Abstract - While regulations on tax evasion via transfer pricing have been thoroughly explored in the literature, little is known about the effects of domestic tax policies under a Bilateral Advanced Price Agreement (hereafter BAPA) on production decisions of multinational companies. I show that, even if BAPAs prohibit income shifting through transfer pricing, inefficiencies arise via distorted production when markets across countries are interrelated through intra–firm trade of multinational companies. The analysis provides governments with useful guidance for how to coordinate tax agreements.

INTRODUCTION

Tax avoidance by multinational companies is a serious problem for governments. Since tax systems differ across countries, multinational companies can reduce their tax burden by shifting their profits to countries with relatively low tax rates. Transfer pricing is one such method. Manipulating the transfer price in intra–firm transactions across countries allows multinational companies to move income among affiliates located in different countries. The last 30 years have been an era of trial and error for governments attempting to regulate such tax avoidance. In the process, policy adjustments were made to coordinate the systematic discrepancies in tax laws across countries. Among them, one of the most noteworthy achievements is the Bilateral Advanced Pricing Agreement (BAPA). The BAPA is an agreement between tax authorities in two countries (and a multinational company) to use the same arm’s length price so as to eliminate the risk of double taxation.

This paper investigates efficiency losses caused by the BAPA case. While the problem of tax avoidance via transfer pricing and the proposed regulations (i.e., mechanism design for governments) have been thoroughly explored in the literature, little is known about the externalities on the production decisions of multinational companies that result from domestic tax policies under BAPAs. I examine the degree to which BAPAs create efficiency loss and propose ways to remedy this loss.
The model falls in between those found in the standard literature on source–based capital tax competition (Wilson, 1986; Zodrow and Mieszkowski, 1986; Mieszkowski and Zodrow, 1989) and the more recent ones which appear in the literature on the transfer pricing of multinational companies. With the expansion of multinational business in the 1970s, transfer pricing issues have figured prominently in the literature (see, for instance, the survey by Gresik (2001)). The earlier literature tries to explain the decisions of multinational companies when tax systems differ across jurisdictions, and shows that the use of transfer pricing can be an arbitrage device used to reduce the tax burden of companies (Horst, 1971; Copithorne, 1971; Eden, 1985; Kant, 1990). More recent papers treat transfer pricing as private information and propose mechanism designs for less–informed governments; i.e., they try to show how these governments can implement tax policies so as to induce an appropriate transfer price from multinational companies (Prusa, 1990; Gresik and Nelson, 1994; Stoughton and Talmor, 1994; Bond and Gresik, 1996).

This paper introduces the BAPA system into the discussion. This allows me to highlight the inefficiency arising from domestic tax policies under the BAPA system. Much of the theoretical work on transfer pricing stems from private information and/or commitment issues among companies and governments (i.e., governments cannot observe the costs of the companies). The analysis in this paper can safely ignore these issues, since a multinational company and two governments can agree on a transfer pricing rule and commit to the agreement. Nevertheless, there is still a problem. I show that, even if the BAPA prohibits the company shifting income through transfer pricing, an efficiency loss arises via distorted production decisions of the multinational company.

The situation is modeled as a tax game of complete information with three players: a multinational company, a tax authority in a foreign host country, and a tax authority in a home country. At the outset, the multinational company and the two governments agree to use the same arm’s length price. This eliminates information problems and the risk of double taxation.1 Given the transfer price (and tax rates) in the home and host countries, the company chooses output to maximize after–tax profits of all affiliates.

The equilibrium turns out to be inefficient. Alternative tax policies have the potential to enlarge the before–tax profits of the company (or the tax base) through a more efficient allocation of the tax burden. The source of the inefficiency is coordination failure between the two governments when both countries share a tax base. The difference in tax rates between the countries hosting the parent and the subsidiary distorts the production decision of the multinational company. This is because the company reduces its tax burden by adjusting output so as to generate more profit in the country with a lower tax rate. The output adjustments thus play an intermediate role in creating the efficiency loss.2

The degree of the inefficiency increases with a larger mark–up rate if the cost–plus

---

1 This paper is also related to Bond and Samuelson (1989) and Janeba (1995). They consider different double taxation conventions in the presence of international capital mobility and show the inefficiency of non–cooperative tax policy when governments influence the location of investment.

2 This is a close analogy to investment distortion in capital tax competition. Countries of different sizes set different tax rates and, thus, the investment decision of internationally mobile capital is distorted (Bucovetsky, 1991; Wilson 1991). The inefficiency here is partly due to the fact that governments in different countries share a tax base when the markets across countries are interrelated through intra–firm trade. In this respect, the analysis is related to the literature on vertical fiscal externalities such as Boadway and Keen (1996), Boadway et al. (1998), and Keen and Kotsogiannis (2002), in which the tax bases of state governments
method detailed in Regulations §482 is used. A positive mark–up ratio operates as a proxy of a hazard index that measures the degree under which BAPAs hinder the vertical integration of multinational companies. While multinational companies are integrated businesses (“groups”) under common control, the companies cannot internalize the costs of intra–firm transactions in the BAPA system. BAPAs segregate the profits earned by two different affiliates within the same company for the purpose of imposing the tax independently. A lower mark–up ratio works to eliminate the hazard and increase the profits of the companies. Similar inefficiencies would result in other arm’s length methods under BAPAs. The loss of efficiency under BAPAs is closely analogous to the inefficiency caused by double marginalization of monopolies in the industrial organization literature.

The analysis provides useful guidance for the coordination of tax agreements. Inefficiency will be overcome once either a zero mark–up ratio or a uniform tax rate eliminates the segmentation between two or more affiliates within the same company. This allows the company to earn more profits by internalizing the cost of the intra–firm transaction. In either case, the coordination needs to involve transfer payments between the two countries since an efficient outcome will be attained at the cost of lower tax revenue in one country.

The plan of the paper is as follows. The second section describes a model for studying the externalities that result from domestic tax policies under BAPAs. The third section clarifies the source of the efficiency loss caused by those policies. Concluding remarks offer directions for future research.

MODEL

The markets across two countries are interrelated through an intra–firm transaction of a vertically integrated multinational company. Trade within a firm is modeled as a manufacturing process from a mother factory in an upstream location to assembly (or distribution) factories at the destination market. A typical example is a parent company in the home country producing and exporting intermediate goods that are further assembled or manufactured by a subsidiary in the host country. Final goods are sold in the local markets of the host country.

Following the traditional vertical integration literature, the intra–firm transaction is characterized by the fixed–coefficient production function (Greenhut and Ohta, 1979). Let \( q > 0 \) be the quantity of the intermediate goods produced by the multinational parent in the home country, and \( Q > 0 \) be the quantity of the final goods processed by the subsidiary in the host country. The production function is denoted as \( Q = \alpha q \), where \( \alpha \) is a positive constant, assuming that the amount of a local input required for the production is proportional to \( q \). I use the special case of \( \alpha = 1 \), as is commonly observed in the transfer pricing literature, with the proper choice of units such that one unit of the intermediate good is required to produce one unit of the final good.

The company is assumed to have a monopoly on its differentiated goods in the overlap with a federal government. On the other hand, tax avoidance (which, in this paper, means that a company tries to generate more profit in the country with a lower tax rate) is similar to horizontal fiscal externalities examined in Zodrow and Mieszkowski (1986), Wilson (1986), Mieszkowski and Zodrow (1989), and Wildasin (1991).

3 A uniform tax rate across regions is also obtained in models of capital tax coordination (e.g., Wildasin (1989), DePater and Myers (1994)), though the mechanisms driving the results are different.

4 The possibility that one country may lose from tax coordination also arises in different models of capital tax competition (Bucovetsky, 1991; Wilson 1991).
host country market. Let the inverse demand function for final goods in the host market be \( p = p(q) \), where \( p \) is the price of the final goods. The price is assumed to be twice continuously differentiable, strictly monotonically decreasing, and concave in the quantity of output. The demand for intermediate goods is derived indirectly from the demand for the final goods in a vertically integrated structure.

The factor markets are characterized to be competitive (either in the home or host country) because many local companies provide non–differentiated parts necessary for production. I use a simple linear cost function \( C_i = c_i q \), where \( c_i \) is a positive constant marginal cost in location \( i \). The location is denoted as \( i = h \) for the home country and \( i = f \) for the host country.

The transfer price of the intra–firm trade is regulated as in the current world tax system, and denoted as \( \theta - ch \) with a positive constant mark–up rate \( k \).\(^5\) The mark–up is an advanced agreement between a multinational company and two governments, and is carefully chosen so that profit of the company is allocated to each country. This is the so called BAPA case, in which tax authorities in the two countries agree to use the same arm’s length price so as to eliminate the risk of double taxation.\(^6\) While this assumption ignores both private information and commitment issues, which are the focus of much work on transfer pricing, it allows me to highlight the inefficiency arising from domestic tax policies under the BAPA system.

Decision of the Company

A multinational company chooses output to maximize after–tax profits of the group given the mark–up rate and tax rates in the home and host countries.\(^7\) Global profit maximization is assumed as is typical in the literature, though companies may have multiple objectives and could possibly benefit from decentralization. Each affiliate pays corporate income taxes calculated at a corporate tax rate \( t_i \) in its resident country. The after–tax global profits of the multinational company are constructed as the sum of profits earned in the two countries:

\[
\Pi = ((1 - t_h)(\bar{\theta} - c_h) + (1 - t_f)(p(q) - \bar{\theta} - c_f))q.
\]

The first–order condition (which is also a sufficient condition) provides the familiar, but slightly modified, relationship with the after–tax marginal revenue equated to the after–tax marginal cost at the group level.

\[
(1 - t_h)(\bar{\theta} - c_h) + (1 - t_f)\left(\frac{dp}{dq} q + p - \bar{\theta} - c_f\right) = 0.
\]

The equilibrium output \( q^* \) is a function of the parameters in the cost and demand functions, including the tax rates in the two countries.

Proposition 1

The equilibrium output \( q^* \) is decreasing in the home tax rate \( t_h \) and increasing in the foreign host tax rate \( t_f \).

\(^5\) This is the so–called cost plus method and is used in cases involving manufacturing, assembly or other production goods that are sold to related parties (Intercompany Transfer Pricing Regulations, §1.482–3(d)(1)). The cost–plus method is the most frequently used method according to surveys on intra–firm transfer pricing (Al–Eryani, Alam, and Akhter, 1990; Hamaekers, 1992; Ernst & Young LLP, 1997).

\(^6\) The model in Elitzur and Mintz (1996), in which tax authorities in each jurisdiction apply a different arm’s length method, may cause international double taxation. Discussion of the BAPA framework is a relevant topic since we have observed an increasing number of applicants to the system (Tax Management Inc., 1999, p. 722).

\(^7\) One interpretation could be that the current model focuses on the second stage of a two–stage tax game, where each tax authority chooses tax instruments in the first stage, and a multinational company chooses output in the second stage.
Proof

The result is obtained immediately by applying the implicit function theorem to the first–order condition [2].

\[
\frac{\partial q}{\partial y} = \frac{(\bar{y} - c_{y})}{(1 - \theta)(2p + \frac{d^2p}{dq^2} q)} < 0, \quad \text{and} \\
\frac{\partial q}{\partial \theta} = \frac{p - \bar{y} - c_{\theta} + \frac{dp}{dq} q}{(1 - \theta)(2p + \frac{d^2p}{dq^2} q)} > 0. 
\]

Figure 1 illustrates output adjustments with respect to tax changes. For this purpose, rearrange equation [2] as

\[
\frac{dp}{dq} q + p = c_{h} + t_{h}(\bar{y} - c_{h}) + c_{f} + t_{f}(\frac{dp}{dq} q + p - \bar{y} - c_{f}),
\]

and denote the left–hand side as MR and the right–hand side as MC. Figure 1(a) shows that an increase in \(t_{h}\) makes MC shift to MC' since \(\theta - c_{h} > 0\) and reduces the equilibrium output of the company. Similarly, in Figure 1(b), an increase in \(t_{f}\) shifts MC down to MC' since \((dp/dq)q + p - \bar{y} - c_{f} < 0\) (from the first–order condition [2]). This increases the equilibrium output.

The latter result of Proposition 1, that output increases with \(t_{f}\), is somewhat unexpected at first glance; however, this can happen when the benefits (cost savings from the income deduction) surpass the costs (an increase in tax payments).\(^8\)Rearranging equation [2] helps explain this result:

\[
(1 - t_{h})\bar{y} + (1 - t_{f})\left(p + \frac{dp}{dq} q\right) = (1 - t_{h})c_{h} + (1 - t_{f})(\bar{y} + c_{f}).
\]

The multinational company maximizes after–tax consolidated profits by choosing the output which equates the after–tax marginal revenue (the right–hand side of equation [2.1]) to the after–tax marginal cost (the left–hand side) at the group level. An increase in the host tax rate makes both marginal cost and marginal revenue decrease. The decrease in marginal revenue is due to an increase in tax payments to the host country; the decrease in marginal cost is due to the cost savings from the income deduction. The effect of decreasing marginal cost always dominates the decrease in marginal revenue (from the inequality \((dp/dq)q + p - \bar{y} - c_{f} < 0\)); therefore, it is profitable for the company to increase production.

The logic of this result follows a similar argument to that used for tax avoidance via transfer pricing. The standard rule for transfer pricing is to avoid the tax burden by manipulating a transfer price so that income is allocated to the country with the lower tax rate. In this case, instead of using transfer pricing (which is regulated), a company tries to increase its after–tax global profits by adjusting its output so as to generate more profits in the country with the lower tax rate. This idea is easily understood once the equilibrium output is solved for in closed form. Let us introduce a linear demand function on final products sold in the host market:

\[ p = a - bq, \]

where \(a > 0\) and \(b > 0\). Then the equilibrium output is simplified as

\[
q' = \frac{1}{2b}(\bar{y} - c_{f} + (1 - t')(\bar{y} - c_{h})),
\]

where \(1 - t' = \frac{1 - t_{h}}{1 - t_{f}},\)

(which is a finite quantity since \(t_{f} \neq 1\)).

Equation [3] shows that the relative tax rates in the two countries play a key role.
Figure 1. Output Adjustments with Respect to Tax Changes

a. An increase in $t_h$

b. An increase in $t_f$
in deciding the equilibrium output. This indicates that the level of corporate profits varies depending on how the company allocates the tax burden between these countries via output adjustments.

POLICY IMPLICATIONS

Inefficiency

The equilibrium level of output is not efficient since alternative tax policies have the potential for enlarging before–tax profits of the company (or the tax base) through a more efficient allocation of the tax burden. The source of inefficiency is coordination failure between the two governments when both countries share a tax base. Output adjustments play an intermediate role in creating the efficiency loss.

Let us examine the mechanism in detail by comparing the equilibrium output \( q^* \) to a key benchmark level of output \( q^M \), which maximizes monopoly rents in the absence of any corporate income taxation (i.e., the solution to equation [2] when \( t_f = t_h = 0 \)). Output \( q^M \) is also the solution to the problem of maximizing the sum of after–tax global profits and joint tax revenues in two countries. In other words, the tax base is maximized at \( q^M \) (but not at \( q^* \)). The causes of an efficiency loss can be seen most easily by rewriting equation [1] as

\[
\Pi = (1 - t_f)(p - c_h - c_f)q + (t_f - t_h)(\theta - c_i)q.
\]

The first term on the right–hand side of [4] is after–tax profits when all of the company’s activities are concentrated in a single jurisdiction, with a corporate income tax rate \( t_f \). These profits are maximized at \( q^M \), whatever the foreign tax rate. However, the equilibrium output \( q^* \) is larger than \( q^M \) if (and only if) the coefficient multiplying \( q \) in the second term on the right side of [4] is positive, i.e., \( (t_f - t_h)(\theta - c_i) > 0 \). The company wants to increase output above \( q^M \) in the presence of the difference in tax rates between countries. Similarly, the output \( q^* \) is less than \( q^M \) if and only if \( (t_f - t_h)(\theta - c_i) < 0 \). In both cases, a lower tax base would result compared to the case of \( q^M \). It is straightforward to show that a multinational company will choose to produce the output level \( q^M \) if either \( t_f = t_h \) or \( k = 0 \) and, thus, the second term on the right side of [4] disappears. We rewrite this result below as a proposition.

Proposition 2

The inefficiency arising from domestic tax policies under the BAPA system will be overcome though either a zero mark–up ratio or a uniform tax rate in both the home and the host country in the case of source–based taxation.

The degree of the inefficiency increases with a larger mark–up rate.\(^9\) This is intuitive since setting \( k = 0 \) results in an efficient outcome. A positive mark–up ratio operates as a proxy of a hazard index that measures the degree under which independent tax systems hinder the vertical integration of multinational companies. The companies cannot internalize the costs of intra–firm transactions under tax systems that segregate profits earned by two different affiliates within the same company for the purpose of imposing the tax independently as in equation [1]. A lower mark–up ratio works to eliminate the hazard and increase the profits of the companies. The loss of efficiency under BAPAs is closely analogous to the ineffi-

\[^9\] Denote before–tax profits (i.e., the tax base) of a company as \( \pi \). Let the before–tax profits when the company produces \( q^M \) be \( \pi^M \) and the profits when the company produces \( q^* \) be \( \pi^* \). The extent of the inefficiency related to \( k \) is measured by differentiating the difference between \( \pi^M \) and \( \pi^* \) with respect to \( k \). The difference \( \pi^M - \pi^* \) is increasing in \( k \) from \( \partial(\pi^M - \pi^*)/\partial k = c_f^2 t_f^2 k/2b > 0 \), (assuming that \( t_f = t_h \)).
ciency caused by double marginalization of monopolies in the industrial organization literature.

**Tax Coordination**

The tax base (i.e., before–tax profits) will increase once governments coordinate their tax policies and let the company enjoy the benefits of an integrated organization. One approach is to set the company’s mark–up at zero when deciding on a transfer–pricing rule. Such a transfer price \( \bar{\theta} = \bar{c}_i \) is eminently justifiable. It would be the arm’s length price if there was a competitive market for the intra–firm trade (and constant returns to scale implies that a competitive market could exist). The other approach is to coordinate tax rates. If governments coordinate and set a uniform tax rate, the efficient outcome would be obtained regardless of the level of the mark–up.

Although tax coordination will increase the tax base, the allocation of gain is a separate issue. For example, the host country increases its tax revenue when setting \( k = 0 \), while the home country loses tax revenue. This can be understood from the fact that the transfer price \( \bar{\theta} = \bar{c}_i \) would lead to an equilibrium in which the host country collects all monopoly rents as tax revenue by setting a tax rate of 100 percent. The same implication also applies to the unitary tax rate case \( t_f = t_h \).

The possibility that one country may lose from tax coordination is similar to results arising in different models of capital tax competition (Bucovetsky, 1991; Wilson 1991). In this case, it is natural that the home government would seek to claim a portion of the increased revenues as its due from coordinating tax policy. Although a formal model of the cooperative game is beyond the scope of this paper, the coordination needs to involve transfer payments between the two countries. Given that such transfer payments are not general practice under the current world tax system (where the source of taxation rights relies on national sovereignty), the result suggests that the system should be integrated to accommodate a flexible revenue allocation across different countries.

A key to attaining an efficient outcome is the method used for revenue allocation. The mark–up rate chosen in a BAPA determines how before–tax profits are split between governments (and a company). In making this point, the analysis provides useful guidance for the coordination of tax agreements. An alternative approach is the profit–split method, in which a company and governments allocate a portion of ex post before–tax profits between jurisdictions based on a contribution rate attributed to each affiliate: each government levies taxes on the allocated tax bases at its own tax rate. Denoting \( g \) as a contribution rate attributed to an affiliate in the country \( f \), the after–tax global profits are expressed as:

\[
\Pi = (1 - gt_f - (1 - g)t_h)(p - c_h - c_f)q.
\]

The company will choose to produce the output level \( q^{M} \), given a positive constant rate \( g \) between zero and one.\(^{10}\) Contrary to a regulated transfer price (the arm’s length methods), the profit split method can avoid a tax distortion despite different tax rates across jurisdictions.

**Applications**

The Resale Price Method

This analytical framework is applicable to other interesting situations. One possible extension is to explore the effects of other transfer pricing methodologies.

---

\(^{10}\) This argument needs to be elaborated once \( g \) is chosen based on the levels of inputs, which increase in proportion to sales.
Suppose a multinational company and two governments agree to use a normal profit rate \( l > 0 \) (i.e., the resale price method) in the BAPA system. Substituting the transfer price \( \theta = (1 - l)p - c_j \) into equation [1] and rearranging it allows us to express the after–tax global profits of the company as

\[
(1 - t_h)(p(q) - c_j - c_h)q + (t_f - t_h)p(q)q.
\]

Equation [6] corresponds to equation [4]. The company will choose to produce the output level \( q^M \), at which the tax base is maximized, if either \( t_f = t_h \) or \( l = 0 \). The transfer price \( \theta = p - c_j \) would lead to an efficient equilibrium.

Worldwide Taxation with a Tax Credit

The production inefficiencies would not arise under worldwide taxation with a tax credit. Suppose the home country (where the headquarters of a multinational company locates) employs the worldwide tax system with a foreign tax credit. Tax payments at home and abroad are expressed as \( (t_h(\theta - c_h) + (t_h - t_f)(p - \theta - c_j))q \) and \( t_f(p - \theta - c_j)q \), respectively. The company can credit taxes paid in the foreign host country up to the level of taxes that the company is supposed to pay in the home country. Let us consider the case where the home tax rate is greater than or equal to the foreign tax rate, \( t_h \geq t_f \). The company can credit all taxes paid in the host country (i.e., no excess credit). Suppose all income earned abroad is immediately repatriated to the home country. After–tax profits of the company are expressed as

\[
(1 - t_h)(p - c_h - c_j)q.
\]

Equation [7] is similar to equation [5]. The company will choose to produce the efficient output level \( q^M \).

CONCLUDING REMARKS

This paper examines the inefficiency of multinational companies’ output decisions arising from domestic tax policies under the BAPA system. I use a model in which transfer pricing issues are absent and the company responds to taxation with an output decision. The analysis shows that inefficiency will arise under the BAPA system since jurisdictional tax policies distort the production decision of the multinational company. When the markets across countries are interrelated through intra–firm trade of multinational companies, the system segregates profits earned by different affiliates within the same company for the purpose of imposing the tax independently. The companies cannot internalize the costs of intra–firm transactions under such a tax system.

The analysis provides governments with useful guidance for coordinating tax agreements. Either a uniform tax rate or a zero mark–up ratio can eliminate the segmentation between the two affiliates and allow the company to earn more profits by internalizing the cost of the intra–firm transaction. The larger tax revenue results from the increase in before–tax profits after the tax distortion on production is eliminated.

This analytical framework is applicable to other interesting but more complicated situations. One possible extension is to consider regimes of multi–agents such as regions, nations, or government agencies. In reality, policy effectiveness is often limited without cooperative planning and implementation among multiple agents because some agents could free ride on the benefits obtained by partial policy coordination. This paper does not cover the case of multi–agents due to its analytical difficulty; however, a numerical simulation would allow further analysis. Another extension is to include the uncertainty of policy commitment. This is a serious problem in developing countries, where governments frequently change their tax policies. Though
these extensions may change some of the results obtained in this paper, the main conclusion will still be valid, since BAPAs cause an efficiency loss in the globally interrelated market through distorted production decisions. All of these topics represent potential future lines of research.

Acknowledgments

I gratefully acknowledge suggestions by Li Ming Dong, Joseph Harrington, Edi Karni, Jennifer Roff, Katie Winder, and Peyton Young on several points of an earlier draft. I would also like to thank seminar participants at Columbia University, the City University of New York, and Johns Hopkins University for their comments. This paper has benefited substantially from comments and suggestions by Rosanne Altshuler and two anonymous referees. Financial support from the World Bank in the form of a dissertation fellowship is gratefully acknowledged. All errors are mine.

REFERENCES


Boadway, Robin, and Michael Keen.

Boadway, Robin, Maurice Marchand, and Marianne Vigneault.

Bond, Eric W., and Thomas A. Gresik.

Bond, Eric W., and Larry Samuelson.

Bucovetsky, Sam.

Copithorne, L. W.


Eden, Lorraine.

Elitzur, Ramy, and Jack Mintz.

Ernst & Young LLP.

Greenhut, M. L., and H. Ohta.

Gresik, Thomas.


Hamaekers, Hubert.
Horst, Thomas.  

Janeba, Eckhard.  

Kant, Chander.  

Keen, Michael, and Christos Kotsogiannis.  

Mieszkowski, Peter, and George Zodrow.  

Prusa, Thomas J.  

Stoughton, Neal, and Eli Talmor.  

Tax Management Inc.  

Wildasin David E.  

Wildasin David E.  

Wilson, John D.  

Wilson, John D.  

Zodrow, George R., and Peter Mieszkowski.  