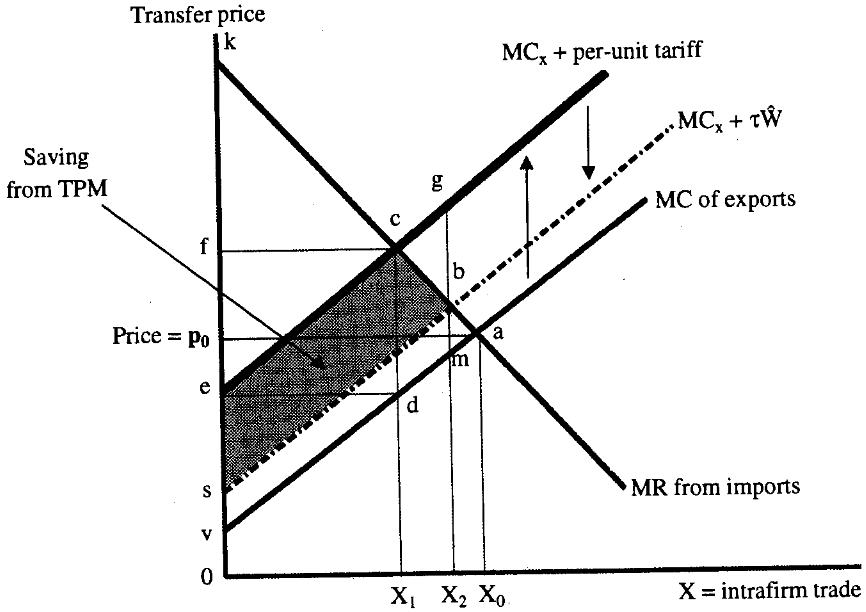
We illustrate the case of a specific tariff, and, later, of an ad valorem tariff, in Figure 6.1. This figure shows a horizontally integrated MNE engaged in intrafirm trade (X). Comparing this figure with the three diagrams in Figure 5.1 in Chapter 5, Figure 6.1 reproduces only the middle diagram from Figure 5.1, for simplicity. The reader might usefully reconstruct the analysis below for all three diagrams, in order to see more clearly the impact of the tariff on the MNE's choice of output, sales, and consumer prices in each country. For our purposes, it is sufficient to focus on the impact on intrafirm trade, i.e., on the middle diagram in Figure 5.1.

In the pre-tariff situation the MNE maximizes profits by equating the marginal revenue from imports (MR_x) to the marginal cost of exports (MC_x) as shown by point a. The initial volume of trade is X_0 and transfer price is $\mathbf{p_0}$, the shadow price of intrafirm trade. The profits of the multinational can be proxied by the area under the MR_x curve and over the MC_x curve – that is, by the triangular area vak, with all profits below the transfer price going to the exporter (area $v\mathbf{p_0}a$) and over the transfer price to the importer (area $\mathbf{p_0}ak$).

The specific tariff raises the costs of intrafirm trade by the per-unit amount of the tariff. We can treat this either as a fall in the returns from importing to firm 2, or as a rise in the cost of exporting for firm 1. In Figure 6.1 this is shown as a parallel shift upwards by the distance ab in the marginal cost of exports function, causing the volume of trade to contract from X_0 to X_1. The MNE pays tariff costs equal to the distance cd times X_1; by construction, this equals the parallelogram vdce. The net-of-tariff profits of the MNE are represented by the area eck. Thus, comparing the pre- and post-tariff situations, we see that the MNE's profit falls by the area vace, which is composed of two parts: (1) the tariff cost vdce, and (2) a real resource cost, known as the ***deadweight loss***, equal to triangle acd.[8]

Since the terms involving **p have disappeared from (2) and (4), this means**

FIGURE 6.1
The Incentives for Transfer Price Manipulation under a Tariff Barrier



that the MNE's choice of transfer price has no effect on its profit maximum or efficient resource allocation. The multinational cannot avoid the specific tariff by adjusting the transfer price since the tariff revenue depends only on the volume of trade, not its value.[9] Thus the efficient transfer price is the shadow price λ_1, as before – that is, the marginal cost of the exporting firm MC_1.

The specific tariff, however, like a quota, can be circumvented to some extent by the MNE shifting to higher value-added activities and thus reducing the effective rate of protection (this is also known as *shipping the good apples out*). That is, if a specific tariff of $5.00 is due on a $50.00 item, the effective tariff rate is 10 per cent; on a $100.00 item, however, the rate falls to 5 per cent. Thus shifting to higher-priced items lowers the effective tariff rate. Multinationals, when faced with a specific tariff that applies equally to higher- and lower-priced items, can reduce the effective tax rate by shipping the good apples (the higher-valued items) out.

MNEs also have two other advantages not available to non-multinationals. First is the ability to ship the product via a country which does not face country **2's tariff; this is known as *trans-shipment*.** If the United States levies a tariff on imports from Japan but not from Thailand, by passing title to a Thai affiliate and shipping the product through Thailand, the MNE may be able to avoid the U.S. tariff (depending on the vigilance of the U.S. customs authorities). Second, the MNE could shift the location where the product is made, or at least the last stage of production before export to the country imposing the tariff, to a country that does not face the tariff. This is known as *investment shunting*. An example would be shifting the last stage of assembly from Japan to Thailand before final export to the United States. Neither of these responses – trans-shipment or investment shunting – directly involves transfer price manipulation per se, but both are methods more easily used by multinationals than by domestic firms.

An Ad Valorem Tariff on Intrafirm Trade

The case of an ad valorem tariff is different since the revenue raised by the customs authorities does depend on the value of the imported product. Assume the importing country, country 2, levies an ad valorem tariff τ on imports so that the total tariff revenue generated is $(1 + \tau)\mathbf{p}X$. The objective function of the MNE, its net-of-tariff profit function with the Lagrangian constraint that all output must be sold, is:

$$\pi^* = [R_1(Y_1) + \mathbf{p}X - C_1(Q_1)] + [R_2(Y_2) - C_2(Q_2) - (1+\tau)\mathbf{p}X] + \lambda_1[Q_1 - Y_1 - X] + \lambda_2[Y_2 - Q_2 - X]$$

which reduces to:

$$\pi^* = R_1(Y_1) - C_1(Q_1) + R_2(Y_2) - C_2(Q_2) - \tau\mathbf{p}X + \lambda_1[Q_1 - Y_1 - X] + \lambda_2[Y_2 - Q_2 - X] \quad (5)$$

First note that the transfer price **p** does remain in the MNE's objective function (5), indicating that for the first time in these models the profit-maximizing transfer price is determinate. The first-order condition for a profit maximum is:

$$MR_1 = MC_1 = MR_2 - \tau\mathbf{p} = MC_2 - \tau\mathbf{p} = \lambda_1 = -\lambda_2 - \tau\mathbf{p}$$

which can be rewritten as:

$$MC_x = MR_x - \tau\mathbf{p} \quad (6)$$

Equation (6) says that the marginal cost of the exporter MC_x should be set equal to the marginal revenue of the importer MR_x minus the per-unit tariff $\tau\mathbf{p}$; thus

the MR_x curve shifts down by the per-unit tariff. Alternatively, instead of shifting firm 2's MR_x curve down, we could have shifted firm 1's MC_x curve up by the tariff. In either case, the MNE reduces its volume of intrafirm trade, and this has effects on the volumes of output and sales in each country as well as on product prices.

The partial equilibrium impact of the tariff on the MNE's profit function can be found by differentiating equation (5) with respect to τ, and using the envelope theorem. We find:

$$\partial\pi^*\backslash\partial\tau = -\mathbf{p}X < 0 \quad (7)$$

Equation (8) says that the tariff has a negative impact on MNE profits, but that this impact can be reduced by underinvoicing the transfer price. Thus an ad valorem tariff can be manipulated through the MNE's transfer pricing policies, even though a specific tariff cannot.

The profit-maximizing transfer price **p*** is found by partially differentiating the MNE's objective function (equation [5]) with respect to **p** and using the envelope theorem, to give us:

$$\partial\pi^*\backslash\partial\mathbf{p} = -\tau X < 0 \quad (8)$$

so **p*** should be set as low as possible to avoid the tariff. Hence, the MNE will underinvoice its exports to affiliates located in countries that levy tariffs on imports. Underinvoicing exports reduces the MNE's tariff burden, effectively shifting the $MC_x + \tau\mathbf{p}$ curve back down towards its initial position. Since the tariff applies only to one-way trade (i.e., on imports) the transfer price should be reduced to zero to completely avoid the tariff.

Even in the absence of income taxes, there are three possible lower bounds to the transfer price. First, if customs officials do not regulate the transfer price, the lower bound on **p*** is zero, eliminating the tariff revenues. Second, as the transfer price is reduced, the exporter's profits on intrafirm trade also fall. Assuming the MNE must declare nonzero profits in each country, a third possible lower boundary is therefore the transfer price that makes firm 1's profits zero (if the exporter sells only within the MNE) or firm 1's profits on intrafirm trade zero (where the firm sells both inside and outside the MNE). Third, the customs authorities are likely to put a floor on the acceptable transfer price based on their assessment of fair market value for the imported good. We therefore assume that the *regulated transfer price* $\hat{\mathbf{W}}$ is positive.

Another way to show that the profit maximizing transfer price **p*** is less than the shadow price $\lambda = MC_1$ in the no-tax situation (so underinvoicing occurs) is to look at the profit on the last unit of intrafirm trade. The profit on the marginal unit is measured by the marginal net revenue to the importer minus the marginal cost to the exporter or, using equation (8), the marginal profit on intrafirm trade is:

$$MR_2 - [MC_1 + \tau \mathbf{p}] = MR_2 - MC_1(1 + \tau \mathbf{p}/MC_1)$$

which we can rewrite as:

$$MR_x = MC_x(1 + \tau \mathbf{p}/MC_1) \qquad (9)$$

If the MNE ignores the tariff and sets $\mathbf{p} = MC_1$, the full burden of the tariff falls on the enterprise. On the other hand, if $\mathbf{p}$ is set below MC_1, the marginal cost of exports falls and marginal profit rises. Underinvoicing is therefore unambiguously profitable. Since the customs authorities know that ad valorem tariffs can be avoided through underinvoicing, most governments have rules in place setting a lower bound on an acceptable price for intrafirm traded goods. This lower bound transfer price is the regulated price $\hat{\mathbf{W}}$.

These results are illustrated in Figure 6.1 where the ad valorem tariff can, to a first approximation, be shown as an upward shift in the MC_x curve to become the $MC_x + \tau\mathbf{p}$ curve.[10] If the MNE keeps its transfer price at the shadow price of intrafirm trade ($\mathbf{p} = MC_1 = \lambda$) the equilibrium is at point c with trade falling to X_1, as in the specific tariff case.

However, if the MNE can underinvoice its imports it would prefer to set as low a transfer price as possible. Assume that the customs authorities set the lower bound at $0 < \hat{\mathbf{W}} < \mathbf{p}^0$. We show the underinvoicing as a parallel downward shift in the $MC_x + \tau\mathbf{p}$ curve to the new curve $MC_x + \tau\hat{\mathbf{W}}$. The new equilibrium is at point b with a volume of trade of X_2. The after-tariff profit of the MNE is now the triangle sbk. After-tariff profit has risen by the shaded area secb. This area represents the net impact on MNE profits of (1) underinvoicing and thus saving tariff costs equal to the distance bg times the volume of trade X_2, which by construction equals the parallelogram segb; and (2) the drop in profits due to the misallocation in resources caused by the expansion in X, equal to the triangle gbc. The gain from underinvoicing clearly exceeds the loss from misallocation of resources, for a net increase in MNE net-of-tariff profit equal to the area secb; thus underinvoicing raises net-of-tariff global profits.

This concludes our discussion of tariffs. We now turn to modelling a tax on pure profits. We discuss the case of the combined impact of taxes and tariffs later in the chapter.

Transfer Pricing and a Tax on Pure Profits

In this section, we model the impact of a pure profits tax on the MNE's resource allocation and transfer pricing decisions.[11] There are two general cases to investigate:

1. Only the home country taxes profits, either on a source basis (only the parent firm's profits are taxed) or on a residence basis (MNE worldwide profits are taxed); the host country does not tax MNE profits. In this case, we have to consider the impact of only one country's tax policy on the MNE's transfer pricing policy choices.
2. Both the home and host countries tax pure profits; the host country taxes on a source basis, while the home country can use either source- or residence-based taxes. In this case, it is the combined impact of the two taxes that affects the MNE's transfer prices.

In the second case (both countries tax profits) we need only look at the home country's choice of residence or source principle since the host country follows the source principle.[12] Where both countries follow the source principle, they do not tax profits earned in the other country. For the home country this is equivalent to exempting foreign-source income from taxation (as Canada does with active business income). Since differences in tax rates are still likely to occur because both governments are levying taxes that are likely to be at different rates, the MNE's behaviour should change so as to take the tax rate differential into account.

Where both countries tax profits and the home country follows the residence principle, its treatment of taxes paid to the host country will be important. We assume the standard practice – that is, that the home country taxes foreign source income as repatriated (i.e., tax deferral is allowed), but gives a foreign tax credit up to the level of the home country rate. This is the practice in the United States. In this case we must consider the impact of deferral on the gap between the tax rates, and deferral becomes one of the choices facing the MNE.

We start with the case of one country, the home country, taxing MNE profits. In all cases we assume pre-tax profit is π and after-tax profit is π^*.

The Home Country Taxes Pure Profits

Assume that we have a vertically integrated MNE in which all the output of the primary division, firm 1, is sold to the downstream division, firm 2. Firm 2 is the parent firm and firm 1 is a wholly owned subsidiary. The home country taxes pure profits and follows either the source principle (only the parent's profits are taxed) or the residence principle (the MNE's worldwide profits are taxed).

The Source Principle

Assume that the home country taxes pure profits at the rate t_2 but follows the source principle so that affiliate profits are not taxed. The after-tax profit function for the MNE is:

$$\pi^* = (1 - t_2)[R_2(Y_2) - C_2(Q_2) - \mathbf{p}Y_1] + [\mathbf{p}Y_1 - C_1(Q_1)] \quad (10)$$

where $Y_1 = Q_1 = Q_2$. Rearranging, substituting Q_1 for Y_1, and adding in the Lagrangian constraints, we have the MNE's objective function:

$$\pi^* = (1 - t_2)[R_2(Y_2) - C_2(Q_2)] - C_1(Q_1) + t_2\mathbf{p}Q_1 + \lambda[Q_2 - Q_1] + \phi[Q_2 - Y_2] \quad (11)$$

Now differentiating (11) with respect to Y_2, Q_1, and Q_2, and rearranging, the MNE's after-tax profits are maximized where:

$$MC_1 + t_2\mathbf{p} = \lambda \quad (12)$$

$$(1 - t_2)MR_2 = (1 - t_2)\,MC_2 + MC_1 - t_2\mathbf{p} = \phi \quad (13)$$

or, rewriting (13) in terms of the net marginal revenue of the importing firm, we have

$$(1 - t_2)[MR_2 - MC_2] = (1 - t_2)NMR_2 = MC_1 - t_2\mathbf{p} = \lambda \quad (14)$$

Equation (12) shows the tax-adjusted shadow price on intrafirm trade; note that the transfer price **p** now affects the shadow price λ and thus resource allocation decisions (more on this below). Equation (14) says that the MNE should equate the net-of-tax marginal profits of the importing firm, $(1 - t_2)NMR_2$, to the marginal costs of the exporter *minus* the tax saving because imports are a tax deductible expense for the parent firm, or, $MC_1 - t_2\mathbf{p}$.

The profit-maximizing transfer price **p*** is found by partially differentiating the MNE's objective function, equation (11), and using the envelope theorem:

$$\partial\pi^*\backslash\partial\mathbf{p} = t_2Q_1 > 0 \quad (15)$$

Thus the profit-maximizing transfer price **p*** should be set high so that profits are shifted to the host country firm where there is no profits tax. The MNE should therefore overinvoice its exports to high-taxed affiliates in order to shift profits out to lower-taxed locations.

Another way to see this is as follows. If there are no taxes or tariffs we know that profits are maximized where $NMR_2 = MC_1 = \mathbf{p} = \lambda$ (see Chapter 5).[13] Suppose the MNE sets $\mathbf{p} = MC_1 = \lambda$, the original shadow price in the no-tax base case. Then equation (14) can be rewritten as:

$$(1 - t_2)(MR_2 - MC_2) = MC_1 - t_2MC_1 = (1 - t_2)MC_1 \quad (16)$$

or, simplifying, as:

$$MR_2 - MC_2 = NMR_2 = MC_1 \quad (17)$$

which is the no-tax equilibrium condition. That is, if the MNE does not alter its transfer price in response to the profits tax, leaving the price equal to the pre-tax shadow price λ, the profit tax does not affect the output, sales, or intrafirm trade decisions of the firm. The firm simply pays a tax equal to $t_2\pi_2$.

If, however, the MNE sets a higher transfer price than the shadow price so that **p** exceeds MC_1, the after-tax profit on the marginal unit is positive. We prove this as follows. The marginal profit is the difference between the marginal return to the importer $(1 - t_2)NMR_2$ minus the marginal cost to the exporter $MC_1 - t_2\mathbf{p}$, or:

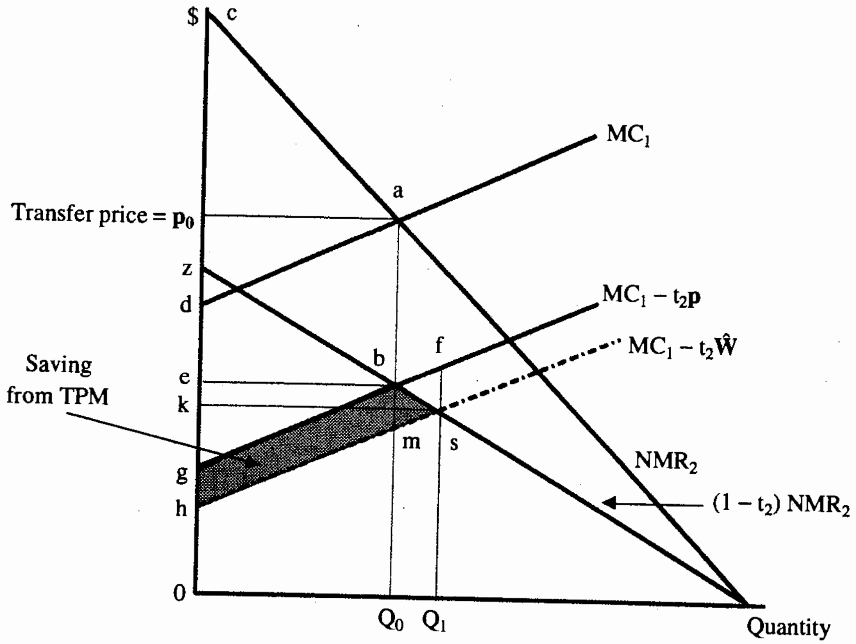
$$(1 - t_2)NMR_2 - [MC_1 - t_2\mathbf{p}] = (1 - t_2)\,NMR_2 - MC_1(1 - t_2\mathbf{p}/MC_1) > 0 \text{ if } \mathbf{p} > MC_1 \quad (18)$$

Hence the MNE has an incentive to *overinvoice* items that are deductible in the high-tax country, and vice versa to *underinvoice* sources of income in that location. Where the MNE changes its overall objective function, shifting from maximizing before-tax income to maximizing after-tax income, the resulting transfer price will change whenever differences in tax rates exist across plant locations. This case is illustrated in Figure 6.2, which is based on the right-hand graph in Figure 5.3 (a vertically integrated MNE).

In the initial situation, the MNE equates MC_1 to NMR_2 at point a, producing Q_0. Pre-tax profits are represented by the area under the NMR_2 curve and over the MC_1 curve, area dac. Let $\mathbf{p}^0$ be the original transfer price. These profits are split between the buying and selling divisions according to the transfer price, with the seller receiving the area dap^0 and the buyer area $\mathbf{p}^0$ac.

The pure profits tax has an *ad valorem tax effect*, which reduces the profit-

FIGURE 6.2
The Incentives for Transfer Price Manipulation under a Pure Profits Tax



ability of the importing division, causing the NMR_2 curve to rotate down to $(1 - t_2)NMR_2$. The profits tax also affects the returns on intrafirm trade since imports are a tax-deductible cost; this effect is shown as a downward shift in MC_1 to the new curve $MC_1 - t_2\mathbf{p}$. This effect is called the *specific tax effect*. If we assume the MNE initially keeps $\mathbf{p} = \lambda = MC_1$, the new equilibrium is at point b, directly below point a. The two effects offset each other and output returns to the pre-tax level Q_0. The MNE simply absorbs the tax costs, equal to the distance ab times Q_0, which, through construction, equals the parallelogram gdab. After-tax profits of the MNE therefore fall to the triangle gbz (i.e., the area under the after-tax NMR_2 curve and over the after-tax MC_1 curve).

However, if the MNE overinvoices and sets $\mathbf{p} > MC_1$, the MC_1 curve shifts even further downwards as the tax deduction for intrafirm imports increases. We assume the tax authorities in country 2 are aware of this and set a floor on the acceptable transfer price of $\hat{\mathbf{W}}$. The new curve is therefore $MC_1 - t_2\hat{\mathbf{W}}$ so that the new equilibrium is at point s with a higher output level Q_1. The new net-of-tax profit of the MNE is the area hzs, which is larger than the old after-tax profit of gbz by the area hgbs. The gain to the MNE from transfer price

manipulation can be calculated as follows. First, overinvoicing saves on tax costs equal to the distance sf times the volume of intrafirm trade; that is, the tax saving is the parallelogram hgfs. Second, the expansion in trade causes a misallocation in resources for the MNE that lowers profit by the triangle bfs. The net impact of these two factors is a rise in after-tax profits equal to the area hgbs. It is therefore profitable for the MNE to overinvoice tax-deductible items like intrafirm imports, and, where possible, to underinvoice taxable items such as revenues from intrafirm transactions.[14]

The Residence Principle

If the home country follows the residence principle, firm 1 is taxed on its worldwide income, that is, on total MNE profits wherever earned. If country 1 taxes on an accrual basis (i.e., tax deferral is not permissible) then all profits are taxed at the same rate t_2 and the MNE cannot avoid the tax through transfer price manipulation. This is clear from looking at a revised version of equation (10), where t_2 now applies to all MNE profits:

$$\begin{aligned}\pi^* &= (1 - t_2)[R_2(Y_2) - C_2(Q_2) - \mathbf{p}Y_1] + (1 - t_2)[\mathbf{p}Y_1 - C_1(Q_1)] \\ &= (1 - t_2)[R_2(Y_2) - C_2(Q_2) - C_1(Q_1)]\end{aligned} \quad (19)$$

From the rearrangement of equation (19) it is clear that the first-order condition for a profit maximum is the same as in the no-tax situation, $MR_1 = MC_1 = MR_2 = MC_2$, and that the profit-maximizing and money transfer prices have no role to play in this model. That is, the MNE cannot avoid the tax in the same way in which a monopoly firm cannot avoid a pure profit tax. The multinational simply absorbs the cost of the tax.

On the other hand, if the taxing authority allows deferral of the tax for profits retained in the host country, the situation changes. In this case the MNE has an incentive to retain profits abroad. Assume the MNE retains the fraction $1 - \beta$ of subsidiary profits in the host country and remits β to the parent firm. Rewriting (19) to allow for deferral we have:

$$\begin{aligned}\pi^* = &(1 - t_2)[R_2(Y_2) - C_2(Q_2) - \mathbf{p}Y_1] \\ &+ (1 - \beta t_2)[\mathbf{p}Y_1 - C_1(Q_1)]\end{aligned} \quad (20)$$

Rearranging and adding in the Lagrangian constraints we have the MNE's objective function:

$$\begin{aligned}\pi^* = &(1 - t_2)[R_2(Y_2) - C_2(Q_2)] - (1 - \beta t_2)C_1(Q_1) + t_2(1 - \beta)\mathbf{p}Q_1 \\ &+ \lambda[Q_2 - Q_1] + \phi[Q_2 - Y_2]\end{aligned} \quad (21)$$

If $\beta = 1$ (full repatriation), then equation (21) collapses to equation (19); that is, the home country taxes on the residence principle and all profits wherever earned are taxed as earned. If $\beta = 0$ (no repatriation), equation (21) collapses to equation (11); that is, the home country taxes on the source principle. The first-order conditions for a profit maximum in the case when $0 < \beta < 1$ are the following:

$$(1 - \beta t_2)MC_1 + t_2(1 - \beta)\mathbf{p} = \lambda \tag{22}$$

$$(1 - t_2)MR_2 = (1 - t_2)MC_2 + (1 - \beta t_2)MC_1 - t_2(1 - \beta)\mathbf{p} = \phi \tag{23}$$

Rearranging (23) we have:

$$(1 - t_2)[MR_2 - MC_2] = (1 - \beta t_2)MC_1 - t_2(1 - \beta)\mathbf{p} \tag{24}$$

That is, the after-tax net marginal revenue to the importing firm should be equated to the after-tax marginal cost of the exporting firm. There are two effects, the ad valorem tax effects on the marginal revenue and cost curves, and the specific tax effect of the transfer price.

Again, the net incentive is to raise the transfer price as long as $\beta < 1$ so that some profits are retained in the host country, where they face a lower effective tax rate.[15] We can see this by partially differentiating equation (21) with respect to **p** and using the envelope theorem:

$$\partial\pi^*\backslash\partial\mathbf{p} = t_2(1 - \beta)Q_1 > 0 \text{ if } \beta < 1 \tag{25}$$

Thus the MNE should overinvoice transfers into high-tax locations in order to shift profits to lower-taxed affiliates.[16]

Both Countries Tax Pure Profits

The Source Principle

Now assume that the host country also taxes the MNE's pure profits. If the home country continues to follow the source principle, both countries tax only profits earned within their borders. In this case the MNE's objective function becomes:

$$\pi^* = (1 - t_2)[R_2(Y_2) - C_2(Q_2) - \mathbf{p}Y_1] + (1 - t_1)[\mathbf{p}Y_1 - C_1(Q_1)] \tag{26}$$

where $Y_1 = Q_1 = Q_2$. Rearranging and adding in the Lagrangian constraints, we have the MNE's objective function:

$$\pi^* = (1 - t_2)[R_2(Y_2) - C_2(Q_2)] - (1 - t_1)C_1(Q_1) + (t_2 - t_1)\mathbf{p}Q_1 + \lambda[Q_2 - Q_1] + \phi[Q_2 - Y_2] \tag{27}$$

The first-order conditions for a profit maximum are:

$$(1 - t_1)MC_1 - (t_2 - t_1)\mathbf{p} = \lambda \tag{28}$$

$$(1 - t_2)MR_2 = (1 - t_2)MC_2 + (1 - t_1)MC_1 - (t_2 - t_1)\mathbf{p} = \phi \tag{29}$$

Again, the MNE compares the after-tax returns to the importer, firm 2, to the after-tax costs of the exporter, firm 1. The profit taxes have two effects, ad valorem and specific. The profit-maximizing transfer price is determined by:

$$\partial\pi^*\backslash\partial\mathbf{p} = (t_2 - t_1)Q_1 > 0 \text{ as } t_2 > t_1 \quad < 0 \text{ as } t_2 < t_1 \tag{30}$$

That is, if the home tax rate (t_2) is higher than the host rate (t_1), MNE profits are increased if the transfer price is set high; if t_2 is less than t_1 the transfer price should be set low. In either case, the profit-maximizing price **p*** is set so that profits are shifted to the lower-taxed affiliate.

The Residence Principle

In our last case we assume both countries tax pure profits on the residence principle, so that t_2 applies to MNE profits as a whole while t_1 applies only to host country profits. We assume as before that the MNE remits the fraction β of its subsidiary profits to the parent firm. This allows us to model three cases at the same time since setting β to different levels is equivalent to different tax regimes in the home country. We identify these cases as: (1) home country taxation on the residence principle with no tax deferral so $\beta = 1$, (2) home country taxation on a source basis so $\beta = 0$, and (3) home country taxation on a residence basis with deferral so $0 < \beta < 0$.

We assume that the home country gives a foreign tax credit up to the level of the home tax rate for any taxes the MNE pays to the host country. We explain the calculation as follows. π_1 is the subsidiary's total before-tax profit on which taxes of $t_1\pi_1$ are paid to the host country, leaving $(1 - t_1)\pi_1$ in after-tax profits. Of this amount, $\beta(1 - t_1)\pi_1$ is remitted to the home country while $(1 - \beta)(1 - t_1)\pi_1$ represents the subsidiary's retained earnings. When the repatriated earnings (the dividends) are remitted to the parent, firm 1 brings them into its income and grosses them up by the amount of foreign taxes paid so that they are now on a pre-tax basis, as follows: $\beta(1 - t_1)\pi_1/(1 - t_1) = \beta\pi_1$. The government then calculates the home country tax due on the dividends as $t_2\beta\pi_1$ and gives a credit for the foreign taxes paid, so that the net effect of the residence principle with deferral and a foreign tax credit is: $t_2\beta\pi_1 - t_1\,\beta\pi_1 = (t_2 - t_1)\,\beta\pi_1$ where, if $t_1 > t_2$, no additional tax is due in the home country.

In effect, the MNE pays t_1 in tax on all of its foreign subsidiary's profits and an additional tax equal to the difference between the home and host tax rates on any profit remittances to the parent firm. The effective tax rate on subsidiary profits is therefore $t_1 + \beta(t_2 - t_1)$, which has a floor of t_1 and a ceiling of t_2. The MNE pays the floor if β is zero (no remittances) or if $t_2 < t_1$ (the home tax rate is lower than the host rate); the MNE pays the ceiling if $\beta = 1$ (no deferral) and $t_2 > t_1$ (the home rate is higher than the host rate). Note that either condition can hold to satisfy the floor but both conditions must hold to satisfy the ceiling. We assume in what follows that the home tax rate is higher than the host rate since if the host rate is higher, no additional taxes are due when the subsidiary remits its profits.

Putting this altogether, the profit function for the MNE is:

$$\begin{aligned}\pi^* &= (1 - t_2)[R_2(Y_2) - C_2(Q_2) - \mathbf{p}Y_1] \\ &\quad + [1 - t_1 - \beta(t_2 - t_1)][\mathbf{p}Y_1 - C_1(Q_1)]\end{aligned} \tag{31}$$

Rearranging and adding in the Lagrangian constraints, we have the MNE's objective function:

$$\begin{aligned}\pi^* &= (1 - t_2)[R_2(Y_2) - C_2(Q_2)] - [1 - t_1 - \beta(t_2 - t_1)]C_1(Q_1) \\ &\quad + (t_2 - t_1)(1 - \beta)\mathbf{p}Q_1 + \lambda[Q_2 - Q_1] + \phi[Q_2 - Y_2]\end{aligned} \tag{32}$$

If $\beta = 1$ (full repatriation), then equation (32) collapses to equation (19); that is, the home country taxes on the residence principle and all profits wherever earned are taxed as earned. If $\beta = 0$ (no repatriation), equation (32) collapses to equation (27), in which both countries tax on the source principle. The first-order conditions for a profit maximum when $0 < \beta < 1$ are the following:

$$[1 - t_1 - \beta(t_2 - t_1)]MC_1 + (t_2 - t_1)(1 - \beta)\mathbf{p} = \lambda \tag{33}$$

$$\begin{aligned}(1 - t_2)MR_2 &= (1 - t_2)MC_2 + [1 - t_1 - \beta(t_2 - t_1)]MC_1 \\ &\quad - (t_2 - t_1)(1 - \beta)\mathbf{p} = \phi\end{aligned} \tag{34}$$

That is, the after-tax net marginal revenue to the importing firm should be equated to the after-tax marginal cost of the exporting firm. Partially differentiating equation (32) with respect to **p** and using the envelope theorem:

$$\partial\pi^*\backslash\partial\mathbf{p} = (t_2 - t_1)(1 - \beta)Q_1$$

$$\text{which is} \quad \begin{cases} > 0 \text{ if } t_2 > t_1 \text{ and } \beta < 1 \\ = 0 \text{ if } t_2 = t_1 \text{ or if } \beta = 1 \text{ and } t_2 > t_1 \\ < 0 \text{ if } t_2 < t_1 \end{cases} \tag{35}$$

Therefore the MNE has an incentive to overinvoice its exports to the home country, shifting profits to the subsidiary, if the effective tax rate on remitted dividends is lower in the host country. The profit-maximizing transfer price in this case should be set at its upper bound, that is, $\mathbf{p}^* = \max \hat{\mathbf{W}}$. Where the tax rates are the same or where the home country rate is higher and all profits are taxed as earned, the profit-maximizing transfer price is indeterminate. And where the host tax rate is higher, $\mathbf{p}^*$ should be set at the lower bound determined by the tax authorities, that is, $\mathbf{p}^* = \min \hat{\mathbf{W}}$. These regulatory bounds will be based on the tax auditor's estimate of the arm's length price, based on the application of the various transfer pricing methods we described in chapters 1 and 5 (e.g., CUP, cost plus, resale price).

Note also that, given the choice, the profit-maximizing deferral rate is zero, that is, no remittances should be made whenever the home rate is higher than the host rate. We can see this by differentiating (32) with respect to β as follows:

$$\partial\pi^*\backslash\partial\beta = -(t_2 - t_1)[\mathbf{p}Q_1 - C_1(Q_1)] \tag{36}$$

The expression in square brackets in (36) is the subsidiary's total pre-tax profit. As long as the home rate is higher and profits are positive, the MNE will reduce its dividend remittances.

Some Related Issues

In this section, we deal with four issues that extend the models we have used so far: (1) whether the MNE uses one or two sets of books, (2) minority shareholders, (3) exchange rate changes, and (4) the combination of taxes and tariffs. The first issue answers the general question of whether the MNE takes external constraints such as taxes and tariffs into account in setting its profit-maximizing transfer price. The last three issues deal with the impacts of common constraints on the MNE's transfer pricing policies.

One or Two Sets of Books?

The issue of whether MNEs use one or more sets of books has been frequently discussed in the transfer pricing literature (Rugman and Eden 1985, Eccles

1985, Cummins et al. 1995, McGuinness 1985). 'Two sets of books' can have different meanings to different readers.

The first meaning is the practical reality in accounting for MNE profits (Cummins et al. 1995). In many countries, such as the United States, MNEs are required to keep two sets of accounting books. U.S. companies are required to publish financial statements that conform to generally accepted accounting principles (GAAP) as set by the Financial Accounting Standards Board (FASB). The published set of books is designed to provide financial information on the firm to potential and actual investors and creditors. At the same time, the IRS has quite different reporting requirements for tax purposes; for example, depreciation rates are different. In other countries, e.g., Germany, for example, one set of books is required for financial and tax purposes. In fact, the MNE may keep several sets of books: one for tax purposes, another for customs valuations, a third for public view by minority shareholders and union groups, and a final one for private allocation decisions. Several sets of books are necessary where different government regulatory authorities require different types of accounting.

The second meaning is quite different; that is, should the MNE keep one set of books for resource allocation decisions, and a second set for tax purposes? The common perception is that MNEs keep several sets of books, and that real decisions and financial decisions are made separately (and by different departments). The evidence, however, is mixed. Surveys of firm pricing policies tend to show MNEs as setting transfer prices without the advice of the MNE's tax department; many studies also document that most MNEs do not have an explicit transfer pricing policy (e.g., Wilson 1993; U.S. Treasury 1988). In addition, keeping different sets of books has real administrative costs, can have create mixed signals for managers, and can have negative impacts on managerial performance where bonus packages are tied to subsidiary returns (Eccles 1985; McGuinness 1985).

Firms that do keep two sets of books and engage in creative financial manoeuvres, in practice, are unlikely to face severe tax penalties in most countries.[17] This encourages transfer price manipulation, particularly of fungible activities such as management fees. Auditing of tax books also has a time dimension as well, since customs valuations are often made years before tax officials may begin a corporate income tax (CIT) investigation, although this is changing as tax and customs authorities, both in Canada and the United States, are now more frequently exchanging transfer pricing information. Firms, especially smaller MNEs, see the probability of a tax audit as small; the audits take place one to three years after the event; and the probability of the dispute's going to court and being settled in favour of the tax authority is very small.[18]

Hence the incentive to use different sets of books to avoid taxes – that is, to keep a *money transfer price* policy different from the firm's *profit-maximizing transfer price* policies – appears clear.

Note, however, that the models we have used above – and most of the economics literature in this area – implicitly assume that the multinational does *not* keep two sets of books; that is, it uses the same transfer price for internal decisions as it uses for tax purposes. Unless the MNE does keep different books for the government authorities and for private allocation decisions, nominal tax rates can and do influence real output, sales, and trade decisions for the integrated firm. Why do the microeconomic models of the MNE assume one set of books when MNEs appear in practice to keep several for different purposes?

We argue that it is more profitable for the MNE to choose one transfer price, a price that takes into account both tariff and tax rates and the impact on resource allocation.[19] If under- or overinvoicing pays in terms of tax relief, the MNE can increase this relief by expanding the volume of intrafirm trade; however, the MNE needs to balance off at the margin the savings from reducing tax penalties against the costs of such misallocated resources. Keeping two or three sets of books also incurs administrative time and effort and increases the risk of state monitoring of MNE transactions.

Let us illustrate why one set of books is more profitable for the MNE than two. Assume that the MNE can keep two sets of books: one in which the transfer price is set at $\mathbf{w}$ (the money or accounting transfer price) for tax/tariff purposes and a second where the transfer price is set at $\lambda = MC_1$ (the shadow transfer price) for resource allocation purposes. We use the simplest tax model – the home country taxes on a source basis, equations (10)–(14) – as our example. In this case the first-order condition (14), substituting $\mathbf{w}$ for $\mathbf{p}$, is:

$$(1 - t_2)(MR_2 - MC_2) = MC_1 - t_2\mathbf{w} \qquad (37)$$

Taxes can be reduced and after-tax profits of the MNE increased through two methods. First, holding the volume of intrafirm trade constant, net profits are higher if the primary firm overinvoices its exports to the downstream parent. We saw this earlier: the MNE benefits if it overinvoices, shifting profits to the lower-taxed subsidiary. The gain to the MNE is $t_2(\mathbf{w} - MC_1)$ per unit of intrafirm trade, where $\mathbf{w} - MC_1$ represents the amount of overinvoicing relative to the shadow transfer price $\lambda = MC_1$ in the no-tax situation. In Figure 6.2, this is equivalent to the tax savings represented by area hgbm.

Second, given that overinvoicing is possible, holding the amount of overinvoicing $(\mathbf{w} - MC_1)$ per unit constant, taxes can be further reduced by expanding the volume of intrafirm trade. Since each unit of intrafirm trade saves the

MNE $t_2(\mathbf{w} - MC_1)$ in taxes, the larger is X the greater total tax savings of $t_2(\mathbf{w} - MC_1)X$. In Figure 6.2 this is shown as the move from Q_0 to Q_1; the additional gain to the MNE is the area mbs. So, raising the volume of intrafirm trade, with its subsequent implications for output and sales in each affiliate, is more profitable than leaving output at its pre-tax levels. Thus, using the same transfer price for resource allocation purposes as for tax/tariff purposes is more profitable than using separate transfer prices. One set of books is preferable! The money transfer price should be the profit-maximizing transfer price, which in this case is the highest possible transfer price allowed by the tax authority in country 2, which is $\hat{\mathbf{W}}$.

Note that this does not mean there are no costs to the MNE from using only one set of books. Choosing a transfer price different from the efficient transfer price $\mathbf{p} = MC_1$ does mean that output and sales are misallocated within the multinational enterprise. Manipulation of the first-order condition clearly shows this:

$$t_2(\mathbf{w} - MC_1) = (1 - t_2)[MC_1 - (MR_2 - MC_2)] \quad (38)$$

The left-hand side of (38) shows the gain to the MNE from overinvoicing, assuming the MNE only uses one set of books; the right-hand side shows the after-tax cost to the MNE from misallocating its resources. If $\mathbf{w} > MC_1$, the firm will choose to set $MC_1 > MR_2 - MC_2$; that is, the MNE will expand output, and thus the volume of intrafirm exports, beyond their initial no-tax levels. The misallocation of resources in Figure 6.2 is shown by the fact that output is expanded beyond the point where $NMR_2 = MC_1$ (point a). Past that point, in pre-tax terms, marginal cost of the exporter exceeds the net marginal revenue of the importer.[20] In post-tax terms, expanding output adds the area mbfs in Figure 6.2 in tax savings, but costs the area bfs in misallocated resources. In net-of-tax terms, the MNE is clearly better off by triangle mbs. Thus it pays to keep one set of books.

In theory then, if not in practice, multinationals maximize their global profits net of taxes and tariffs. In doing so, they find it more profitable to choose transfer prices that are either higher or lower than those the no-tax situation would suggest. Over- or underinvoicing, per unit of intrafirm trade, creates tax savings and higher after-tax profits for the MNE. These tax savings are magnified when the firm increases the volume of intrafirm trade. The MNE will therefore *increase* its trade volume in *both* under- and overinvoicing cases.[21] Thus the money transfer price **w** and the profit-maximizing price **p*** are the same: the MNE keeps one set of books.

Minority Shareholders: A Tax on the Multinational?

Does it make a difference to the MNE if the foreign subsidiary has minority shareholders? The answer is: it depends. Where a host government imposes minority shareholder requirements on the MNE (as many developing countries have done and continue to do as part of their FDI regulations), the firm no longer has complete control over its activities and its affiliate profits must be shared with unrelated foreign parties. In effect, such regulations act as a tax on host country profits earned by the MNE. One would expect the outcome therefore to be similar to that of a pure profits tax levied by the host country on the affiliate's income.

On the other hand, this is only true if minority shareholders are imposed on the MNE. If the enterprise freely enters into a strategic alliance or a joint venture, then presumably the MNE perceives the benefits (access to the partner's capital, knowledge of the local market, and so on) as outweighing the costs (sharing the profits with the foreign partner). In such cases, we would not expect the minority shareholder to be seen as a tax.[22]

In this section, we assume the host government requires the MNE to have minority shareholders, and investigate the reactions of the MNE in terms of transfer pricing. Assume the foreign affiliate in country 2 is partly owned by host country nationals, and the proportion of foreign minority ownership is k, where $0 < k < 1$. Assume the foreign shareholders must be paid the fraction k of the profits declared by the subsidiary in the host country.[23] Then there is a clear incentive to use transfer pricing manipulation to shift profits out of the host country. Assume we have a horizontally integrated multinational in which firm 1, the MNE parent, is exporting a finished good to firm 2, the joint venture, as we modelled in Figure 5.1 and in Figure 6.1. The global profit function for the MNE, taking minority ownership into account, is:

$$\pi^* = [P_1Y_1 + \mathbf{p}X - C_1(Q_1)] + (1 - k)[P_2Y_2 - \mathbf{p}X - C_2(Q_2)] \quad (39)$$

which can be rearranged as the objective function:

$$\begin{aligned} \pi^* = {} & P_1Y_1 - C_1(Q_1) + (1 - k)[P_2Y_2 - C_2(Q_2)] \\ & + k\mathbf{p}X + \lambda_1[Q_1 - Y_1 - X] + \lambda_2[Y_2 - Q_2 - X] \end{aligned} \quad (40)$$

from which the profit-maximizing condition is:

$$\mathbf{MR_1 - kp = MC_1 - kp} = (1 - k)MR_2 = (1 - k)MC_2 \quad (41)$$

The presence of minority shareholders therefore has two effects on the MNE. First, it acts like an ad valorem tax on profits earned in the host country, since the MNE only receives 1 − k of the profits. This ad valorem effect tends to reduce the volume of intrafirm trade. On the other hand, there is also a specific tax effect; the **kp** term means that if the MNE charges a high transfer price for parent exports to its affiliates, these costs are deductible in country 2 and thus reduce declared profits in the host country. The profit-maximizing transfer price is positive since:

$$\partial\pi^*\backslash\partial \mathbf{p} = kX > 0 \qquad \text{as long as } k > 0 \tag{42}$$

So there is an incentive to overinvoice subsidiary imports and underinvoice its exports in order to shift profits from the subsidiary to the parent firm. We can see this more clearly by examining the equation for net profit on the marginal unit of intrafirm trade:

$$(1-k)MR_2 - MC_1(1 - \mathbf{kp}/MC_1) \quad \begin{aligned} &> 0 \text{ if } k > 0 \text{ and } \mathbf{p} > MC_1 \\ &< 0 \text{ if } k > 0 \text{ and } \mathbf{p} < MC_1 \\ &= 0 \text{ if } k = 0 \end{aligned} \tag{43}$$

which is positive (negative) if the MNE over- (under-) invoices. If the MNE is free to set its transfer price it will therefore shift profits out of the minority affiliate where possible.

Transfer Pricing and Exchange Rate Changes

Assume first that we have a vertically integrated MNE in which the foreign subsidiary, firm 1, exports a raw material for processing and final sale by its parent firm, firm 2. Let all transactions take place in the home country currency and the exchange rate **e** be used to convert the foreign currency. The exchange rate **e** is the price of the home currency divided by the price of the foreign currency, so the subsidiary's profit π_1 (measured in units of the foreign currency) multiplied by **e** gives us $\mathbf{e}\pi_1$ (measured in units of the home currency). When **e** rises, the home currency depreciates, which means one unit of the home currency buys less of the foreign currency, and the host currency appreciates. When **e** falls, the reverse occurs: the home currency appreciates and the host currency depreciates.

As before, but now with the exchange rate added, the MNE maximizes the overall profit function:

$$\pi^* = \mathbf{e}[\mathbf{p}X - C_1(Q_1)] + [R_2(Y_2) - C_2(Q_2) - \mathbf{ep}X] \tag{44}$$

Rearranging, and adding in the Lagrangian constraints, we have:

$$\pi^* = R_2(Y_2) - C_2(Q_2) - \mathbf{e}C_1(Q_1) + \lambda[Q_2 - Q_1] + \phi[Q_2 - Y_2] \tag{45}$$

It is clear from equation (45) that the exchange rate **e** has no immediate link with the transfer price since the **p** terms have disappeared.

The first-order conditions are basically unchanged since the MNE still sets $NMR_2 = \mathbf{e}MC_1$; however, they are now adjusted for the exchange rate **e**. What the exchange rate does affect is the cost of production in the foreign country. Taking the partial differential of (45) with respect to **e** and using the envelope theorem, we have:

$$\partial\pi^*\backslash\partial\mathbf{e} = -C_1(Q_1) < 0 \tag{46}$$

Therefore increases in **e** (which mean the host currency appreciates relative to the home currency) raise the cost of production in the foreign subsidiary and have a negative impact on MNE profits. Intuitively, if production becomes more expensive abroad because the foreign currency has appreciated relative to the home currency, the MNE will react by curtailing production abroad and importing less. However, the exchange rate shock cannot be avoided through transfer pricing since the money and profit-maximizing transfer prices are not determinate in this model.

To pin down the transfer price, we need to introduce an additional constraint such as a tariff. Let the home country levy a tariff on imports from firm 1.[24] The MNE's objective function is:

$$\pi^* = \mathbf{e}[\mathbf{p}X - C_1(Q_1)] + [R_2(Y_2) - C_2(Q_2) - \mathbf{e}(1+\tau)\mathbf{p}X] \tag{47}$$

In this case, the effect of the exchange rate change is:[25]

$$\partial\pi^*\backslash\partial\mathbf{e} = -C_1(Q_1) - \tau\mathbf{p}X \quad \begin{aligned} &< 0 \text{ if } \tau > 0 \\ &= 0 \text{ if } \tau = 0 \end{aligned} \tag{48}$$

which again is negative if there is a tariff. To the extent that the MNE underinvoices to avoid the tariff (which it would do if possible), the exchange rate shock is also moderated. This is clear from deriving the profit-maximizing transfer price:

$$\partial\pi^*\backslash\partial \mathbf{p} = -\mathbf{e}\tau X < 0 \tag{49}$$

The transfer price **p*** should be set as low as possible to avoid the tariff, as before. The higher the exchange rate, the more tariff revenue is saved from underinvoicing. We can show this by looking at the marginal profit per unit of intrafirm trade, which is:

$$NMR_2 - \mathbf{e}(MC_2 + \tau\mathbf{p}) = NMR_2 - \mathbf{e}MC_2(1 + \tau\mathbf{p}/MC_2) \tag{50}$$

Thus the lower is **p** the greater the marginal profit, and, vice versa, the higher is **e** the lower the marginal profit per unit of intrafirm trade.[26]

Transfer Pricing, Taxes, and Tariffs

The last case we want to look at before turning to a full-blown model of transfer pricing and the corporate income tax is a model that includes both a tariff and a tax on pure profits. Let us take our second model in this chapter (the ad valorem tax) and add in taxes levied by both countries on a source basis. The MNE's objective function becomes:

$$\begin{aligned}\pi^* = {} & (1 - t_1)[R_1(Y_1) + \mathbf{p}X - C_1(Q_1)] \\ & + (1 - t_2)[R_2(Y_2) - C_2(Q_2) - (1 + \tau)\mathbf{p}X]\end{aligned} \tag{51}$$

which, adding in the Lagrangian constraint, reduces to:

$$\begin{aligned}\pi^* = {} & (1 - t_1)[R_1(Y_1) - C_1(Q_1)] + (1 - t_2)[R_2(Y_2) - C_2(Q_2)] \\ & + ((1 - t_1) - (1 - t_2)(1 + \tau))\mathbf{p}X \\ & + \lambda_1[Q_1 - Y_1 - X] + \lambda_2[Y_2 - Q_2 - X]\end{aligned} \tag{52}$$

Leaving the first-order conditions for the reader as an exercise, we turn directly to the profit-maximizing transfer price. Partially differentiating (52) with respect to **p** and using the envelope theorem, we have:

$$\begin{aligned}\partial\pi^*\backslash\partial\mathbf{p} = {} & ((1 - t_1) - (1 - t_2)(1 + \tau))X = [(t_2 - t_1) - r(1 - t_2)]X \\ = {} & (1 - t_2)[(t_2 - t_1)/(1 - t_2) - r]X\end{aligned} \tag{53}$$

which, after some rearranging, says that the MNE should set its transfer price high (low) if the tax differential $(t_2 - t_1)$ exceeds (is less than) the effective tariff rate adjusted for the tax saving $\tau(1 - t_2)$. Thus the MNE trades off the tax differential against the tax-adjusted tariff cost in determining whether the profit-maximizing transfer price should be set high or low. Alternatively, the MNE's pricing policy should depend on whether the tax gap ratio $(t_2 - t_1)/(1 - t_2)$ is greater or less than the tariff τ. If the tax differential exceeds (is less than) the tariff, the MNE should over- (under-) invoice the transfer price. This is the standard theoretical result in the tax transfer pricing literature (Eden 1976, 1985; Horst 1971).

Transfer Pricing and the Corporate Income Tax

We turn now to a more realistic model of MNE transfer pricing. In the section below the corporate income tax (CIT) is modelled as falling both on pure profits and on the return to equity capital. The MNE is faced with an ad valorem tariff, two corporate income taxes, and various withholding taxes on remittances to the parent. This model more closely approximates the problems facing multinationals as they are taxed in the real world.

Setting Up the Model

Assume we have a horizontally integrated multinational enterprise, consisting of a parent firm, firm 1, and a subsidiary, firm 2. The subsidiary produces a finished good and also imports the same good from its parent, depending on relative costs and the transfer price charged on these intrafirm imports, for sale in its domestic market. The subsidiary pays head office fees to cover various services provided by the parent, and also makes dividend payments to the parent.

Each firm produces output, Q_i, for sale locally as Y_i, or for export X, where i = 1, 2. Parent sales are therefore $Y_1 = Q_1 - X$ with revenues $R_1(Y_1)$, while affiliate sales are $Y_2 = Q_2 + X$ with revenues $R_2(Y_2)$. Country 2 levies a tariff at rate τ on imports. Since intrafirm imports are priced at transfer price **p** for a total trade value of **pX**, customs revenues are τ**pX**. We assume the exchange rate between the two countries is **e** and that all variables are measured in the home currency.

Each firm's net profit function is based on its taxable income, defined as its economic profit minus tax-deductible expenses. The initial tax payable is the corporate income tax (CIT) rate times taxable income, from which tax credits are subtracted to determine the actual tax bill. Subtracting the actual tax bill from the economic profit determines the net profit of the firm, π^*_i. We look first at the foreign subsidiary's, and then at its parent's, after-tax profit function.

The net profit function for the *foreign subsidiary* is:

$$\begin{aligned}\pi^*_2 = \mathbf{e}[& (1 - t_2)\{R_2 - (1/e)(1 + \tau)\mathbf{p}X - H\} \\ & - C_{k2} - w_hH - (1 + w_d)D]\end{aligned} \tag{54}$$

where C_{k2} is the net-of-CIT cost of capital. We ignore labour costs for simplicity. There are three variables that complicate this model and that are not present in our earlier tax models: capital costs, head office fees, and dividends. We outline the impact of each variable below.

First, we assume that each firm owns $P_{Ki}K_i$ in physical capital in the form of machinery and equipment. The capital stock depreciates at a uniform rate d, and the MNE must invest to replace and expand $P_{Ki}K_i$.[27] The annual economic cost of capital to the firm is its opportunity cost, that is, its real return plus the depreciation rate. The real return is assumed given (the firm is a price taker in the capital market), and equals r, the going world rate of return. Arbitrage between countries ensures that the rate of return is equal across markets so that $r = r_1 = r_2$.[28] The cost of capital is therefore $(r + d)P_{Ki}K_i$ to each firm in the absence of taxation. This is the *pre-tax cost of capital.*

The MNE will purchase units of capital as long as the economic benefits of an additional unit exceed the costs of an additional unit.[29] The economic benefit to the firm can be measured as the extra output an additional unit of capital can produce (i.e., the marginal product of capital) times the additional revenue earned when that output is sold (i.e., the marginal revenue); this is called the ***marginal revenue product of capital*** (MRP_{Ki}) or the firm's demand for capital. The economic cost of capital is its opportunity cost; this is the price of capital to the firm. In the absence of taxes, the MNE would therefore invest in capital up to the point where MRP_{Ki} equals $(r + d)P_{Ki}$, that is, where the demand and supply curves of capital intersect.

The corporate income tax complicates the firm's estimate of capital costs because interest costs on debt capital are tax deductible, whereas the costs of equity capital are not.[30] The tax deductibility of interest costs reduces the true opportunity cost of capital to the firm by the CIT rate times the leverage ratio, L_i (the ratio of long-term liabilities to long-term liabilities plus equity). The *after-tax cost of capital* to each firm is therefore $C_{ki} = (r + d - t_iL_ii_i)P_{Ki}K_i$.[31]

Second, in terms of head office fees, we assume that the subsidiary is charged for H as its share of head office services (these may be specific services or some share of group services – see Chapter 5), and that H is deductible against the host country's CIT. The subsidiary remits H to the parent firm after paying a withholding tax at rate w_h to the host government, state 2. Third, we treat dividends as a residual payment out of after-tax subsidiary profits; the subsidiary remits its dividends, D, to the parent, net of a withholding tax w_d paid to state 2.

The after-tax profit function of the *parent firm* is also calculated in a manner similar to equation (54), as the CIT times taxable income, minus tax credits. Taxable income equals domestic economic profits plus remittances from the subsidiary (after grossing up the dividends by the host CIT) minus other tax-

deductible expenses. Both the CIT and withholding tax paid by the subsidiary are creditable up to the level of the home country CIT rate. The profit function of the *parent* is therefore:

$$\begin{aligned}\pi^*_1 = {} & (1 - t_1)(R_1 + \mathbf{p}X) - C_{k1} + \mathbf{e}[1 - (t_1 - w_h)]H \\ & + \mathbf{e}[1 - f\{(t_1 - (t_2 + w_d(1 - t_2))/(1 - t_2))\}D\end{aligned} \tag{55}$$

where C_{k1} is the net-of-tax cost of capital to the parent. The variable f must be either zero (if the subsidiary has a surplus of foreign tax credits) or positive (a deficit of credits).

Now we put these two profit functions together to get the MNE's overall profits. We assume the goal of the MNE is to maximize global net profits, $\pi^* = (\pi^*_2 + \pi^*_1)$, subject to the constraints that $\Sigma Q_i = \Sigma Y_i$ (everything is sold) and $r_2 = r_1 = r$ (the net return to capital is everywhere equal, so capital markets are perfect), where π^* is:

$$\begin{aligned}\pi^* = {} & \mathbf{e}[(1 - t_2)R_2 - C_{k2} + (1 - t_1)R_1 - C_{k1} + \mathbf{e}\,[t_2 - t_1]H \\ & + [(1 - t_1) - (1 - t_2)(1 + \tau)]\mathbf{p}X \\ & + \mathbf{e}[-w_d - f\{(t_1 - (t_2 + w_d)(1 - t_2))/(1 - t_2)\}]D\end{aligned} \tag{56}$$

Given this objective function, what should the multinational do?

The Various Transfer Pricing Choices

In this particular model the MNE has several possible transfer pricing alternatives, broadly defined, one for each type of intrafirm flows. First, the standard case is the transfer price on the tangible, the final good imported by the foreign affiliate from the parent firm. As we have seen above, both the transfer price and the volume of trade can be manipulated to shift profits out of the high-tax location. The tax saving, however, if it involves overinvoicing, must be traded off against the extra tariffs paid to the customs authorities. Second, the parent firm provides support services to the subsidiary, for which the affiliate pays head office fees. Third, the subsidiary pays dividends to the parent firm. A fourth possibility, which we have not modelled, is that the parent firm lends financial capital to the subsidiary, so there are intracorporate interest charges to be considered. As a result, the MNE has several degrees of freedom in terms of choosing which flows it wants to manipulate, in what direction, and by what amount. The goal of the tax and customs authorities, therefore, is to prevent these various types of transfer price manipulations. Let us now turn to the MNE's choice of transfer pricing policies.

The real decision variables for the MNE are K_2 and K_1 (the capital stocks and therefore output levels of each firm) and X (the volume of intrafirm trade and therefore, indirectly, sales and output levels of each firm). The financial decision variables for the MNE are H (the price of head office services), D (the amount of repatriated dividends), and **p** (the transfer price on the tangibles transaction). We turn first to the real decision variables.

Differentiating with respect to K_2, K_1, and X, we have the following conditions for maximizing MNE overall profits net of taxes and tariffs. The first two conditions determine the optimal amounts of plant investment and output levels in the two plants:

$$MRP_{Ki}/P_{Ki} = (r + d - t_i L_i i_i)/(1 - t_i) = C_{Gi} \qquad \text{(where i = 1, 2)} \quad (57)$$

This says that the marginal revenue product of capital divided by its price should equal C_{Gi}, the *tax-adjusted or gross cost of capital* per dollar of capital expenditures. Note that although capital arbitrage ensures that the net return, r, is equalized between the firms, the gross costs of capital are unlikely to be equalized, since CIT rates, tax deductions, and credits are likely to differ between countries, and depreciation and leverage ratios to differ between firms.

The third first-order condition determines the optimal volume of intrafirm trade (and hence sales levels):

$$(1 - t_2)[\mathbf{e}MR_2 - (1 + \tau)\mathbf{p}] = (1 - t_1)[MR_1 - \mathbf{p}] \qquad (58)$$

which says that the MNE should balance the subsidiary's net marginal revenue from imports against the parent's net marginal cost of exports. The marginal revenue from imports equals affiliate marginal revenue from domestic sales, $\mathbf{e}MR_2$, net of importing costs, $(1 + \tau)\mathbf{p}$, in after-tax terms. The net marginal cost of exports equals the parent's forgone marginal revenue on local sales, MR_1, minus its marginal earnings from exports, **p**, after tax.

The profit-maximizing financial decisions concerning D and H (what Brean [1985] calls the *fiscal transfer prices*) and profit-maximizing transfer price **p*** are found by partially differentiating the profit function (56) with respect to **p**, D, and H, and using the envelope theorem. First, the optimal level of D depends on whether a surplus or deficit of tax credits applies to dividends. In the deficit case, since $t_1 > t_2$ and $f > 0$, the optimal amount of remitted dividends to the parent firm is determined by:

$$\partial\pi^*/\partial D = \mathbf{e}[t_2 - t_1]/(1 - t_2) < 0 \qquad (59)$$

In the surplus case where $f = 0$, the optimal amount of dividend repatriation is:

$$\partial\pi^*/\partial D = -\mathbf{e}w_d < 0 \qquad (60)$$

Thus the important tax variables influencing dividend repatriation are the statutory CIT rate and the dividend withholding tax rate. In both deficit and surplus of credits cases, the MNE maximizes profits by setting dividends at their lowest possible level. The optimal amount of head office charges, H, is:

$$\partial\pi^*/\partial H = \mathbf{e}(t_2 - t_1) + (\partial\pi^*/\partial D)[\mathbf{e}\alpha\{-(1 - t_2) - w_h\}] \qquad (61)$$

which implies that normally, where t_2 exceeds (is less than) t_1, head office charges should be raised (lowered) since they are tax deductible in the host country and taxable at home. The withholding tax usually has no effect on the optimal amount of H because w_h is so low relative to t_1 that the tax is normally all deductible when head office fees are remitted. The net cost of H to the subsidiary therefore is $\mathbf{e}(1 - t_2 + w_h)H$ and to the parent is $\mathbf{e}(-1 + t_1 - w_h)H$, for a total net cost to the MNE of $\mathbf{e}(t_1 - t_2)H$.

The profit-maximizing transfer price **p*** is determined by:

$$\begin{aligned}\partial\pi^*/\partial \mathbf{p} = {} & [t_2 - (t_1 + \tau(1 - t_2)]X \\ & + (\partial\pi^*/\partial D)[\alpha\{-(1 - t_2)(1 + \tau)X\}]\end{aligned} \qquad (62)$$

The square-bracketed term may be either positive or negative depending on the tax and tariff costs; for example, a higher host country CIT rate than in the home country tends to encourage overinvoicing, whereas host country tariffs encourage underinvoicing. If the host CIT exceeds (is less than) the combined home CIT rate plus the tax-adjusted host tariff, the MNE should over- (under-) invoice its exports to the foreign affiliate.

The general conclusions from this more complicated model of the impacts of corporate income taxes, withholding taxes, and tariffs on MNE behaviour are that (1) marginal rates of tax tend to affect the real decision variables such as investment and output; (2) nominal or statutory tax and tariff rates affect financial variables such as the choice of transfer price, repatriation of dividends, and the remittance of head office fees; (3) the MNE chooses its transfer pricing policy for intrafirm trade in goods by comparing the tax differential with the tax-adjusted tariff rate; and (4) the MNE should attempt to minimize financial payments, and maximize tax-deductible expenses, in high-tax locations through financial transfer pricing, if the goal of the enterprise is to maximize global after-tax profits. We come back to these conclusions in the next chapter when we explore the impact of tax differentials and tariffs on Canada–U.S. intrafirm trade.

Tax Penalties for Transfer Price Manipulation

In this chapter we have developed several cases in which the MNE has an incentive to manipulate transfer prices. *Transfer price manipulation (TPM)*, as we have used the term so far in this chapter, means to set a price different from the shadow price of intrafirm trade in order to reduce tax and/or tariff payments to the government. TPM occurs because the multinational increases its global net-of-tax-and-tariff profits by over- or underinvoicing its transfers and expanding the volume of intrafirm trade. Governments historically have been concerned about transfer prices for just this reason, that is, the ability of MNEs to alter their prices so as to reduce their overall tax and/or tariff costs.

So far we have explored one way in which governments can affect the incentives for TPM: the regulated price $\hat{\mathbf{W}}$. By setting a floor or ceiling on the MNE's transfer price, governments constrain the tax and/or tariff savings that are available to the enterprise from over- or underinvoicing. Whether or not the constraint is successful depends on (1) the government's ability to predict which way the MNE will want to move the transfer price (higher or lower), (2) whether $\hat{\mathbf{W}}$ is set correctly (i.e., in the right direction), and (3) whether or not the constraint binds (i.e., the MNE would have preferred to set **p** outside the regulated price).[32]

In this section we look at another way to reduce TPM. The United States has recently introduced penalties for misstatement of transfer prices; that is, if the transfer price lies too far outside the regulated price a penalty is levied on the differential. This is a different view of TPM than the one we have been using. In our economic models in this chapter, TPM occurs when the profit-maximizing transfer price **p*** is different from the shadow transfer price λ.

The federal government, however, defines TPM as a transfer price declared by the MNE for tax purposes or when paying customs duties (this will be either the profit-maximizing price **p*** or the money transfer price **w**, depending on whether the MNE keeps one or two sets of books) that is different from the price determined by the regulator (i.e., that is different from the regulated transfer price $\hat{\mathbf{W}}$). So the regulator looks at the gap between **p*** and $\hat{\mathbf{W}}$ (assuming one set of books), not between **p*** and λ.

Let us call the regulator's view of TPM *regulatory transfer price manipulation (RTPM)* and the economist's view of TPM *economic transfer price manipulation (ETPM)*. Clearly, RTPM and ETPM will only be the same if $\hat{\mathbf{W}} = \lambda$ and

if the MNE keeps one set of books. Since the regulatory concept is the arm's length standard, the key question is: Will the MNE, in the absence of taxes and tariffs, select the arm's length price?

We know that the answer to this question depends on the answers to three subquestions, that is, whether: (1) an exact comparable exists in the open market, (2) the related parties are free to buy or sell on the external market, and (3) there are no interdependencies in demand or supply within the MNE. If the answer to all three questions is yes, then the Hirshleifer rule says that the MNE will choose the external price. In this case the arm's length price is the efficient shadow price and $\hat{\mathbf{W}} = \lambda$. When the MNE reacts to a tax or tariff by moving the profit-maximizing price away from λ, the amount of ETPM is exactly the same, and in the same direction, as the amount of RTPM.

However, in practice, we would not expect this to occur. Exact comparables seldom occur in real life; even in the case of undifferentiated tangibles such as crude oil and wood pulp, MNEs and tax authorities are forced into making adjustments to an inexact comparable to estimate a transfer price. In the case of differentiated goods, group services, and intangibles, the chance of finding a good comparable is even less likely. Therefore the regulator's price may be far away from the free trade shadow price λ, and the transfer price manipulation the tax authority is trying to measure will be quite different from that an economist would be estimating.

We first quickly summarize the U.S. rules in section 6662 (see Chapter 9 for more details). A *substantial valuation misstatement (SVM)* occurs if the transfer price **p** is 50 per cent or less, or 200 per cent or more, than the regulated transfer price ($\hat{\mathbf{W}}$); a tax penalty of 20 per cent of the amount of estimated underpayment of tax is due in this case. A *gross valuation misstatement (GVM)* occurs if **p** is 25 per cent or less, or 400 per cent or more, than $\hat{\mathbf{W}}$; if this occurs, a 40 per cent penalty is levied on the amount of underpaid taxes.

We model this as follows. Let α be the penalty rate levied by country 2 on the underpayment of taxes. Under section 6662, α equals 20 per cent for an SVM and 40 per cent for a GVM. We ignore these complications and simply assume α is the same percentage regardless of the size of the misstatement or its direction. Let the percentage gap between **p** and $\hat{\mathbf{W}}$ that triggers the penalty be ϕ. We assume ϕ is some fixed percentage and does not vary with the size of the misstatement or its direction (e.g., the penalty applies if the transfer pricing gap is plus or minus 50 per cent of the regulated price).

The total penalty amount (call it $\mathscr{P}$) is α times the additional tax owed by the MNE, which is α times t_2 multiplied by the absolute value of the change in the transfer price $|(\mathbf{p} - \hat{\mathbf{W}})|$ times the volume of intrafirm trade X. Thus the penalty is:

$$\mathscr{P} = \alpha t_2|(\mathbf{p} - \hat{\mathbf{W}})|X \quad \begin{aligned} &> 0 \text{ if } |(\mathbf{p} - \hat{\mathbf{W}})| \geqslant \phi\hat{\mathbf{W}} \\ &= 0 \text{ if } |(\mathbf{p} - \hat{\mathbf{W}})| < \phi\hat{\mathbf{W}} \end{aligned} \tag{63}$$

which is positive if the absolute value of the transfer pricing gap $\mathbf{p} - \hat{\mathbf{W}}$ equals or exceeds the percentage ϕ of the regulated transfer price.

In order to model the impacts of penalty regulation on the MNE's choice of transfer pricing policy, it is critical to know whether or not ETPM and RTPM are the same measures. We make the following assumptions to simplify the analysis. Since there is no outside market, the internal shadow price λ is the profit-maximizing transfer price for the MNE in the absence of taxes or tariffs. The regulatory authorities accept this transfer price as the arm's length price, so $\lambda = \hat{\mathbf{W}}$. Therefore ETPM and RTPM are one and the same in this model.

Now assume we have a horizontally integrated MNE in which both governments levy taxes on their own firm's pure profits on a source basis; assume firm 1 is the exporter. Note that this model is similar to that in equations (51–3), but without the tariff. The MNE's objective function with the penalty is:

$$\begin{aligned} \pi^* = {} & (1-t_1)[R_1(Y_1) + \mathbf{p}X - C_1(Q_1)] + (1-t_2)[R_2(Y_2) - C_2(Q_2) - \mathbf{p}X] \\ & - \alpha t_2|(\mathbf{p} - \hat{\mathbf{W}})|X + \lambda_1[Q_1 - Y_1 - X] + \lambda_2[Y_2 - Q_2 - X] \end{aligned} \tag{64}$$

where α is positive only if $|(\mathbf{p} - \hat{\mathbf{W}})| \geq \phi\hat{\mathbf{W}}$. The first-order conditions for a profit maximum, which are found by differentiating (64) with respect to Q_1, Q_2, Y_1, Y_2, and X, are:

$$(1 - t_1)MR_1 = (1 - t_1)MC_1 = \lambda_1 \tag{65}$$

$$(1 - t_2)MR_2 = (1 - t_2)MC_2 = -\lambda_2 \tag{66}$$

$$(t_2 - t_1)\mathbf{p} - \alpha t_2|(\mathbf{p} - \hat{\mathbf{W}})| = \lambda_1 + \lambda_2 \tag{67}$$

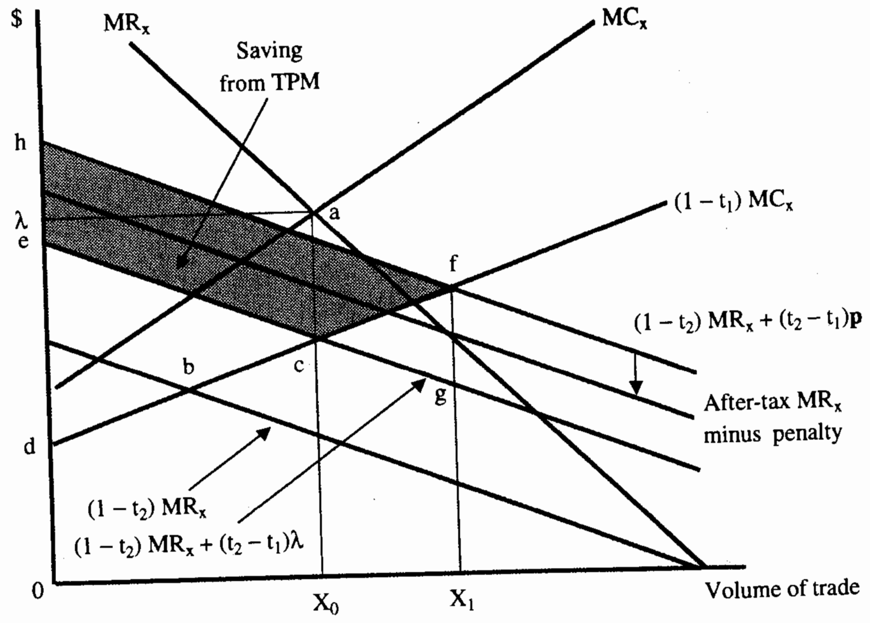
Putting these together and rearranging, we have:

$$\begin{aligned} & (1 - t_1)MR_1 = (1 - t_1)MC_1 = \\ & (1 - t_2)MR_2 + [(t_2 - t_1)\mathbf{p} - \alpha t_2|(\mathbf{p} - \hat{\mathbf{W}})|] = \\ & (1 - t_2)MC_2 + [(t_2 - t_1)\mathbf{p} - \alpha t_2|(\mathbf{p} - \hat{\mathbf{W}})|] \end{aligned} \tag{68}$$

Equation (68) is not as complicated as it looks; in the absence of the penalty, (68) collapses to:

$$\begin{aligned} & (1 - t_1)MR_1 = (1 - t_1)MC_1 = (1 - t_2)MR_2 + (t_2 - t_1)\mathbf{p} \\ & = (1 - t_2)MC_2 + (t_2 - t_1)\mathbf{p} \end{aligned} \tag{69}$$

FIGURE 6.3
Tax Penalties for Transfer Price Manipulation



which can be rewritten as:

$$(1 - t_1)MC_x = (1 - t_2)MR_x + (t_2 - t_1)\mathbf{p} \tag{70}$$

Thus the impacts of the two source taxes on the MNE, if we ignore the penalty, are straightforward. There are two effects. First, the *ad valorem tax effects* shift the curves for the marginal cost of exporting MC_x and the marginal revenue from importing MR_x downwards by their respective tax rates. Second, the *specific tax effects* shift the MR_x curve up by the tax-adjusted transfer price.

When we combine these two effects, if t_1 equals t_2, equation (70) simply reduces to $MC_x = MR_x$ so that complete tax harmonization eliminates the incentive to transfer price. Similarly, if the MNE sets $\mathbf{p}$ equal to $MC_x = MR_x$, (70) also collapses to the free trade condition; therefore, setting $\mathbf{p} = \lambda$ returns the enterprise to its initial sales, output, and trade levels.

Assume, however, that t_2 exceeds t_1. Then the MNE has an incentive to over-invoice firm 1's exports to firm 2 in order to shift profits to the lower-taxed country. We illustrate this in Figure 6.3.

The MNE in Figure 6.3 initially sets $MR_x = MC_x$ at point a with intrafirm trade of X_0. The ad valorem tax effects are shown by the downward rotation of each of these curves. We assume that t_2 exceeds t_1 so that the intersection of the after-tax curves (point b) lies to the left of X_0. The specific tax effect causes the after-tax MR_2 curve to shift upwards by the tax differential times the transfer price. If the MNE keeps its transfer price at the shadow price, the new equilibrium at point c lies directly below point a and the volume of intrafirm trade remains at X_0. After-tax MNE profit is the triangular area dce.

The MNE, however, would prefer to overinvoice firm 1's exports to firm 2 since these are tax-deductible expenses and jurisdiction 2 is the high-tax jurisdiction. We show overinvoicing as a further upward shift in the MR_x curve. The new equilibrium is at point f with X_1 traded. The new after-tax MNE profit is the triangle dfh. There are two effects on MNE profits from overinvoicing: the gain in profit from the tax saving on the transfer price represented by the parallelogram ehfg, and the fall in profit from the misallocation of resources caused by the expansion in the trade volume, equal to triangle cfg. The net gain to the MNE from overinvoicing is the shaded area ehfc.

With the penalty rules in place, equation (68) can be rewritten, like equation (70), as:

$$(1 - t_1)MC_x = (1 - t_2)MR_x + [(t_2 - t_1)\mathbf{p} - \alpha t_2|(\mathbf{p} - \hat{\mathbf{W}})|] \qquad (71)$$

We now have a third factor influencing the MNE: the *specific penalty effect*, which acts so as to offset the specific tax effect. If the transfer price **p** differs from the regulated price $\hat{\mathbf{W}}$, the MNE pays an additional tax to the government if the differential exceeds the stated percentage.

The profit-maximizing transfer price can be found by partially differentiating (64) with respect to **p** and using the envelope theorem, as follows:

$$\partial\pi^*\backslash\partial\mathbf{p} = [(t_2 - t_1) - \alpha t_2)X = [(1 - \alpha)t_2 - t_1)]X \qquad (72)$$

According to (72), in the absence of the tax penalty ($\alpha = 0$), the MNE would simply overinvoice its transfer price when $t_2 > t_1$, or underinvoice if $t_2 < t_1$. With the penalty, however, the incentive to manipulate the transfer price is reduced if the manipulation moves **p** sufficiently (higher or lower than $\hat{\mathbf{W}}$ so as to trigger the penalty. For example, suppose t_2 is 40 per cent, t_1 is 34 per cent, and α is 20 per cent. Since t_2 exceeds t_1, the MNE would like to overinvoice **p**, in the sense of charging a transfer price higher than the free trade marginal cost of the exporting firm (the shadow price λ). However, if doing this triggers the penalty, the MNE must pay an additional eight per cent tax (20 per cent times 40 per cent). Thus the gain from transfer price manipulation is reduced by the amount of the additional tax.

We show this in Figure 6.3 by shifting the after-tax MR_x curve downward to reflect the smaller (or possibly zero) net gain to the MNE from transfer price manipulation. Thus the penalty can be an effective way to reduce incentives to over- and underinvoicing.

We turn now to another proposal that would reduce the multinational's incentive to manipulate transfer prices: unitary taxation.

Transfer Pricing and Unitary Taxation

Unitary taxation is taxation of the worldwide income of a *unitary business*. A unitary business consists of all the related affiliates of an enterprise that do business within the taxing jurisdiction. For example, if the jurisdiction is California, and one affiliate of IBM is located in California, all the related affiliates of IBM could be considered as a unitary business and the worldwide income of IBM taxed by the state of California. Normally, unitary taxation is based on a *formula apportionment or worldwide combined reporting* method whereby California IBM's share of certain factors (e.g., employment, sales, capital stock – the so-called 'apportionment factors') as a percentage of the worldwide IBM amount of these factors, however weighted, multiplied by the total worldwide income of IBM, is used to compute the tax to be paid by IBM to the state of California.

Separate accounting, on the other hand, defines the borders of a firm (a permanent establishment) according to national boundaries, the so-called 'water's edge.' Domestic affiliates are consolidated with the parent for tax purposes (as are foreign branches), but foreign subsidiaries and other affiliates of the MNE are treated as separate firms. Transfer price rules are used to ensure that such transactions approximate arm's length prices.

In this section we investigate the implications of unitary taxation for transfer pricing. We concentrate on the case in which one jurisdiction adopts a formula apportionment approach, and the other jurisdiction does not, and examine the resulting distortions and opportunities for transfer price manipulation.[33]

Setting Up the Model

Assume the MNE is a horizontally integrated multinational consisting of two firms that share joint overhead costs, where firm 2 exports an intrafirm traded good to firm 1. The pre-tax global profit function of the MNE is:

$$\pi = (R_1(Y_1) - W_1L_1 - \mathbf{p}X) + (R_2(Y_2) - W_2L_2 + \mathbf{p}X) - F \qquad (73)$$

where π is pre-tax global MNE profit, $R_i(Y_i)$ is total revenue of firm i from sales Y_i, **pX** is the value of intrafirm trade, W_iL_i is the wage bill in firm i, and F is overhead costs (i = 1, 2). We assume, for simplicity, that all costs are labour costs; therefore, total cost C_i equals W_iL_i for each firm i.

Assume country 2 follows the water's-edge principle and taxes only profits arising in its jurisdiction, whereas country 1 applies formula apportionment to the worldwide income of its residents. Assume the ratio used to determine the share of MNE profits taxable in country 1 is firm 1's share of worldwide labour costs of the MNE.[34] The MNE's objective is to maximize its after-tax global profit function:

$$\begin{aligned}\pi^* = {} & (R_1(Y_1) - W_1L_1 - \mathbf{p}X) + (R_2(Y_2) - W_2L_2 + \mathbf{p}X) - F \\ & - t_1\{\beta[(R_1(Y_1) - W_1L_1 - \mathbf{p}X) + (R_2(Y_2) - W_2L_2 + \mathbf{p}X) - F]\} \\ & - t_2[(R_2(Y_2) - W_2L_2 + \mathbf{p}X) - \alpha_2 F] \qquad (74)\end{aligned}$$

where π^* is post-tax MNE global profit and α_i is the tax-deductible share of overhead costs allocated to jurisdiction i where $\alpha_1 + \alpha_2 = 1$. Firm 1's taxable income is calculated as worldwide MNE income, net of expenses, multiplied by the labour factor ratio β,[35] the weighting factor used to determine firm 1's taxes in country 1, where:

$$\beta = (W_1L_1)/(W_1L_1 + W_2L_2) \qquad (75)$$

Thus the first line in (74) represents pre-tax global profits of the MNE (π), the second line the taxes paid by firm 1, and the third line the taxes paid by firm 2.

The Various Transfer Pricing Choices

What should the MNE do to maximize its after-tax global profits in these circumstances? It is sufficient to look at L_1, L_2, and X in order to determine optimal output, sales, and trade volumes. Looking first at the national factor markets, where the two firms hire units of labour, we differentiate (74) with respect to L_i, recalling that $Y_1 = Q_1 + X$ and $Y_2 = Q_2 - X$:

$$\partial\pi^*\backslash\partial L_1 = (1 - \beta t_1)(MR_1\partial Q_1\backslash\partial L_1 - W_1) = 0 \qquad (76)$$

$$\partial\pi^*\backslash\partial L_2 = (1 - \beta t_1 - t_2)(MR_2\partial Q_2\backslash\partial L_2 - W_2) = 0 \qquad (77)$$

Each firm should hire units of labour up to the point where the marginal revenue product of labour[36] (MRP_{Li}) equals the wage rate W_i, both measured on an after-tax basis. Putting the two equations together, we have:

$$(1 - \beta t_1)(MRP_{L1} - W_1) = (1 - \beta t_1 - t_2)(MRP_{L2} - W_2) \qquad (78)$$

Equations (76, 77) say that the MNE as a whole should allocate labour between the two firms such that the after-tax marginal revenue product, net of the wage rate, is equalized between the two firms.

Note that the tax rate for firm 1 is βt_1, whereas the tax rate on firm 2's profits is $t_2 + \beta t_1$. The reason for this is straightforward. All profits wherever earned are taxed at βt_1 in country 1; in addition, profits earned by firm 2 are taxed in country 2 at rate t_2. Unless country 2 credits country 1's tax, or vice versa, this double taxation of MNE profits on intrafirm trade persists as long as $\beta > 0$. Therefore the effective tax rate is higher on firm 2's profits, by the amount t_2, as long as country 1 practises unitary taxation.

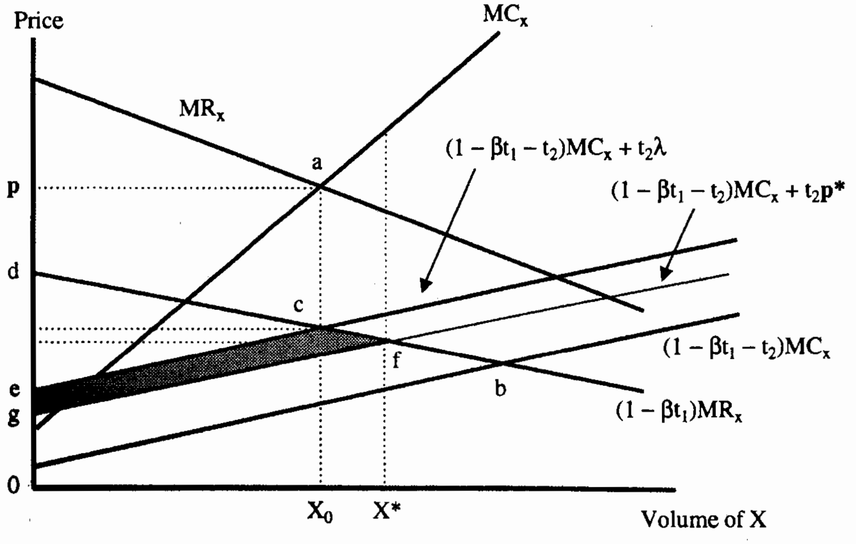
Since β is a fraction, it can lie either above or below t_1, depending on firm 1's share of total MNE labour income. Under separate accounting, the MNE would have paid taxes of $t_1\pi_1$ to country 1; under unitary taxation the MNE pays $t_1\beta(\pi_1 + \pi_2)$. Thus, when country 1 moves from separate accounting to unitary taxation, the tax rate on firm 1's profits falls (as long as $\beta < 1$), but country 1 now taxes firm 2's profits at the rate $t_1\beta$, which is positive as long as $\beta > 0$. Whether the effective tax rate on firm 1's profits (i.e., the total tax divided by firm 1's profits) rises or falls depends on whether β, the apportionment factor, is larger or smaller than $\pi_1/(\pi_1 + \pi_2)$.[37]

This situation has two effects. First, as long as the MNE has the choice between locating in country 1 or country 2, the firm will not choose to locate in the unitary tax jurisdiction because doing so in effect opens its other MNE affiliates up to taxation by that jurisdiction, since the 'water's edge' is ignored. Unless either taxing authority is willing to provide a full foreign tax credit for the double taxation of firm 2's profits, the MNE's overall tax costs rise. Second, if the MNE is already located in country 1, and therefore is paying unitary taxes, the enterprise will shift activities out of the other jurisdictions to avoid the additional taxation by country 1 unless country 2 credits the unitary taxes or country 1 credits the taxes already paid in country 2 on π_2.[38]

Turning to the effects on intrafirm trade, we differentiate (74) with respect to X, the volume of intrafirm trade:

$$\begin{aligned}\partial\pi^*\backslash\partial X &= (1 - \beta t_1)(MR_1 - \mathbf{p}) - (1 - \beta t_1 - t_2)(MR_2 - \mathbf{p}) \\ &= (1 - \beta t_1)MR_1 - (1 - \beta t_1 - t_2)MR_2 - t_2\mathbf{p} = 0 \qquad (79)\end{aligned}$$

FIGURE 6.4
Transfer Pricing under Mixed Systems: Unitary Taxation and Separate Accounting



which says the MNE should balance firm 1's after-tax net marginal revenue from imports, $(1-\beta t_1)\mathbf{MR_x}$, against firm 2's after-tax net marginal cost of exports $(1-\beta t_1-t_2)MC_x-t_2\mathbf{p}$), or:

$$(1-\beta t_1)MR_x = (1-\beta t_1-t_2)MC_x + t_2\mathbf{p}$$
$$= (1-\beta t_1)MC_x - t_2(MC_x-\mathbf{p}) \qquad (80)$$

If $MC_x = \lambda$ (the transfer price is set equal to the shadow price on intrafirm trade), equation (80) collapses to $MR_x = MC_x$ and the volume of intrafirm trade does not change from its pre-tax level. We illustrate the impact of setting the transfer price in Figure 6.4, which is based on the middle graph in Figure 5.1.

In the pre-tax situation, the MNE chooses the volume of intrafirm trade where $MR_x = MC_x = \mathbf{p}$ at point a in Figure 6.4 (also labelled as point v in Figure 5.1. When unitary taxes are levied on firm 1, and separate accounting on firm 2, the MNE reacts by setting the after-tax marginal cost of exports to the after-tax marginal revenue from imports, as in equation (70). The MR_x curve rotates downward to $(1-\beta t_1)MR_x$, and the MC_x curve rotates downward to $(1-\beta t_1-t_2)MC_x$; these are the *ad valorem tax effects*. This intersection is labelled point b. The marginal cost curve then shifts upward by $t_2\mathbf{p}$; this is the ***specific tax effect.***

Assume initially that **p** is set equal to $MC_x = \lambda$ so the MNE absorbs the tax. The new equilibrium is at point c with the same output as before, X_0. The after-tax global MNE profits are the triangle dce.[39]

These are the first-order conditions for an after-tax profit maximum; however, the MNE can affect its overall tax payments in three ways: changes to (1) β, the weighting factor in the unitary tax formula, (2) α_2, the share of overhead costs allocated to the country using the separate accounting approach, and (3) **p**, the transfer price.

First, if the MNE can affect β, the weighting factor should be set as *low* as possible, as can be seen from the equation below, where we differentiate (74) with respect to β and use the envelope theorem:[40]

$$\partial\pi^*\backslash\partial\beta = -t_1\pi < 0 \qquad (81)$$

Therefore, if the MNE can reduce the factor ratio used to determine its effective tax rate, overall MNE profits are higher.[41]

If the MNE can affect the allocation of the fixed costs F between the two countries, the MNE should set α_2 as *high* as possible:

$$\partial\pi^*\backslash\partial\alpha_2 = t_2F > 0 \qquad (82)$$

When one government gives a tax deduction for an affiliate's share of overhead expenses and the other government does not, it makes sense to maximize the affiliate share in the country with the deduction.

Lastly, the MNE should set its transfer price **p** as *low* as possible:

$$\partial\pi^*\backslash\partial\mathbf{p} = -t_2X < 0 \qquad (83)$$

Manipulation of the transfer price **p** no longer affects the taxes paid in country 1 (since βt_1 applies to pre-tax profits for the MNE as a whole), but still affects the taxes paid in country 2. Since firm 2 is the exporter, any income it makes from intrafirm trade is taxable. Therefore the MNE should minimize the transfer price to reduce its overall tax bill. This is clear from Figure 6.4, where the lower is **p** the lower the effective tax on the MNE and the greater the after-tax profits. Underinvoicing is shown as the new price $\mathbf{p^*} < \lambda$, causing a downward shift in the after-tax marginal cost of exports curve; this causes an expansion of X to **X***, and an increase in after-tax profits represented by the shaded area ecfg.

Thus, where one country taxes on the basis of formula apportionment, while the other country or countries follow traditional separate accounting methods, there are still ways in which transfer price manipulation can be used to reduce

MNE tax payments. Would unitary taxation work if all countries used this approach? Our immediate reaction is to say 'yes.' We show below, however, that there is still at least one 'slip between the cup and the lip' that allows MNEs to avoid paying taxes under a global unitary tax system.

Global Unitary Taxation

Assume both governments follow a unitary tax approach and that they define and measure the MNE's global pre-tax income identically and accurately as π, and each government taxes a share β_i of the worldwide income of the MNE where the country allocation factors sum to unity so that $\beta_1 + \beta_2 = 1$.[42] Assume initially that the tax rates t_1 and t_2 differ. The MNE's after-tax global profit function is:

$$\pi^* = (1 - t_1\beta_1 - t_2\beta_2)\pi \qquad (84)$$

Since $\beta_2 = 1 - \beta_1$, we can rewrite (84) as follows:

$$\pi^* = [1 - \beta_1(t_1 - t_2) - t_2]\pi \qquad (85)$$

The first-order conditions are straightforward (and are left to the reader). With all profits wherever earned taxed at the same rate, the MNE simply absorbs the tax and does not change its output, sales, or trade volumes. As a result, there are no resource allocation effects, or deadweight losses, imposed on the MNE; the tax is completely neutral. This is one of the key benefits argued by the proponents of unitary taxation (see Chapter 12).

The MNE, however, still has some ability to manipulate its tax payments. Differentiating (85) with respect to β_1, **p**, and α_i, we have:

$$\partial\pi^*\backslash\partial\beta_1 = (t_2 - t_1)\pi \qquad (86)$$

which is positive (negative) if t_2 exceeds (is less than) t_1, thus, the MNE should raise (lower) β_1 whenever country 1's tax rate is less (greater) than country 2's rate. Therefore differences in tax rates can still be exploited by manipulating the factor allocation ratio at the national level, reducing it in high-tax countries and raising it in low-tax countries.

The allocation of fixed costs between the two countries however makes no difference to the total tax paid, as long as the costs are deductible in both countries; see equation (87) below:

$$\partial\pi^*\backslash\partial\alpha_2 = 0 \qquad (87)$$

Lastly, the transfer price **p** also no longer affects total tax payments:

$$\partial\pi^*\backslash\partial\mathbf{p} = 0 \qquad (88)$$

The MNE therefore has its degrees of freedom significantly reduced, but not eliminated. In practice, where formula apportionment is applied, generally a three-factor formula is used. That is, the multi-factor ratio for country i (F_i) is an arithmetic average of all the factors from 1 through n, defined by:

$$F_i = (w_{1i}F_{1i} + w_{2i}F_{2i} + w_{3i}F_{3i} + \ldots + w_{ni}F_{ni})/n \qquad (89)$$

where F_{ni} is factor n in country i and w_{ni} is the weight ($0 < w_{ni} < 1$) attached to F_{ni}. Once F_i is calculated, the MNE's pre-tax profit in jurisdiction i is estimated as F_i times total MNE profit, or $\Pi_i = F_i\,\pi$. The taxes paid in jurisdiction i are then determined as the actual tax rate multiplied by the estimated MNE profit. As before, Π_i, the estimated profit, may be greater or less than π_i, the profit actually declared by firm i.

The F_i formula makes it clear that MNEs can manipulate their taxes in several ways, such as: (1) misrepresenting the size of variables that carry a high weight in the formula, (2) physically moving high-weight activities out of high-tax jurisdictions, and (3) lobbying governments to reduce the weights and/or change the factors in the formula. The only way to eliminate transfer price manipulation is for all governments to use exactly the same formula, applied to all global income sources, at the same tax rate – an unlikely occurrence at best.

Conclusions

This concludes Chapter 6 on the theory of taxing multinationals. The chapter developed a general microeconomic theory of transfer pricing behaviour by multinational enterprises in response to taxes and trade barriers. We showed that a profit-maximizing MNE will attempt to arbitrage the imperfections in product and factor markets induced by government regulations, such as tariffs, profit and corporate income taxes, and minority shareholder requirements. The models explained the MNE's choice of transfer pricing policy, both for intrafirm trade in tangibles and for financial manoeuvres such as dividend repatriation and head office fees. We now turn to the empirical work that has been done on taxing multinationals, focusing in particular on the tax treatment of MNEs in North America. We will see that there is some evidence supporting the theoretical predictions of the above model, but that the data are, for the most part, inconclusive.