Gradualism in Tax Treaties with Irreversible Foreign Direct Investment

Richard Chisik[*]

and

Ronald B. Davies[**]

June 2000

Revised: June 2002

Abstract: Bilateral international tax treaties govern the host country taxation for the vast majority of the world's foreign direct investment (FDI). Of particular interest is the fact that the tax rates used under these treaties are gradually falling although the treaties themselves do not specify any such reductions. Since there is no outside governing agency to redress treaty violations, such reductions must be both mutually beneficial and self-enforcing. Furthermore, the optimal tax rates must be less than those initially set, otherwise no reductions would be necessary. To explain such behavior, we model a two-country setting with two-way capital flows. In particular, only part of FDI is immediately reversible. As the extent of irreversibility increases, the likelihood of Pareto optimal tax rates obtaining as a self-enforcing outcome in the initial period is reduced. More modest tax reductions, from the non-treaty levels, are still possible. These limited tax reductions generate an increase in bilateral FDI. As countries increase the stock of capital in one another, further reductions in taxes become self-enforcing. Depending on the extent of irreversibility and asymmetry, Pareto optimal tax rates may be obtainable in the long run. Thus, the amount of inbound investment a country can attract may be related to the commitment to which its outbound investment binds it. This final insight provides an additional rationale for the observed pattern of capital flows in which those countries with the greatest outbound capital flows are also those with the highest inbound flows.

JEL Classification: F21, F23, F13, C73.

Key Words: Foreign Direct Investment, Tax Treaties, Multinational Enterprise, Gradualism, Irreversibilities, Dynamic Games.

[*]Department of Economics DM-309C, Florida International University, Miami, FL, 33199; Phone: (305) 348-3286; Fax: (305) 348-1524, E-mail: chisikr@fiu.edu.
[**]1285 University of Oregon, Department of Economics, 533 PLC Building, Eugene, OR, 97403-1285; Phone: (541) 346-4671; Fax: (541) 346-1243; E-mail: rdavies@oregon.uoregon.edu.
We thank seminar participates at Georgetown University, the University of Miami, Oregon State University, the University of Western Ontario, the University of Copenhagen, the University of Kentucky, the 2001 Canadian Economic Association meetings, the Spring 2001 Midwest International Economics Group meetings, the Winter 2001 Econometric Society meetings, the Fall 2000 Southeast Theory and International Economics meetings, All errors are entirely the responsibility of the authors.

I. Introduction

The use of dynamic tax incentives to attract foreign direct investment (FDI) is well known. The vast majority of economic studies, however, examine only the use of tax holidays in which a host government offers a firm an initial period of reduced tax rates after which taxes increase to their standard levels. While this is a powerful tool for attracting FDI, other types of tax incentives are also used. One of the most common alternative methods is the bilateral tax treaty, which is the focus of this paper.

A global network of nearly 2000 bilateral tax treaties governs international taxation (Radaelli, 1997). In 1998, such treaties with the U.S. covered approximately 78 percent of total U.S. investment abroad and 96 percent of FDI within the U.S. (BEA, 1998). Similar to tax holidays, bilateral treaties are intended to promote investment between treaty partners, however they differ from tax holidays in two key respects. First, treaties are negotiated between two governments, not a government and a single firm. As such, they apply to all investment between the two countries. Second, rather than "front-loading" incentives as tax holidays do, treaty tax rates tend to gradually fall over time. While this difference in the timing of tax reductions is interesting in and of itself, what is particularly intriguing is that rather than a simple one-time, permanent reduction in taxes, treaty tax rates fall incrementally. Our goal is to offer one explanation for this gradual reduction in tax rates over time.

While tax treaties provide several avenues for FDI promotion, we will concentrate on their effect on withholding taxes. When a firm invests overseas, it typically does so through a subsidiary that repatriates profits to its parent through dividends, interest, and royalty payments. Since these payments are a cost to the subsidiary, they are not taxable by the host country as part of the subsidiary's income. Nevertheless, most host governments capture part of this parent income through a withholding tax levied on the repatriation. For example, the standard, non-treaty U.S. withholding tax on dividends, interest payments, and royalties is 30 percent. Under a tax treaty, each government agrees to lower the withholding tax rates it applies

to payments made to recipients who reside in the partner country. Specifically, tax treaties reduce withholding taxes by specifying maximum allowable tax rates. These maximum rates are as low as zero and are almost always identical for both treaty partners.[1] Similar to the tariff reductions specified in many bilateral and regional trade agreements (such as NAFTA, Mercosur, and the U.S.-Israel free trade agreement) reductions in taxes often occur incrementally. Some sequential tax reductions are specified by timetables written into the treaties themselves, while others are achieved through renegotiation. Table 1 provides examples of both types of sequential reductions. As Table 1 shows, the first U.S.-Canadian treaty was signed in 1941 and implemented a royalty withholding tax of 15 percent. This rate was lowered to 10 percent during renegotiations in the late 1970's. According to the Price-Waterhouse *Corporate Taxes: A World-Wide Summary*, this tax was reduced to zero in 1997, although the treaty itself does not yet reflect this change. Additionally, the parental dividend tax fell from 15 to 10 to 5 percent over the same period.

TABLE 1 GOES APPROXIMATELY HERE

As stated in the introduction to the OECD's model tax treaty, the goal of a treaty is to "remove the obstacles that double taxation presents to the development of economic relations between countries" (OECD, 1997, pg. I-1). Since reducing tax rates may mitigate the investment distortions due to tax policy, these reductions can be efficiency-improving. However, if lower tax rates are preferable, why are the lowest rates not chosen at the outset of the treaty? One answer is that they are not self-enforcing. Since there is no external body that can enforce tax treaties, a treaty must be self-enforcing as well as mutually beneficial.

Key to our explanation is the irreversible nature of bilateral FDI. A treaty that specifies a large tax reduction, and thus a large inflow of irreversible FDI, may create more temptation than a country can resist. This leads the violator to tax the irreversible inbound investment heavily and

[1] For information on the specifics of treaties, see the treaties themselves which are collected by Diamond and Diamond (1998).

to minimize outward investment so that it has less FDI trapped overseas which can be taxed in retribution. Alternatively, a small tax decrease can be self-enforcing for both countries. Since this tax reduction applies to both nations' investors, it increases FDI in both directions. Now, if a country deviates, it has more captive FDI in its partner, allowing the partner to exact a harsher punishment if a violation occurs. Additionally, when FDI is less reversible, this increases the length of the punishment phase in which the other country can implement a punitively high tax rate. These two effects imply that a country with a relatively large amount of outbound FDI is less likely to deviate from a tax treaty and will find a larger set of tax rates to be self-enforcing. A further implication is that a reduction in tax rates in the current period increases the current amount of outbound FDI, expanding the following period's set of self-enforcing tax rates. Thus, one reduction in tax rates makes additional reductions self-enforcing. Therefore, through gradual tax reductions, countries can approach an efficient allocation of capital even though they could not credibly commit to such a policy at the treaty's outset.

We formalize the above argument by presenting a model of two countries with bilateral FDI. Without a treaty, governments cannot credibly commit to the efficient tax rates, resulting in an inefficient equilibrium with high tax rates and low levels of FDI. A tax treaty improves on this outcome by allowing governments to coordinate on a pair of Pareto improving tax rates. Because of the initially low levels of FDI, Pareto-optimal tax rates may not be self-enforcing at the outset of the treaty. Despite this, as long as investment is not entirely irreversible and governments care sufficiently about the future, some Pareto improving tax reduction is credible. This leads to a small increase in FDI and expands the set of self-enforcing tax rates. This makes additional tax reductions possible and implies that through gradually falling tax rates countries can achieve mutually beneficial tax rates that were not sustainable in the early stages of the treaty. If countries are sufficiently asymmetric or if investment is sufficiently irreversible, then there exists an inefficient limit to the tax reductions, however, this cooperative outcome is still a Pareto improvement over the non-treaty equilibrium.

Although we are unaware of any work studying gradualism in tax treaties, there are obvious parallels between this gradualism and that found for tariffs under self-enforcing trade agreements. Staiger (1995), Furusawa and Lai (1998), Devereux (1997), and Chisik (forthcoming) all analyze gradual tariff reductions that are determined endogenously along with the evolution of an economy-wide state variable. In an approach similar to ours, Chisik finds that an initial reduction in tariffs leads to an increase in irreversible country-specific export capacity. This increase in capacity increases the cost of a trade war and makes further tariff reductions possible. This idea of partner-specificity is particularly appealing for FDI, since by the nature of its physical location, it is partner-specific. When combined with the bilateral FDI tax competition model of Davies (forthcoming), this approach works well to describe gradualism in tax treaties.

As noted above, tax treaties and tax holidays differ in both the negotiating parties (government-government versus government-firm) and the timing of tax reductions. These differences are best illustrated by contrasting our results with those of Thomas and Worrell (1994).[2] In their paper, the host government is unable to credibly commit to low taxes on inbound FDI, leading to inefficiently low levels of investment. Nevertheless, it is incentive compatible for the host to offer the firm a tax holiday in early periods to induce additional investment that it can then later tax.[3] In their unidirectional framework with one host and one firm, the only possible punishment for a host that breaks an agreement is the withholding of

[2] Bond and Samuelson (1989a) and Doyle and Van Wijnbergen (1994) also consider the relationship between tax holidays and the obsolescing bargaining problem, however, they do not discuss gradualism. Eaton and Gersovitz (1984) suggest that when a country would prefer to commit but cannot, the threat of punishment by a foreign country can improve a country's credibility and raise its welfare. We formalize this idea by explicitly considering how the foreign punishment endogenously evolves in the context of tax treaties. Bond and Samuelson (1986) offer a different rationale for tax holidays in an environment in which two countries with private information about a country-specific productivity parameter compete for FDI. They find that the country with the higher productivity can signal this private information to the firm by offering a longer-lasting tax holiday. Aizenman (1996) does not consider gradualism or tax holidays, but does include FDI in a model of managed trade. He finds that since FDI is a substitute for exports, FDI can reduce the time inconsistency problem in trade negotiations.

further investment. For this reason, the tax policy is backloaded: taxes are initially low and increase over time. FDI, meanwhile, increases over time (although it remains inefficiently small). Their unidirectional framework captures the dynamics of tax holidays along with the phenomenon of gradually increasing investment. Our bi-directional framework, on the other hand, finds that under tax treaties bi-directional FDI gradually increases while tax rates gradually *decrease* over time. Furthermore, our approach shows that bilateral investment irreversibility may help to ameliorate the time inconsistency problem that is evidenced in the unidirectional framework and, therefore, brings in to question this rationale for tax-holidays.

We proceed as follows. The next section presents the model and derives the equilibrium without a tax treaty. Section III introduces tax treaties and discusses their implications for the path of tax rates over time. Section IV considers asymmetries between the treaty partners. Section V concludes.

II. A Simple Model of Bilateral FDI and Taxation

A. Investors

Consider two countries, home and foreign, each of which are endowed with inelastic capital stocks, K and K^*, respectively.[4] In each period, investors can invest either at home or overseas. The amount of home (foreign) FDI in foreign (home) during period t is denoted Z_t (Z_t^*). Home capital that remains at home is used to produce a good according to the production function $h(K-Z_t)$. Home capital located in foreign produces the same good using the production function $h^s(Z_t)$. The price of this good is constant and normalized to one. Both production functions satisfy the Inada conditions and are increasing and strictly concave in capital.

---

[3] Bond and Park (forthcoming) utilize a similar methodological approach to show that gradualism may arise as a symptom of time-inconsistency in tariff setting by a large country which trades with a small country when the small country desires to smooth consumption over time. In contrast with the other papers in the gradualism literature Bond and Park and Thomas and Worrell demonstrate the existence of gradualism even in the absence of an endogenously evolving state variable.

[4] Alternatively, we could permit capital to grow at some exogenous rate. This makes FDI more attractive over time, but does not change the intuition regarding the role of irreversibility. Therefore, for the sake of brevity, we retain the assumption of fixed capital stocks.

Investment irreversibility is captured by $\gamma \in [0, 1]$ and the constraint that $Z_t \geq \gamma Z_{t-1}$. This irreversibility can represent the time to dismantle or sell off assets, the effort of training immobile factors of production such as foreign workers, or the cost of tailoring production to the overseas market. If $\gamma = 0$, then all FDI is immediately reversible. If $\gamma = 1$, it is impossible to recall overseas investment. The irreversibility parameter is exogenous.[5]

A more general interpretation of the model is to replace $h(K-Z_t)$ with a function of Z_t, $\tilde{h}(Z_t)$, where this function simply represents the increasing and convex cost of FDI.[6] This interpretation is particularly attractive for some of the most irreversible types of FDI such as tourist facilities or natural resource extraction. In this case, the irreversibility parameter reflects the difficulty of liquidating this investment. Alternatively, if investors have access to a limited amount of resources, then this cost can represent the opportunity cost of dedicating physical and human capital to investment projects in the foreign country. Although we discuss this as a drop in domestic output, it could just as easily represent the decrease in subsidiary activity in other (non-modeled) countries. For example, if production location decisions, plant development, and worker training have to be planned far in advance, an increase in a subsidiary's output implies reducing output (or not making expansions) elsewhere by the firm. A third interpretation is that a recurring direct cost of FDI may represent the cost of acquiring funds from global capital markets. Here, an increasing marginal cost of capital may be particularly appropriate for a firm with large overseas investments since overseas projects may be considered more vulnerable to both exchange rate and expropriation risks as well as agency problems. A final interpretation of

[5] Although we make the irreversibility parameter exogenous, this need not be true in practice. For example it may be possible for investors to choose alternative technologies or split their investment between easily-reversible portfolio investment and less reversible direct investment. While this case is highly intriguing, given the complexity of the current model we leave this extension to future research.

[6] It is not necessary that this cost is strictly convex in the investment level, Z_t, (i.e. that $h(K-Z_t)$ is strictly concave), for it is possible that this cost merely reflects the constant cost of obtaining capital on world markets. For an interior solution we do require strict concavity of the benefit (or strict convexity of the cost) of investing. If $\tilde{h}(Z)$ truly reflects a capital cost, then it is important to note that we are assuming that the cost is paid every period rather than a one time cost.

$\tilde{h}$ is simply an increased cost of global coordination by the firm. Thus, although we explicitly model the cost of FDI as the opportunity cost of home production forgone, it is not necessary to include domestic production in our model. Our adopted approach does present an advantage in that it allows us to easily consider one potential difference between treaty partners - their available capital stocks. In any case, we only require that FDI comes at some cost and that when reversibility is not perfect some portion of this cost must still be borne by investors.

Similarly, foreign investors split their capital between foreign production, $f(K^*-Z_t^*)$, and production in the home country, $f^s(Z_t^*)$, under analogous assumptions regarding the behavior of the production functions and irreversibility.[7] The price of the foreign good is also inelastic and equal to one. It is important to note that there is no substitutability between the home and foreign capital stocks, indicating that capital is specific to the good produced by each country. This would be the case if the differentiated products embody proprietary technology, patents, brand names or country-of-origin reputations. [8] According to Caves (1996), these are common features of multinational corporations.

FIGURE 1 GOES APPROXIMATELY HERE

The timing of the model is summarized in Figure 1. At the beginning of each period, governments simultaneously announce tax rates to domestic and overseas investors. Although it is not necessary for our results, we permit governments the ability to announce different tax rates to each group of investors. Thus, a government can warn its investors in advance of a planned

[7] For now we assume that $\gamma = \gamma^*$ to conserve on notation. We explicitly consider the case in which $\gamma \neq \gamma^*$ in section IV.C. below.

[8] As Markusen and Venables (1998) show, such two-way capital flows can arise even between identical countries in a differentiated products industry. This non-substitutability assumption is not necessary under our alternative interpretation of $h(K-Z_t)$ as a cost.

deviation from the cooperative treaty tax rates.[9] Such a warning could be an explicit signal to investors, or it could simply be the result of home investors being more knowledgeable about the political climate in the home country than are the relatively distant foreign investors. The home (foreign) government's announced tax rate to the home (foreign) investors in period t is τ_t^A (τ_t^{*A}). The home (foreign) government's announced tax rate to foreign (home) investors is τ_t^a (τ_t^{*a}). Although we permit the possibility, we show below that neither government will announce anything other than the cooperative tax rate to overseas investors. Following tax rate announcements, investors simultaneously commit capital allocations for the period (subject to their irreversibility constraints). After capital allocations are made and observed by both governments, both countries set actual tax rates (τ_t, τ_t^*) and production values are realized.[10] The ability to revise tax rates from their announced levels implies that tax rates can be changed faster than investment decisions. Given the observed efficiency of most governments, this is perhaps the most fanciful of our assumptions.

Home investors maximize the present value of lifetime profits by choosing a stream of FDI. These profits, where $\delta \in [0, 1]$ is the discount factor and $E(\tau_t^*)$ is the expected foreign tax rate in time t, are given by:[11,12,13]

---

[9] The ability to notify one's own investors before a deviation allows the deviating country to begin reducing its FDI earlier than the surprised country. We believe that this possibility for coordination between the government and its own investors is the more natural case to consider. On the other hand, if governments do not (or cannot) warn their own investors, then no investors can respond until the period following the deviation. In Appendix B, we analyze when the government would choose not to inform their own investors and we show that gradualism still occurs.

[10] This timing is similar to that used by Lapan (1988), McLaren (1997) and Bond and Park (forthcoming) in their discussion of trade agreements. As in those studies, this timing creates the potential for time inconsistency in government policy. The important difference, here, is that bilateral investment irreversibility can mitigate the time inconsistency problem.

[11] Although our model is completely deterministic, investor's decisions in the current period take into account their expectation of future tax rates. Thus, $E(\tau^*_t)$ is the anticipated future foreign tax rate which is conditional on the current period's history rather than on some stochastic process.

[12] Note that home investors pay no home tax. Since the home government has no incentive to distort home investor decisions, this is optimal from the home government's perspective. As long as the home government can set discriminatory taxes on foreign investors operating within its borders, such an optimum is obtainable. Hines (1988) and Hufbauer (1992) provide evidence of such discrimination.

$$\sum_{t=1}^{\infty}\delta^t\pi(K, Z_t, \tau_t^*) = \sum_{t=1}^{\infty}\delta^t[h(K - Z_t) + (1 - E(\tau_t^*))h^s(Z_t)]. \quad (1)$$

The irreversibility constraint requires that:

$$Z_t \geq \gamma Z_{t-1}. \quad (2)$$

Investors are forward-looking and allocate their capital taking into account both expected future tax rates and how their decisions affect government actions (such as deviations from a treaty). As shown below, if foreign taxes are not expected to increase in the future, (2) is non-binding. With this in mind, as long as the expected foreign tax rate is less than one, the home investor will pick a positive level of investment, $Z(K, E(\tau_t^*))$ which equates the domestic and overseas after-tax rates of return. This yields the home capital market equilibrium condition in period t:

$$(1 - E(\tau_t^*))h_k^s(Z_t) - h_k(K - Z_t) = 0. \quad (3)$$

If taxes are expected to rise at some point in the future, then a condition similar to (3) can be derived. Since efficiency-improving tax treaties will exhibit non-increasing tax rates, this result is not presented.[14] An interior solution to equation (3) yields to the following comparative statics (suppressing the arguments of functions):

$$\frac{dZ_t}{dE(\tau_t^*)} = \frac{h_k^s}{(1 - E(\tau_t^*))h_{kk}^s + h_{kk}} < 0; \quad (4)$$

$$\frac{\partial Z_t}{\partial K} = \frac{h_{kk}}{(1 - E(\tau_t^*))h_{kk}^s + h_{kk}} \in (0,1). \quad (5)$$

---

[13] Equation (1) indicates that profits are repatriated at the end of the period in which they are earned. Since production is a function only of capital which is inelastically supplied, there is no incentive for the subsidiary to retain its earnings. Hartman (1985) considers a case in which a subsidiary can be financed either through retained earnings or capital flows from the parent. Because retained earnings avoid withholding taxes, this represents a cheaper form of finance than parent equity. Thus, Hartman suggests that withholding taxes will only influence the behavior of mature, i.e. non-expanding, subsidiaries. If taxes do not influence FDI, then there would be little reason to negotiate a treaty. Therefore, we have chosen what we feel is the simplest model of subsidiary finance in which taxes affect FDI.

Blonigen and Davies (2000) find empirical evidence of a negative relationship between withholding taxes and FDI for inbound U.S. investment. This result indicates that as long as tax rates do not rise in the future, investors will not choose to retract investment, implying that the irreversibility constraint is non-binding. Blonigen and Davies, Brainard (1997), and others have found a positive correlation between FDI and a country's available funds.

Given the similarities between the home and foreign investors, if home taxes are not expected to rise in the future, then an interior equilibrium in the foreign capital market in period t is given by:

$$f_k(K^* - Z_t^*) = (1 - E(\tau_t))f_k^s(Z_t^*). \quad (6)$$

A comparable set of comparative statics can be derived.

B. Governments

Each government chooses its tax rate in each period to maximize the present value of its national income subject to the capital market equilibrium conditions (3) and (6), irreversibility constraints (2), and investor expectations. Denote the actual home and foreign tax rates in each period by τ_t and τ_t^*. Tax rates are constrained to the unit interval. Governments share the same discount rate as investors. This implies that home and foreign national incomes are:[15]

$$\sum_{t=1}^{\infty}\delta^t I_t = \sum_{t=1}^{\infty}\delta^t I(K, K^*, Z_t, Z_t^*, \tau_t, \tau_t^*) = \sum_{t=1}^{\infty}\delta^t[h(K - Z_t) + (1 - E(\tau_t^*))h^s(Z_t) + \tau_t f^s(Z_t^*)], \quad (7)$$

and

$$\sum_{t=1}^{\infty}\delta^t I_t^* = \sum_{t=1}^{\infty}\delta^t I(K^*, K, Z_t^*, Z_t, \tau_t^*, \tau_t) = \sum_{t=1}^{\infty}\delta^t[f(K^* - Z_t^*) + (1 - E(\tau_t))f^s(Z_t^*) + \tau_t^* h^s(Z_t)]. \quad (8)$$

As equations (7) and (8) indicate, taxes both redistribute income between countries and distort capital flows. World income in each period, which is invariant to redistributions between

[14] If expected future tax rates are lower than the current expected value, then there is still no incentive to engage in investment in excess of what (3) dictates. This occurs since additional investment can be put off until the expected tax reduction, avoiding the cost of presently doing so. If there were some cost to installing FDI, this need not be the case.

countries, is maximized when tax rates are zero. Define these optimum FDI levels as $Z^{opt} = Z(K, 0)$ and $Z^{*opt} = Z^{*}(K^{*}, 0)$. As we show in Proposition 1 below, these optimal investment levels and the maximization of world income cannot be obtained without a tax treaty.

C. Equilibrium in the Absence of a Tax Treaty

As a benchmark case, consider the equilibrium that arises in the absence of a tax treaty. Without a tax treaty, governments do not condition their action on history and choose myopically optimal tax rates based solely on the current period FDI levels. These Markov strategies yield a particularly stark outcome. Because FDI is fixed in the third stage of every period, the best response for both governments is to fully tax away the profits on any inbound FDI. Correctly forecasting this third stage outcome, investors do not add any FDI in the second stage. As a result of this time inconsistency problem, the government's announced tax rate in the first stage is non-credible and immaterial in the resulting equilibrium. Furthermore, because no FDI is added, the physical environment does not change in the following period and the same inefficient outcome is obtained. A subgame-perfect equilibrium in these Markov strategies is a Markov-perfect equilibrium (MPE). It is the unique equilibrium in the absence of a tax treaty and it is described in Proposition 1.[16]

---

[15] The assumption of national income maximization follows Bond and Samuelson (1989b), Janeba (1995) and others.

[16] Alternative assumptions on belief structures, irreversibility, or timing can temper the severity of the non-treaty tax rates. Bond and Samuelson (1989a), Doyle and Van Wijnbergen (1994), and Thomas and Worrall (1994) demonstrate the existence of positive FDI equilibria even without a tax treaty. When the host government does not hold all the bargaining power after investment occurs, their gain from continued future investment can outweigh their gain from short-term expropriation. This desire for reputation provides credibility to an announced tax rate less than one. It will still be the case, however, that investment is inefficiently low, leaving room for a tax treaty. Greif, Milgrom, and Weingast (1994) provide a fascinating historical analysis of how medieval merchant guilds provided the multilateral reputation incentives that reduced the desire to fully tax away the profits on any inbound FDI. Incorporating this initial unilateral FDI would not change our main result that increases in bilateral FDI can generate lower tax rates, however, it can alter the initial dynamic of a tax treaty. (See footnote 19 below.) Alternatively, as Eaton and Gersovitz (1984) point out, credibility can be gained through side agreements such as a trade treaties. These provide an avenue for punishment similar to the one we discuss.

Proposition 1: *(a.) In the unique MPE of the entire game,* $Z_t^m = Z_t^{*m} = 0$ *and* $\tau_t^m = \tau_t^{*m} = 1$ *for all t. (b.) There is a unique MPE in the continuation game starting in period s:* $\tau_t^m = \tau_t^{*m} = 1$ *for all* $t \geq s$ *and* $Z_t^m = \gamma^{t-s+1} \cdot Z_{s-1}$, $Z_t^{*m} = \gamma^{t-s+1} \cdot Z_{s-1}^*$ *for t = (s, s+1,…).*

Proof: Consider the home country in an arbitrary period t. Their optimal choice of τ_t in the third stage of period t maximizes (7). The first order condition for τ_t is:

$$\frac{\partial I_t}{\partial \tau_t} = f^s(Z_t^*) + \tau_t f_k^s(Z_t^*) \frac{\partial Z_t^*}{\partial \tau_t}.$$

Note that Z_t^* is fixed in stage three, therefore, $\partial Z_t^* / \partial \tau_t = 0$. This implies that home will set its tax rate as high as possible, i.e. $\tau_t = 1$. Knowing this, foreign investors will minimize FDI, indicating that $Z_t^* = \text{Max}\{0, \gamma Z_{t-1}^*\}$. It is straightforward to verify that this is the only equilibrium. Similarly we can show that $\tau_t^* = 1$ is optimal for the foreign country implying that $Z_t = \text{Max}\{0, \gamma Z_{t-1}\}$. Finally note that the above result holds for any period t.

The first part of Proposition 1 shows that without a tax treaty no FDI occurs. Part b of the proposition is a direct extension of part a. It indicates that if an existing tax treaty is ever voided and nations return to competitively set tax rates, then the unique outcome is similar to the austere no-treaty outcome: taxes are set at their maximal level, and investors remove FDI as fast as the irreversibility constraint permits. Since the equilibrium without a treaty is clearly inefficient, we now depart from this benchmark case to consider the role of tax treaties in ameliorating this dire equilibrium.

III. Gradualism in Cooperative Tax Rates

An efficiency-improving tax treaty specifies a path of tax rates $\{\tau_t^c, \tau_t^{*c}\}_{t=1}^{\infty}$ that does not move farther away from the Pareto efficient zero tax rates as time progresses. Given the lack of an international enforcement agency, we concentrate on those treaties that are self-enforcing in

each period. We also assume that the only instrument that each government has at its disposal is its tax rate.[17] Finally, since almost all treaties specify identical withholding taxes for both treaty partners, we impose the additional constraint that both countries levy the same tax rate, $\underline{\tau}_t$. This constraint will not be binding in the symmetric case considered in section III. In section IV we consider asymmetries between the treaty partners and show how these differences, coupled with the same tax rate constraint, can affect the treaty outcome.

A. Tax-Treaty Strategies

A common form of self-enforcing mechanism relies upon "grim" strategies that mandate that any deviation from the specified cooperative action generates an infinite reversion to a punishment stage. A credible punishment threat relies on continuation payoffs from a perfect equilibrium (often the most undesirable static Nash equilibrium). The evolution of the state variable, here, precludes reversion to a static Nash equilibrium. A natural counterpart in dynamic games such as this one where past play can alter the current period state variable is an infinite reversion to the worst MPE for the offending party.[18] From Proposition 1, we know that there is a unique MPE following any deviation and we use this MPE in describing the subgame-perfect tax-treaty strategies. These tax-treaty strategies imply that if a deviation occurs, then governments return to their non-cooperative tax rates.[19]

To facilitate a formal representation of these tax-treaty strategies we introduce the following notation. The history through period t is $H^t = \{\tau^t, \tau^{*t}, Z^t, Z^{*t}\}$, where $\tau^t = \{\tau_0, \tau_1, ..., \tau_t\}$

[17] Although side payments, or concessions on alternative international agreements, may constitute a portion of all international negotiations, we follow the tax treaties themselves which make no mention of side payments.

[18] For early examples of such punishments in dynamic games with an evolving state variable, see papers by Ausubel and Deneckere (1987), Cave (1987) and Radner and Benhabib (1992).

[19] A self-enforcing, individually-rational treaty requires that the payoff to the treaty exceeds not only the payoff to deviating, but also the payoff to not entering the treaty in the first place. Under our assumption of no initial FDI (see footnote 16), any incentive-compatible treaty is also individually rational. Under alternative assumptions, the non-treaty equilibrium still requires high positive tax rates, however, they need not equal one. The initial self-enforcing tax rates analyzed in Propositions 2, 3 and 4 may, therefore, exceed the non-treaty rates and fail to be individually rational. As Proposition 5 suggests, this may be a particular difficulty for highly asymmetric countries.

is a vector of past home tax rates, $Z^t = \{Z_0, Z_1, ..., Z_t\}$ is a vector of past home investment, and τ^{*t} and Z^{*t} are the foreign counterparts. When adhering to the tax treaty these taxes and FDI levels take on their cooperative values: $\tau^{ct} = \{ \tau_0, \tau_1^c, ..., \tau_t^c \}$ and $Z^{ct} = \{ Z_0, Z_1^c, ...Z_t^c \}$. To see explicitly the conditioning of these strategies upon history, denote θ_{t+1} as the function that maps histories onto the government's action in period t+1: $\theta_{t+1}(H^t, Z_{t+1}^*, \tau_{t+1}^{*a}) \rightarrow \tau_{t+1}$. Similarly the home investor's action is $\zeta_{t+1}(H^t, \tau_{t+1}^A, \tau_{t+1}^{*a}) \rightarrow Z_{t+1}$. The tax-treaty strategies are defined as follows:

$$\theta_{t+1} = \tau_{t+1}^c \text{ if } H^t = \{\tau^{ct}, \tau^{*ct}, Z^{ct}, Z^{*ct}\},\ Z_{t+1}^* = Z_{t+1}^{*c} \text{ and } \tau_{t+1}^{*a} = \tau_{t+1}^{*c}$$

$$\tau_{t+1}^m = 1 \text{ otherwise;}$$

$$\zeta_{t+1} = Z_{t+1}^c = Z(K, \tau_t^{*c}) \text{ if } H^t = \{\tau^{ct}, \tau^{*ct}, Z^{ct}, Z^{*ct}\},\ \tau_{t+1}^A = \tau_{t+1}^c \text{ and } \tau_{t+1}^{*a} = \tau_{t+1}^{*c}$$

$$Z_{t+1}^m = \gamma Z_t^c \text{ otherwise.} \qquad (9)$$

These tax-treaty strategies are straightforward. If the tax treaty has been adhered to in the past, and if it appears that the foreign country intends to adhere to it in the current period, then the home country sets its current taxes according to the tax treaty. After any other history they abandon the treaty. Investors have similar strategies. If there have been no deviations in the past and there is no indication that either government intends to deviate from the treaty in the current period, then home investors invest according to the expected treaty tax rate. Similar tax-treaty strategies can be described for the foreign country. In the presentation that follows we focus, without loss of generality, on the incentives of the home country.

The foreign government may inadvertently be alerted to a deviation in period d in two ways. First, if home announces something other than τ_d^c to the foreign investors, all agents correctly anticipate both that $\tau_d = 1$ and that the agreement will collapse in period d+1. This leads all investors to immediately begin reducing FDI and causes the foreign government to break from their announcement and set $\tau_d^* = 1$. Since such a move by home reduces both their overseas

profits in the deviation period and their tax revenues, this is clearly not an optimal strategy on home's part. Therefore, the home country always announces the cooperative tax to foreign investors.

Second, if the home government informs home investors of a planned deviation, then the tax-treaty strategies indicate that the home investors would choose to reduce investment and inadvertently signal this deviation to the foreign country. At first glance, it may not appear obvious that this information revelation is in the best interests of the home investors. If home investors ignore the announcement and increase their stock of irreversible FDI, they will maintain the cooperative foreign tax for period d but will have greater trapped FDI (and greater losses) in the future. Therefore, home investors reduce FDI and tip off the foreign government only if:

$$\sum_{t=d}^{\infty} \delta^{t-d}(h(K-\gamma^{t-d+1}Z_{d-1})-h(K-\gamma^{t-d}Z_d)) \geq (1-\tau_d^{*c})h^s(Z_d), \qquad (10)$$

where Z_d and Z_{d-1} are functions of the expected cooperative tax rates. We show in Appendix B that because home's national income is the sum of home investors' profits and tax revenues from inbound FDI, equation (10) is also the necessary and sufficient condition for the home government to choose to warn its investors. Hence, if home finds it optimal to warn its investors, then their optimal action is to terminate cooperative behavior immediately. We show in Appendix B that (10) is more likely to be satisfied if the degree of irreversibility and/or the discount rate is larger. Since our main result shows that gradualism occurs if the degree of irreversibility is sufficiently large, we proceed under the assumption that (10) holds, implying that the home government will forewarn its investors of a planned deviation, and defer the alternative case to Appendix B. We show, there, that if (10) does not hold or if governments cannot forewarn their investors, then the magnitude of tax reductions is reduced, which strengthens our gradualism result.

B. Tax-Treaty Payoffs

Whether or not equation (10) holds, foreign investors are taken by surprise when a deviation occurs. When (10) holds home investors begin reducing FDI in the deviation period. This asymmetry of FDI during the deviation period and subsequent punishment stage is a distinguishing characteristic of this case. Using the investors' tax-treaty strategies, this implies that the deviation period FDI levels by the informed home investors and the surprised foreign investors are respectively given by:

$$Z_d = \gamma \cdot Z^c_{d-1} = \gamma \cdot Z(K, E(\tau^{*c}_{d-1})); \quad Z^*_d = Z^{*c}_d = Z^*(K^*, E(\tau^c_d)). \tag{11}$$

If both countries abide by the tax-treaty, then the home country's payoff in period t is:

$$I^c_t = I(K, K^*, \tau^c_t, \tau^{*c}_t) = h(K - Z^c_t) + (1 - \tau^{*c}_t) \cdot h^s(Z^c_t) + \tau^c_t \cdot f^s(Z^{*c}_t).$$

As a direct corollary of Proposition 1, we know that the optimal tax for a deviating home government is $\tau_d = 1$. Using this fact and equation (11), the home country's payoff from a period d deviation is given by $I^d_d = I(K, K^*, \gamma \cdot Z^c_{d-1}, \tau^c_d) = h(K - \gamma \cdot Z^c_{d-1}) + f^s(Z^{*c}_d)$. From equation (11), Proposition 1, and the tax-treaty strategies, the home country's payoff in each period after their period d deviation can be written as $I^{wd}_t = I(K, K^*, \gamma, Z^c_{d-1}, \tau^c_d) = h(K - \gamma^{t-d+1} \cdot Z^c_{d-1}) + f^s(\gamma^{t-d} \cdot Z^{*c}_d)$. The gain to deviating from the tax treaty in period d can then be succinctly expressed as:

$$\Psi_d = \Psi(K, K^*, \gamma, Z^c_{d-1}, \tau^c_d, \tau^{*c}_d) = I^d_d - I^c_d =$$

$$h(K - \gamma \cdot Z^c_{d-1}) - [h(K - Z^c_d) + (1 - \tau^{*c}_d) \cdot h^s(Z^c_d)] + (1 - \tau^c_d) \cdot f^s(Z^{*c}_d). \tag{12}$$

Similarly, the cost of deviating from the tax treaty in period d can be written as:

$$\Omega_d = \Omega(K, K^*, \gamma, Z^c_{d-1}, \tau^c_d, \tau^c_t, \tau^{*c}_t, \delta) = \sum_{t=d+1}^{\infty} \delta^{t-d}(I^c_t - I^{wd}_t) =$$

$$\sum_{t=d+1}^{\infty} \delta^{t-d}[h(K - Z^c_t) + (1 - \tau^{*c}_t) \cdot h^s(Z^c_t) + \tau^c_t \cdot f^s(Z^{*c}_t) - h(K - \gamma^{t-d+1} \cdot Z^c_{d-1}) - f^s(\gamma^{t-d} \cdot Z^{*c}_d)]. \tag{13}$$

Ψ^*_d and Ω^*_d are defined similarly. A self-enforcing cooperative tax path requires that the cost of deviation outweighs its benefits. This gives the following incentive compatibility constraints:

$$\Omega_d - \Psi_d \geq 0\,; \qquad \Omega_d^* - \Psi_d^* \geq 0 \qquad \text{for } d \in (1, 2, \ldots..) \qquad (14)$$

We are interested in the lowest tax rates that satisfy each of these constraints: $\{\underline{\tau}_t^c, \underline{\tau}_t^{*c}\}$ Tax treaties generally specify a single tax rate. We follow this convention and look for the lowest tax rate that satisfies both sets of incentive constraints: $\underline{\tau}_t = \max\{\underline{\tau}_t^c, \underline{\tau}_t^{*c}\}$. In the symmetric case K = K* and $Z_t^c = Z_t^{*c}$, therefore, $\Omega_d = \Omega_d^* = \Omega(K, \gamma, Z_{d-1}^c, \underline{\tau}_d, \underline{\tau}_t, \delta)$, $\Psi_d = \Psi_d^* = \Psi(K, \gamma, Z_{d-1}^c, \underline{\tau}_d)$ and both incentive constraints bind at the same tax rate: $\underline{\tau}_t^c = \underline{\tau}_t^{*c} = \underline{\tau}_t$.[20]

FIGURE 2 GOES APPROXIMATELY HERE

The home country's incentive constraint is illustrated for the first period of the treaty as a function of this common tax rate, $\underline{\tau}_t$, in Figure 2. The mathematics underlying this figure is provided in Appendix A and in the proof of Proposition 2. First, as shown in Appendix A, Ψ_t is decreasing in $\underline{\tau}_t$. A higher cooperative tax rate implies less inbound FDI from foreign. This reduces the tax base the home country can expropriate and reduces Ψ_t. Second, as a direct result of Proposition 1, in the initial period of the tax treaty, when previous FDI levels are zero, the non-cooperative outcome for all time is self-enforcing: $\Omega(K, K^*, \gamma, 0, 1_1, 1_t, 1_t, \delta) =$

$\sum_{t=2}^{\infty} \delta^{t-1}[h(K) + f^s(0) - h(K) - f^s(\gamma^{t-1} \cdot 0)] = 0 = h(K) - h(K) = \Psi(K, K^*, \gamma, 0, 1_1, 1_1)$. To see this result note that if the home country deviates in the first period, then its FDI is zero regardless of the cooperative tax.

C. The Main Results in the Symmetric Case

In this section we show that there exists a connected set of self-enforcing tax rates in the initial period that Pareto dominate the non-cooperative tax rates. We then ask whether Pareto optimal tax rates are immediately obtainable, i.e. whether the tax path $\tau_t^c = \tau_t^{*c} = 0$ for all t is self-enforcing. If there is no irreversibility (i.e. $\gamma = 0$) and home and foreign are symmetric, then our

setting is equivalent to a repeated game and the standard folk theorem results obtain, i.e. that this outcome occurs for a discount factor close to one.[21] As shown in the following proposition, irreversibility limits the set of initially self-enforcing outcomes so that if the extent of irreversibility is sufficiently large, then the standard result can not arise. Put another way, if the extent of irreversibility is large enough, then even the lowest of these initial tax rates are inefficiently high, indicating that further reductions are desirable. Finally, in Proposition 3, we show that the increase in FDI generated by this initial tax reduction creates slack in the incentive constraint making further tax-reductions self-enforcing.

Proposition 2: *Consider the symmetric case. (a.) If* $\delta > 0$ *and* $\gamma < 1$, *then there exists a self-enforcing path of tax-rates such that* $\underline{\tau}_t = \tau_t^c = \tau_t^{*c} < 1$ *for all t. (b.) If the extent of irreversibility,* γ, *is sufficiently large, then for all* $\delta < 1$, *the Pareto optimal tax path,* $\tau_t^c = \tau_t^{*c} = 0$ *for all* $t \geq 1$, *is not a stationary outcome. (c.) Furthermore, the cooperative tax* $\underline{\tau}_1$ *is increasing in* γ *and decreasing in* δ. *(d.) There exists a* $\delta^F(\gamma)$ *that is increasing in* γ, *such that for all* $\delta \geq \delta^F(\gamma)$, $\tau_t^c = \tau_t^{*c} = 0$, *for all t is self-enforcing, where* $\delta^F(0) < 1 < \delta^F(1)$.

Proof: First recall that there is no FDI before the tax treaty is signed so that $Z_0 = 0$. Second note that $K = K^*$, $Z_t^c = Z_t^{*c}$ and $\underline{\tau}_t^c = \underline{\tau}_t^{*c} = \underline{\tau}_t$ in the symmetric case. Therefore, $(1 - \tau_t^{*c})\ h^s(Z_t^c) + \tau_t^c f^s(Z_t^{*c}) = h^s(Z_t^c) = f^s(Z_t^{*c})$. Writing the incentive constraint for an arbitrary first period tax rate yields (recall that Z and Z^* are functions of tax rates):

$$\Omega(K, \gamma, 0, \underline{\tau}_1, \underline{\tau}_t, \delta) = \sum_{t=2}^{\infty} \delta^{t-1}[h(K - Z_t^c) + f^s(Z_t^{*c}) - h(K) - f^s(\gamma^{t-1} \cdot Z_1^{*c})] \geq$$

[20] In the asymmetric case both equations in (14) need not always hold with strict equality at the common tax rate. We present results suggesting which country's incentive constraint will bind in section IV.C.
[21] We show in Proposition 5 below that the symmetric case presents the easiest case for Pareto optimal taxes to be obtained and that they may not ever be obtained in the asymmetric case. Proposition 2, therefore, considers the easiest case for Pareto optimal taxes to obtain immediately.

$$\Psi(K,\gamma,0,\underline{\tau}_1) \;=\; h(K) - h(K - Z_1^c)\,. \tag{15}$$

(a.) To show that some tax reduction is possible we start by assuming that no tax reduction is ever possible and establish a contradiction. In this case, taxes are set a constant level $\tau_t^c = \tau_1^c = \tau_t^{*c} = \tau_1^{*c} = \tau_1$ for all t. Substituting this constant tax rate into equation (15), totally differentiating the resulting expression with respect to τ_1, using an envelope result for the optimality of Z (from equations 3 and 6), and simplifying, yields:

$$\frac{d(\Omega_1 - \Psi_1)}{d\tau_1} = \sum_{t=1}^{\infty} \delta^{t-1}[\tau_1 \cdot f_k^s(Z_1^{*c})\frac{\partial Z_1^{*c}}{\partial \tau_1} - \gamma^{t-1} f_k^s(\gamma^{t-1} Z_1^{*c})\frac{\partial Z_1^{*c}}{\partial \tau_1}] \tag{16}$$

Taking the limit of the above derivative as $\tau_1 \to 1$ from below and remembering from equation (3) that $Z_1^c(K,1) = Z_1^{*c}(K^*,1) = 0$ yields:

$$\lim_{\tau_1 \to 1} \frac{d(\Omega_1 - \Psi_1)}{d\tau_1} = \frac{\delta(1-\gamma)}{1-\delta\gamma} f_k^s(0)\frac{\partial Z_1^{*c}}{\partial \tau_1} < 0 \tag{17}$$

as long as δ is positive and γ is less than one. As established above $\Omega_1 - \Psi_1 = \Omega_1^* - \Psi_1^* = 0$, when $\tau_t = 1$ for all t. Therefore, from equation (17), $\Omega(K,\gamma,0,\underline{\tau}_1,\delta)$ must lie above $\Psi(K,\gamma,0,\underline{\tau}_1)$ for $\underline{\tau}_t$ sufficiently close to one and Ω_1 must be steeper than Ψ_1 for $\underline{\tau}_1$ in this range. A similar result holds for the foreign country's incentive constraint. Hence, there exists a self-enforcing tax that is strictly less than one as long as both nations place some weight on the future and investment is not completely irreversible.

(b.) Suppose that $\underline{\tau}_t = 0$ for all t is self-enforcing. This implies that the first period incentive constraint must satisfy:

$$\sum_{t=1}^{\infty} \delta^{t-1}\left([h(K - Z^{opt}) - h(K)] + [f^s(Z^{*opt}) - f^s(\gamma^{t-1} \cdot Z^{*opt})]\right) \geq 0\,. \tag{18}$$

Now, when γ approaches 1, the second bracketed term approaches zero. The first bracketed term is strictly negative for all t, therefore, equation (18) is not satisfied. This establishes the necessary contradiction. Since Ω_1 lies below Ψ_1 for $\underline{\tau}_1 = 0$ and lies above it for $\underline{\tau}_1$

near one there must be at least one other intersection of Ψ_1 and Ω_1 for $\underline{\tau}_1 \in (0, 1)$ and Ψ_1 must cut Ω_1 from above at the lowest of these intersections.[22] Parts (a) and (b) are represented in Figure 2, where $\underline{\tau}_1^c$ represents the lowest tax rate for which home is willing to maintain the agreement.

(c.) Differentiating both sides of equation (15) with respect to γ and δ we have:

$$\frac{\partial \Psi_1}{\partial \delta} = \frac{\partial \Psi_1}{\partial \gamma} = 0. \quad \frac{\partial \Omega_1}{\partial \gamma} = \sum_{t=2}^{\infty} \delta^{t-1}[-f_k^s(\gamma^{t-1} \cdot Z_1^{*c}) \cdot (t-1)\gamma^{t-2} \cdot Z_1^{*c}] < 0.$$

$$\frac{\partial \Omega_1}{\partial \delta} = \sum_{t=2}^{\infty} (t-1)\delta^{t-2}[h(K - Z_t^c) + f^s(Z_t^{*c}) - h(K) - f^s(\gamma^{t-1} \cdot Z_t^{*c})] \geq 0$$

In interpreting the sign of $\partial\Omega_1/\partial\delta$, first note that Ω_1 must be positive in any self-enforcing agreement and the bracketed term must therefore be positive in some time period. Second, the bracketed expression is not decreasing over time (and is strictly increasing in future periods if either $\gamma < 1$ or if future tax reductions are expected). An increase in the discount factor shifts weight to these future periods. Figure 3 shows how changes in γ and δ shift the Ω_1 curve and, therefore, how the lowest self-enforcing tax rate, $\underline{\tau}_1^c$ is affected by changes in these parameters.

(d.) From equation (18) we see that when $\gamma = 0$ and δ approaches 1 that $\underline{\tau}_1^c = 0$ satisfies the incentive constraint. This is a direct result of the assumed desirability of FDI occurring. Denote the lowest δ that satisfies equation (18) when $\gamma = 0$ as $\delta^F(0)$. We have just seen that $\delta^F(0) < 1$ and part (b) showed that $\delta^F(1) > 1$. Finally, we show that $\delta^F(\gamma)$ is increasing in γ. From part (c) we know that Ω_1 is strictly increasing (decreasing) in δ (γ). Furthermore we have that

$$\frac{\partial^2 \Omega}{\partial\gamma\partial\delta} = \sum_{t=2}^{\infty} (t-1)\delta^{t-2}[-f_k^s(\gamma^{t-1} \cdot Z_1^{*c}) \cdot (t-1)\gamma^{t-2} \cdot Z_1^{*c}] < 0.$$

[22] Although Ω_d is drawn concave we do not require this result or make this claim. All we require for our results is that Ψ_d intersects Ω_d from above at their lowest intersection and from below at a first period tax rate of one. We point out, however, that a sufficient condition for concavity of Ω_d is symmetric countries and homogeneity of $h^S(Z)$. Finally, since Ψ_d and Ω_d are both continuous transformations of continuous functions they are also continuous and the existence of this intersection is assured.

Hence, an increase in γ requires an increase in δ in order for the incentive constraint to be satisfied in the initial period.

FIGURES 3 AND 4 GO APPROXIMATELY HERE

The idea behind Proposition 2 is that when FDI is irreversible (or almost so), the costs to deviation are small since reductions in inbound FDI are slow to happen. Alternatively, with immediate reversibility, a standard folk-theorem-type result is produced. Increasing the extent of irreversibility from its minimal level has a monotonic effect on the critical discount factor necessary for Pareto optimal taxes to occur in the initial period. We assume in what follows that governments are sufficiently short-lived so that their discount factor does not exceed this critical level. Even here, however, some cooperative tax rate less than one is self-enforcing. This tax reduction, in turn, generates the irreversible FDI that permits further tax reductions.

Proposition 3: *After the initial (and each subsequent) tax reduction, additional reductions that were previously not incentive-compatible become self-enforcing.*

Proof: We now show that after this initial tax reduction to $\underline{\tau}_1$ further tax cuts become possible. From equation (4) we know that Z_1 increases as $\underline{\tau}_1$ falls. We see the effect of Z_1 on the incentive constraint in period 2 by differentiating equations (12-14):

$$\frac{\partial \Omega_d}{\partial Z_{d-1}} = \sum_{t=d+1}^{\infty} \delta^{t-d} \cdot \gamma^{t-d+1} \cdot h_k(K - \gamma^{t-d+1} Z^c_{d-1}) > 0,$$

$$\frac{\partial \Psi_d}{\partial Z_{d-1}} = -\gamma \cdot h_k(K - \gamma \cdot Z^c_{d-1}) < 0, \qquad \frac{\partial(\Omega_d - \Psi_d)}{\partial Z_{d-1}} > 0. \qquad (19)$$

Equation (19) shows that the increase in Z_{d-1} creates slack in the period d incentive constraint. As shown in Figure (4) this slack can be taken up by a lower period d tax rate. Since the first-period tax reduction generates positive FDI for both countries, slack is introduced into the foreign incentive constraint as well. A reduction is, therefore, produced in the lowest mutually self-enforcing tax rate: $\underline{\tau}_{t+1} < \underline{\tau}_t$.

Note that this cooperative tax path is not unique. There exists a continuum of less efficient tax rates between $\underline{\tau}_t$ and 1 that are also self-enforcing, however, we feel it is natural to expect that the treaty would pick the lowest mutually self-enforcing tax rate in each period.

IV. Asymmetries

A. Gradualism in the Asymmetric Case

We start by showing that our main result also holds in the more general asymmetric case in which capital stocks, irreversibility, and/or discount rates may differ between countries.

Proposition 4: *If* $\delta, \delta^* > 0$ *and* $\gamma, \gamma^* < 1$, *then there exists a self-enforcing path of tax-rates such that* $\{\tau_t^c, \tau_t^{*c}\} < \{1, 1\}$ *for all t. Furthermore, after the initial (and each subsequent) tax reduction, additional reductions that were previously not incentive compatible become self-enforcing.*

Proof: The proof of Proposition 4 is almost identical to that of Propositions 2a, 2b and 3. We only discuss the few differences here. The incentive constraint for the first period of the tax treaty, in the non-symmetric case, can be written as follows:

$$\Omega_1 - \Psi_1 = \sum_{t=2}^{\infty} \delta^{t-1}[\{h(K - Z_t^c) + (1-\tau_t^{*c})h^s(Z_t^c) + \tau_t^c f^s(Z_t^{*c})\} - \{h(K) + f^s(\gamma^{*t} Z_1^{*c})\}]$$

$$- [\{h(K) + f^s(Z_1^{*c})\} - \{h(K - Z_1^c) + (1-\tau_1^{*c})h^s(Z_1^c) + \tau_1^c f^s(Z_1^{*c})\}] \geq 0. \qquad (20)$$

We are still interested in examining the behavior of this constraint as tax rates approach 1. The derivative of this expression with respect to the common tax rate is:

$$\frac{d(\Omega_1 - \Psi_1)}{d\tau_1} = \frac{1}{1-\delta}[f^s(Z_1^{*c}) - h^s(Z_1^c) + \tau_1 \cdot f_k^s(Z_1^{*c})\frac{\partial Z_1^{*c}}{\partial \tau_1}] - \sum_{t=1}^{\infty}(\delta\gamma)^{t-1} f_k^s(\gamma^{*t-1} Z_1^{*c})\frac{\partial Z_1^{*c}}{\partial \tau_1} \qquad (21)$$

It is straightforward to verify that under symmetry equation (21) becomes equal to equation (16). In any case, because $Z_1^c(K,1) = Z_1^{*c}(K^*,1) = 0$, when taking the limit as $\tau \to 1$ from below equation (21) still yields equation (17).

To see that Pareto-optimal taxes may not be initially self-enforcing we rewrite the home country's incentive constraint under zero taxes for the asymmetric case:

$$\sum_{t=1}^{\infty}\delta^{t-1}\left([h(K-Z^{opt})-h(K)]+[h^{s}(Z^{opt})-f^{s}(\gamma^{*t-1}\cdot Z^{*opt})]\right)\geq 0\,. \qquad (18')$$

In the asymmetric case, this incentive constraint may be satisfied, even as γ^{*} approaches one, if the home country has a lot more outbound FDI than does the foreign country making the positive second term larger the negative first term. In this case, however, the second bracketed term in the foreign counterpart of (18′) would be negative so that foreign's incentive constraint would not be satisfied. Hence, even in the asymmetric case the requirement an agreement be incentive-compatible for both countries limits the possibility of Pareto-optimal tax rates from obtaining. Furthermore, although the gain to deviating from the treaty may not be monotonically declining for a country that has much more outbound FDI (see Appendix A) the requirement that both constraints are satisfied indicates that we only have to consider the country with less outbound FDI, for whom this function Ψ_d (or Ψ_d^{*}) is monotonically declining.

The rest of the proof proceeds identically to that of Proposition 3, with the following modification. The tax rate specified in the agreement is $\underline{\tau}_1 = \max\{\underline{\tau}_1^{c}, \underline{\tau}_1^{*c}\}$ since this is the lowest tax rate that satisfies both incentive constraints. Hence, the necessary reduction is in $\underline{\tau}_{t+1} = \max\{\underline{\tau}_{t+1}^{c}, \underline{\tau}_{t+1}^{*c}\} < \underline{\tau}_t = \max\{\underline{\tau}_t^{c}, \underline{\tau}_t^{*c}\}$.

Because of the conventions of tax treaties in practice, we restrict attention to agreements in which taxes are identical between countries. In the asymmetric case, this means that one nation's incentive compatibility constraint may remain slack. Thus, there may exist Pareto preferred tax reductions whereby both equations in (14) hold with equality. Given this source of potential inefficiency we next ask what outcomes are obtainable when treaties are restricted to a common tax rate.

B. Attainable Outcomes

We now examine the limit of this gradual tax-reduction path and ask if the Pareto efficient tax rate pair {0,0} is ever self-enforcing. Consider home country welfare along the cooperative path as given by $I(K, K^*, \tau_t^c, \tau_t^{*c})$. Differentiating this expression with respect to the common tax rate $\underline{\tau}_t$ and using equation (3) we have:

$$\frac{\partial I_t^c}{\partial \underline{\tau}_t} = -h^s(Z_t^c) + f^s(Z_t^{*c}) + \underline{\tau}_t \cdot f^s(Z_t^{*c}) \frac{\partial Z_t^{*c}}{\partial \underline{\tau}_t}. \qquad (22)$$

In the symmetric case, $h^s(Z_t^c) = f^s(Z_t^{*c})$ and the above derivative is negative for any positive tax rate. Therefore, moving the common tax rate towards zero is always Pareto improving. Since in the symmetric case the same is true for foreign, both countries would agree to all self-enforcing tax reductions until taxes reach the efficient level. On the other hand, consider the case in which $h^s(\cdot)$ is zero. Now (22) is non-monotonic in $\underline{\tau}_t$ and is positive for small values of $\underline{\tau}_t$, implying that the home country will not agree to completely eliminate withholding taxes. Hence, when $h^s(\cdot)$ is small relative to $f^s(\cdot)$, there exists a strictly positive tax rate τ^{min} such that, without side-payments, home will not agree to a tax rate below this level. This demonstrates that zero taxes may not be obtainable between highly asymmetric countries.[23] This result is summarized as Proposition 5.[24]

Proposition 5: *Although a Pareto improvement over the non-treaty tax rates is always possible, Pareto efficient tax rates are never self-enforcing for sufficiently asymmetric countries.*

Instead of side payments, we can consider asymmetric taxes. Although equal tax reductions below τ^{min} are not welfare improving for a small home country, a small tax reduction by the home country coupled with a large foreign tax reduction may improve home welfare.

[23] Mayer (1981) and Kennan and Reizman (1988) show that, in the absence of side payments, globally efficient free trade may not be obtainable between countries of differing sizes.

[24] In Appendix B we provide an alternative proof of Proposition 5 and also show how the constraint implied by equation (10) may limit the possibility of Pareto optimal taxes obtaining for symmetric countries if the discount factor is sufficiently small.

Because $f^s(\cdot)$ is larger than $h^s(\cdot)$ in the asymmetric case we have been considering, this tax cut can improve foreign country welfare as well. Despite this, unless the foreign country chooses negative (and distortionary) taxes the home country will never agree to zero taxes, implying that fully efficient FDI flows are still unattainable. Although of theoretical interest, asymmetric tax rates are rarely encountered in real-world tax treaties.[25] As our main focus is on characterizing existing tax agreements we leave the specification of optimal tax reduction paths for asymmetric countries for future research.

C. Comparative Statics and Asymmetries.

Although our results hold for both symmetric and asymmetric countries, it is instructive to consider how differences in γ and K between the two countries influences the path of self-enforcing tax rates. The first of these results is contained in the following proposition.

Proposition 6: *A once and for all increase in the extent of irreversibility of one country's FDI in period t, will: (a.) lower taxes in that period and in all future periods if that country's incentive constraint is binding; (b.) raise taxes in that period and in all future periods if the other country's incentive constraint is binding.*

Proof: Assume, without loss of generality, that home's incentive constraint is binding. Differentiating the incentive constraint with respect to γ and γ^* yields:

$$\frac{\partial \Omega_d}{\partial \gamma} = \sum_{t=d+1}^{\infty} \delta^{t-d} \cdot h_k(K - \gamma^{t-d+1} Z_{d-1}^c) \cdot (t-d+1) \cdot \gamma^{t-d} \cdot Z_{d-1}^c > 0;$$

$$\frac{\partial \Omega_d}{\partial \gamma^*} = \sum_{t=d+1}^{\infty} \delta^{t-d} \cdot [-f_k^s(\gamma^{t-d} Z_d^{*c})] \cdot (t-d) \cdot \gamma^{t-d-1} \cdot Z_d^{*c} < 0;$$

$$\frac{\partial \Psi_d}{\partial \gamma} = -h_k(K - \gamma \cdot Z_{d-1}^c) \cdot Z_{d-1}^c < 0 = \frac{\partial \Psi_d}{\partial \gamma^*}.$$

[25] One example of this is the U.S. treaty with Pakistan. While Pakistani firms enjoy reduced withholding tax rates in the U.S., U.S. firms do not receive preferential withholding taxes in Pakistan. They do, however, benefit from increased coordination in tax laws.

For an intuitive explanation of the above proposition, consider an increase in home's irreversibility parameter. This reduced salvageability of home FDI reduces home's gain from deviation. Similarly, since more FDI will be trapped in foreign following a deviation, home's future cost of deviation increases. Combining these indicates that an increase in γ creates slack in home's incentive constraint and reduces its lowest self-enforcing tax rate both when the change occurs and in all future periods. On the other hand, the increase in home's γ reduces foreign's future cost of deviation because home cannot recall its capital as swiftly. Since a foreign deviation takes home by surprise, foreign's gain from deviation is invariant to changes in γ. Together, these results show that foreign's lowest self-enforcing tax rate is increasing in the home γ. Although these changes in $\underline{\tau}_t^c$ and $\underline{\tau}_t^{*c}$ move in opposite directions, it is important to remember that only one country's incentive constraint need be binding in the asymmetric case. If home's constraint is binding, then the common tax, $\underline{\tau}_t$, drops, otherwise it increases. One implication of Proposition 6 is that if countries are identical except in the ease with which their FDI is reversible, then the country with the relatively reversible FDI will drive the treaty. Put succinctly, if $\gamma > \gamma^*$ is the only difference between the countries, then $\underline{\tau}_t = \underline{\tau}_t^{*c} = \max\{\underline{\tau}_t^c, \underline{\tau}_t^{*c}\}$, for all t.

An alternative type of asymmetry arises when countries differ in their capital stocks. Starting with identical countries, suppose that there is a one-time permanent increase in the home's capital stock, so that $K > K^*$. This increases the attractiveness of FDI for home (recall that $\partial Z_t/\partial K > 0$ from equation (5)). The effect on home's incentive constraint is given by (where we again use the envelope result for the optimality of home's FDI):

$$\frac{\partial(\Omega_d - \Psi_d)}{\partial K} = \sum_{t=d}^{\infty} \delta^{t-d}[h_k(K - Z_t^c) - h_k(K - \gamma^{t-d} Z_{d-1}^c)] > 0, \qquad (23)$$

implying that home's incentive constraint loosens. Put simply, an increase in home's capital stock makes FDI is more profitable and home will prefer to encourage it even more than it

currently does. This is not always the case for the now relatively small foreign country. The total effect on foreign's incentive constraint, again holding taxes constant, is

$$\frac{\partial(\Omega_d^* - \Psi_d^*)}{\partial K} = \sum_{t=d}^{\infty} \delta^{t-d}[\tau_t^{*c} \cdot h_k^s(Z_t^c)\frac{\partial Z_t^c}{\partial K} - \gamma^{t-d} \cdot h_k^s(\gamma^{t-d} Z_d^c)\frac{\partial Z_d^c}{\partial K}] \qquad (24)$$

which is ambiguous in sign. When K rises, this increases foreign's inbound FDI. A greater current level of FDI for foreign to expropriate increases foreign's gain from deviation, that is $\frac{\partial \Psi_d^*}{\partial K} = (1-\tau_d^{*c})h_k^s(Z_d^c)\frac{\partial Z_d^c}{\partial K} > 0$. The effect of this rise in inbound FDI on the cost of a foreign deviation is less clear. Since more present FDI implies that more FDI will remain in foreign after a deviation, this reduces foreign's cost of breaking the treaty. At the same time, however, inbound FDI along the cooperative path also increases. Unless the treaty specifies zero tax rates both in the current period and ever after, this rise in inbound FDI increases foreign's future tax revenues and increases its cost of deviation. These effects are at odds and, therefore, yield an ambiguous effect. Note that if all future taxes are zero, then an increase in K causes foreign's incentive constraint to bind at a higher $\underline{\tau}_t^{*c}$, implying that it would choose to deviate. This mirrors the results of Proposition 5 and it suggests that the smaller, less advanced treaty partner may face greater difficulty in maintaining the agreement.

V. Conclusions

Before concluding, it is instructive to cast our model in a different light by reinterpreting tax rates as probabilities of expropriation. Although none of our results change, this does suggest two intriguing lines of thought. First, as the BEA's (1998) data shows, the largest hosts for U.S. investment also send the most FDI to the U.S. and numerous empirical studies have found that FDI activity is greatest between countries that are similar in size and in relative factor endowments.[26] These results contradict Helpman's (1984) vertical FDI model in which

[26] Studies using U.S. data include Brainard (1997) and Markusen and Maskus (2001). Blonigen, Davies, and Head (2002) find similar patterns in OECD data.

differences in relative factor endowments drive FDI. Instead, the results are generally viewed as supportive of Markusen's (1984) horizontal model in which FDI occurs between similar countries to realize firm-level economies of scale and avoid trade costs. Our results provide an alternative explanation for the data, since it is only through bilateral investment that countries can punish one another after "inhospitable" behavior. Similar countries have similar FDI outflows and, therefore, can credibly commit to larger reductions in taxes or in the risk of expropriation. Furthermore, similar countries are able to eventually attain a greater degree of economic integration so that FDI can come closer to the efficient zero-tax (or zero-risk) allocation of FDI. As a result, our model predicts that FDI will be greatest between similar countries, exactly the pattern found in the data.

Second, a policy implication of this expropriation interpretation is particularly interesting for developing countries. These nations have long sought out FDI as a tool for growth and expend a great deal of effort to attract foreign investors. Our model suggests that in order to encourage inbound investment, it may also be useful to promote outbound investment or, more generally, the creation of other hostages since this decreases the perceived probability of future expropriation. This mirrors the export promotion advice discussed in the development literature (Pack, 1989), although it occurs from changes in the time inconsistency problem rather than comparative advantage.

In summary, we have considered one facet of the commonly-used bilateral tax treaties on FDI. We find that bilateral tax treaties can indeed increase FDI and improve the global allocation of capital relative to non-treaty outcomes. Furthermore, when governments are unable to commit to their tax policies, we find that in some cases Pareto efficient tax rates can only be obtained when taxes are gradually reduced. Even in those situations for which globally optimal tax rates are not self-enforcing, we find that improvements over the non-treaty equilibrium are still

obtainable.[27] This suggests that, much like international trade agreements, bilateral tax treaties are a powerful tool for increasing world welfare.

[27] On the other hand, if government preferences place a much higher value on the profits of the multinational than on tax revenues, then their initial incentive to deviate is reduced. We thank an anonymous referee for bringing this point to our attention. Continuing on this line of thought can also provide an interesting political economy explanation for the observed gradualism. In particular, if these same preferences gradually shift from a placing a greater weight on tax revenues to placing a greater value on firm profits, then the incentive constraint would also slacken over time allowing a decrease in treaty tax rates and an increase in FDI.

References

Aizenman, Joshua (1996), "Foreign direct investment as a commitment mechanism in the presence of managed trade," *International Economic Journal*, 10:4, pp. 1-28.

Ausubel, L., and R. Deneckere, (1987) "One is Almost Enough for Monopoly," *Rand Journal of Economics* 18, Summer, 1987, pp. 255-274.

Benhabib, J., and R. Radner, (1992) "The Joint Exploitation of a Productive Asset: a Game-theoretic Approach," *Economic Theory* 2, 1992, pp.155-190.

Blonigen, Bruce and Ronald B. Davies (2000), "The effect of bilateral tax treaties on U.S. FDI activity," NBER Working Paper Series, No. 7929, Cambridge, MA.

Blonigen, Bruce; Davies, Ronald B. and Head, Keith "Estimating the Knowledge-Capital Model of the Multinational Enterprise: A Comment." *NBER Working Paper*, No. 8929.

Bond, Eric and Jee-Hyeong Park (forthcoming), "Gradualism in trade agreements with asymmetric countries," *Review of Economic Studies.*

Bond, Eric and Larry Samuelson (1986), "Tax holidays as signals," *American Economic Review*, 76 (4), pp. 820-26.

Bond, Eric and Larry Samuelson (1989a), "Bargaining with commitment, choice of techniques, and direct foreign investment," *Journal of International Economics*, 26, pp. 77-97.

Bond, Eric and Larry Samuelson (1989b), "Strategic behaviour and the rules for international taxation of capital," *Economic Journal*, 99:1099-1111.

Brainard, S. Lael (1997), "An empirical assessment of the proximity-concentration trade-off between multinational sales and trade," *American Economic Review*, 87:4, pp. 520-544.

Bureau of Economic Analysis (1998), Detailed annual balance of payments and position estimates, Bureau of Economic Analysis International Data, Washington, D.C.

Cave, J., (1987) "Long-term Competition in a Dynamic Game: The Cold Fish-War," *Rand Journal of Economics* 18, Winter, 1987, pp. 596-610.

Caves, Richard E. (1996), *Multinational Enterprise and Economic Analysis*, Second edition, Cambridge Press.

Chisik, Richard (forthcoming), "Gradualism in free trade agreements: A theoretical justification," *Journal of International Economics*.

Davies, Ronald B. (forthcoming), "The OECD model tax treaty: tax competition and two-way capital flows," *International Economic Review*.

Devereux, Michael B. (1997), "Growth, specialization, and trade liberalization," *International Economic Review*, 38:3, pp. 565-585.

Diamond, Walter and Dorothy Diamond (1998), *International Tax Treaties of All Nations*, Ocean Publications, Dobbs Ferry, New York.

Doyle, Chris and Sweder Van Wijnbergen (1994), "Taxation of foreign multinationals: a sequential bargaining approach to tax holidays," *International Tax and Public Finance*, 1:3, pp. 211-25.

Eaton, J. and M. Gersovitz (1984), "A theory for expropriations and deviation from perfect capital mobility," *Economic Journal*, 94, pp. 16-40.

Furusawa, T., and E. Lai (1999), "Adjustment Costs and Gradual Trade Liberalization," *Journal of International Economics*, December, 49:2, pp. 333-361.

Greif, Avner, Paul Milgrom and Barry Weingast (1994), "Coordination, Commitment and Enforcement: The Case of the Merchant Guild," *Journal of Political Economy* 102:4, pp. 745-776.

Hartman, David (1985), "Tax policy and foreign direct investment," *Journal of Public Economics*, V. 26, pp. 107-121.

Helpman, Elhanan (1984), "A Simple Theory of International Trade with Multinational Corporations," *Journal of Political Economy*, *94*(3), pp. 451-71.

Hines, James R. Jr. (1988), "Taxation and U.S. multinational investment," in *Tax Policy and the Economy*, Lawrence Summers ed., MIT Press.

Hufbauer, Gary (1992), *U. S. Taxation of International Income: Blueprint for Reform*, Institute for International Economics.

Janeba, Eckhard (1995), "Corporate income tax competition, double taxation treaties, and foreign direct investment," *Journal of Public Economics*, 56:311-325.

Kennan, J. and R. Riezman, (1988) "Do Big Countries Win Tariff Wars?" *International Economic Review*, February 1988, 29:1, pp. 81-85.

Lapan, Harvey (1988), "The Optimal Tariff, Production Lags, and Time Consistency," *American Economic Review*, June 1988, 78:3, pp. 395-401.

Markusen, James R. (1984), "Multinationals, Multi-plant Economies, and the Gains from Trade," *Journal of International Economics*, *16*(3/4), pp. 205-26.

Markusen, James R. and Keith E. Maskus (2001), "Multinational Firms: Reconciling Theory and Evidence," in Magnus Blomstrom and Linda S. Goldberg, eds., *Topics in Empirical International Economics: A Festschrift in Honor of Robert E. Lipsey*. Chicago, IL: University of Chicago Press for National Bureau of Economic Research.

Mayer, Wolfgang (1981) "Theoretical Considerations on Negotiated Tariff Adjustments," *Oxford Economics Papers*, March 1981, 33:1, pp.135-153.

McLaren, John (1997), "Size, Sunk Costs, and Judge Bowker's Objection to Free Trade," *American Economic Review*, June 1997, 87:3, pp. 400-420.

Markusen, James and Anthony Venables (1998), “Multinational firms and the new trade theory,” *Journal of International Economics*, V. 46, No. 2, pp. 183-203.

Organization for Economic Cooperation and Development (1997) *Model Tax Convention on Income and on Capital*, Paris.

Pack, Howard (1989), "Industrialization and Trade," *Handbook of Development Economics*, Vol. III, eds. Hollis B. Chenery and T. N. Srinivasan, North-Holland Publishers, pp. 333-380.

Price-Waterhouse (1995), *Corporate Taxes: A Worldwide Summary*, I. Richard Wood ed., New York.

Radaelli, Claudio M. (1997), *The Politics of Corporate Taxation in the European Union*, Routledge Research in European Public Policy, New York.

Staiger, Robert (1995), “A theory of gradual trade liberalization,” in *New Directions in Trade* Theory, eds. J. Levinsohn, A. Deardorff , and R. Stern, University of Michigan Press.

Thomas, Jonathan and Tim Worrall (1994), “Foreign direct investment and the risk of expropriation,” *The Review of Economic Studies*, 61:81-108.

APPENDIX A: *The shape of the incentive constraint.*

First, in the asymmetric case,

$$\frac{\partial \Psi_d(\cdot)}{\partial \underline{\tau}_d} = (1-\underline{\tau}_d)\cdot f_k^s(\cdot)\cdot\frac{\partial Z_d^{*c}(\cdot)}{\partial \underline{\tau}_d} - f^s(Z_d^{*c}) + h^s(Z_d^c);$$

$$\frac{\partial \Omega_d(\cdot)}{\partial \underline{\tau}_d} = \sum_{t=d+1}^{\infty}\delta^{t-d}[\underline{\tau}_t\cdot f_k^s(Z_t^{*c})\cdot\frac{\partial Z_t^{*c}(\cdot)}{\partial \underline{\tau}_d} - \gamma^{t-d}\cdot f_k^s(\gamma^{t-d}Z_d^{*c})\cdot\frac{\partial Z_d^{*c}(\cdot)}{\partial \underline{\tau}_d} + f^s(Z_t^{*c}) - h^s(Z_t^c)].$$

In the symmetric case $f^s(Z_t^{*c}) = h^s(Z_t^c)$ and $\frac{\partial \Psi_d(\cdot)}{\partial \underline{\tau}_d} < 0$. In the asymmetric case this result must also hold as long as $h^s(Z_t^c)$ is not too large as compared with $f^s(Z_t^{*c})$. Similarly, if the countries are not too asymmetric, then the sign of $\frac{\partial \Omega_d(\cdot)}{\partial \underline{\tau}_d}$ is dependant on the first two terms in brackets.

In this case, for $\underline{\tau}_t$ near zero for all t, $\frac{\partial \Omega_d(\cdot)}{\partial \underline{\tau}_d}$ is increasing. In the main text we show that for the opposite case of no tax reductions ever occurring (i.e. $\underline{\tau}_t = 1$ for all t) that Ω_d is decreasing for $\underline{\tau}_t$ near one and that it is steeper than Ψ_d in this range.

APPENDIX B: *Informing Outbound Investors of a Planned Deviation.*

Suppose that the government does not inform the outbound investors. In this case, the home investor's FDI allocations in the deviation period can be written as $Z_d^c = Z(K, E(\tau_d^{*c}))$ whereas the foreign allocation remains as described in equation (11). Because the foreign government is not tipped off as to the coming deviation they adhere to the cooperative tax in period d. Using the superscript (^{alt}) to denote payoffs in this case we have $I_d^{d,alt} = h(K - Z_d^c) + (1 - \tau_d^{*c}) \cdot h^s(Z_d^c) + f^s(Z_d^{*c})$. Naturally the cooperative payoff doesn't change, therefore, the gain from deviating can be expressed as $\Psi_d^{alt} = I_d^{d,alt} - I_d^c = (1 - \tau_d^c) \cdot f^s(Z_d^{*c})$.

In the subsequent punishment stage, the only difference with the case in the main text is in home's FDI allocation: $I_t^{wd,alt} = h(K - \gamma^{t-d} \cdot Z_d^c) + f^s(\gamma^{t-d} \cdot Z_d^{*c})$. The cost of deviating is

$$\Omega_d^{alt} = \sum_{t=d+1}^{\infty} \delta^{t-d}[h(K - Z_t^c) + (1 - \tau_t^{*c}) \cdot h^s(Z_t^c) + \tau_t^c \cdot f^s(Z_t^{*c}) - h(K - \gamma^{t-d} \cdot Z_d^c) - f^s(\gamma^{t-d} \cdot Z_d^{*c})].$$

The tax-treaty tax rate, therefore, must satisfy $\Omega_d^{alt} \geq \Psi_d^{alt}$ as well as $\Omega_d \geq \Psi_d$. As long as $\Omega_d \geq \Psi_d$ the additional constraint is satisfied if $(\Omega_d^{alt} - \Psi_d^{alt}) - (\Omega_d - \Psi_d) \geq 0$. Note that if this condition is satisfied then a deviation is less attractive (the costs minus the benefits are greater) if the government does not inform its investors. Hence, a sufficient condition for the home government to inform its outbound investors is:

$$(\Omega_d^{alt} - \Psi_d^{alt}) - (\Omega_d - \Psi_d) = \sum_{t=d}^{\infty} \delta^{t-d}[h(K - \gamma^{t-d+1} Z_{d-1}^c) - h(K - \gamma^{t-d} Z_d^c)] - (1 - \tau_d^{*c}) \cdot h^s(Z_d^c) \geq 0,$$

which is precisely equation (10).

We now analyze the behavior of equation (10). Because $Z_{d-1}^c \leq Z_d^c$ and $\gamma \leq 1$, the term in the [] brackets is positive, therefore, the above expression is increasing in δ. Differentiating this expression with respect to γ we have:

$$\frac{\partial}{\partial \gamma}(\cdot) = \sum_{t=d}^{\infty} \delta^{t-d}\{[h_k(K - \gamma^{t-d+1} \cdot Z_{d-1}^c) \cdot (t - d + 1) \cdot \gamma^{t-d} \cdot Z_{d-1}^c] - [h_k(K - \gamma^{t-d} Z_d^c) \cdot (t - d) \cdot \gamma^{t-d-1} \cdot Z_d^c]\}$$

Remembering that h(·) is strictly concave, we see that the above derivative must be positive as long as δ is greater than zero. Differentiating (10) with respect to $\underline{\tau}_t$, and again using the envelope result from equation (3) yields:

$$\frac{\partial}{\partial \underline{\tau}_t}(\cdot) = \sum_{t=d+1}^{\infty} \delta^{t-d}[h_k(K - \gamma^{t-d} \cdot Z_d^c) \cdot \gamma^{t-d} \cdot \frac{\partial Z_d^c(\cdot)}{\partial \tau_d^{*c}}] + h^s(Z_d^c),$$

which is negative for $\underline{\tau}_t$ near 1. Hence, small tax reductions create slack in equation (10). As taxes approach zero the above derivative may become positive, so that equation (10) may become binding as taxes drop further and may limit the magnitude of tax reductions. To further see this effect of equation (10) in later periods of the tax treaty, note that it is decreasing in Z_{d-1}^c so that increases in Z_{d-1}^c reduces the slack in equation (10).

As an alternative method of seeing how the additional constraint may limit tax reductions, consider the incentive constraints in the first period in which zero taxes are obtained: period T. In the case presented in the main text this can be written as:

$$\Omega_T - \Psi_T = \sum_{t=T}^{\infty} \delta^{t-T}[h(K - Z^{opt}) + h^s(Z^{opt}) - h(K - \gamma^{t-T+1} \cdot Z_{T-1}^c) - f^s(\gamma^{t-T} \cdot Z^{*opt})].$$

This expression is clearly positive if $\gamma < 1$ and $h^s(Z^{opt}) \geq f^s(Z^{*opt})$.

The additional incentive constraint in period T is:

$$\Omega_T^{alt} - \Psi_T^{alt} = \sum_{t=T+1}^{\infty} \delta^{t-T}[h(K - Z^{opt}) + h^s(Z^{opt}) - h(K - \gamma^{t-T} \cdot Z^{opt}) - f^s(\gamma^{t-T} \cdot Z^{*opt})] - f^s(Z^{*opt}).$$

This expression is positive if $\gamma < 1$ and $h^s(Z^{opt}) \geq f^s(Z^{*opt})$ and if δ is sufficiently large. Hence, we see not only the result of Proposition 5, but also the claim in footnotes 21 and 24.

TABLE 1: *Tax Rate Reductions.*

Reductions in bold.

Treaty	Year	Dividend Paid to Parent [1]	Dividend Paid to Other	Interest Payments [2]	Royalties [3]
U.S. Non-treaty Tax		30	30	30	30
Argentina-Sweden	1962 [4]	no reductions	no reductions	no reductions	**15**
	1995	**10**	**15**	**12.5**	**10**
Australia-U.S.	1953 [4]	**15**	**15**	30	30
	1984	15	15	**10**	**10**
Barbados-U.S.	1987 [4]	**5**	**15**	**12.5**	**12.5**
	1993	5	15	**5**	**5**
Canada-U.S.	1941 [4]	**15**	**15**	**15**	**15**
	1985	**10**	15	15	**10**
	1996 [5]	**5**	15	**10**	**0**
France-U.S.	1940 [4]	**5**	**15**	**10**	**5**
	1987	5	15	**0**	5
Germany-U.S.	1954 [4]	**15**	**15**	**0**	**0**
	1991	**5**	15	0	0
India-U.S.	1990 [4]	**15**	**25**	**15**	**20**
	1995 [6]	15	25	15	**15**
Indonesia-U.S.	1990 [4]	**15**	**15**	**15**	**15**
	1995 [6]	15	15	15	**10**
	1997	**10**	15	15	10
Israel-Netherlands	1973 [4]	**15**	**15**	**10**	**5**
	1996	**5**	15	10	5
Japan-Norway	1959 [4]	**15**	**15**	**15**	**15**
	1967	**10**	15	**10**	**10**
	1992	**5**	15	10	10

Source: Diamond and Diamond (1998).

[1] The most favorable withholding tax available. The percentage of equity holdings necessary to receive this tax rate varies from treaty to treaty.

[2] Withholding tax on interest payments to non-financial institutions.

[3] Withholding tax on royalty payments for industrial royalties.

[4] Year treaty began.

[5] Not part of an actual treaty revision, tax rates reported by Price-Waterhouse (1996).

[6] Reduction specified by the initial version of the treaty.

Figure 1

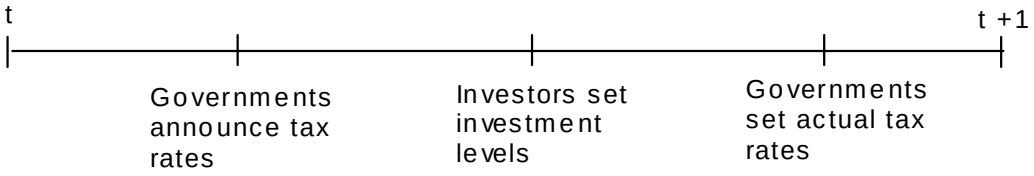


Figure 2: Determination of $\underline{\tau}_1^c$

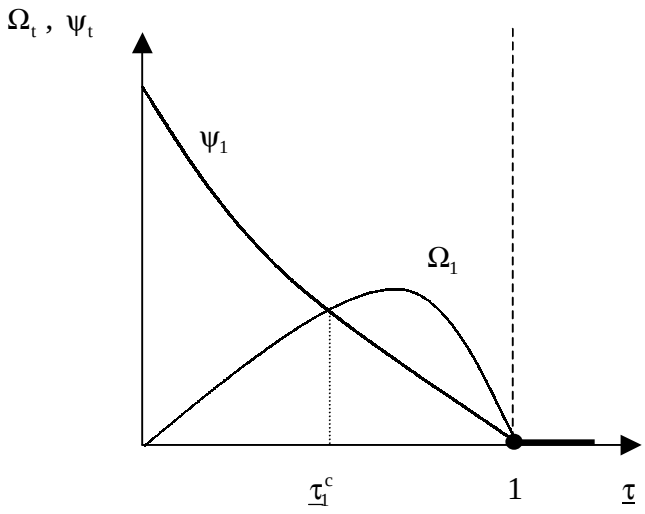


Figure 3: $\gamma' > \gamma \quad \delta' > \delta$

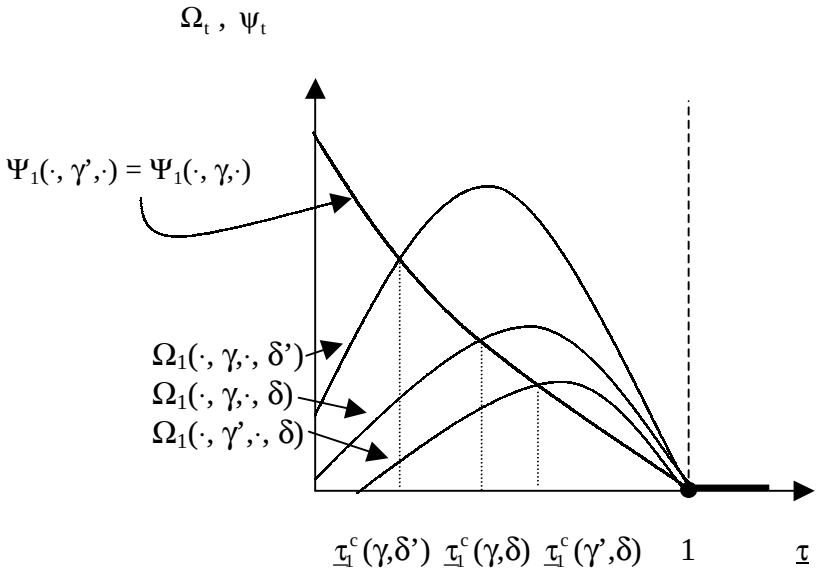


Figure 4: $Z_{d-1} > Z_{d-2}$

