Time Inconsistency of Trade Policy and Multilateralism

By: Mohammed Amin
Third draft: June 2003

Abstract: In this paper we address the issue of multilateralism versus bilateralism in a situation where a home government's optimal policy is time-inconsistent and the time-consistent policy sub-optimal. Short-run production and wage rigidities create incentive for the government to surprise private agents, *ex post*, with excessive protection. The problem is shown to be particularly severe when the intended beneficiaries (members of the lobby) cannot coordinate on their (investment) decisions fully. A commitment to multilateral free trade may then be welfare enhancing and politically feasible. However, once a discriminatory Preferential Trade Agreement is formed, it undermines any incentive for further multilateral trade liberalization. Thus, we propose another reason why Article XXIV of GATT/WTO may be a stumbling bloc for wider multilateral trade liberalization. Our result is based on the *trade deflection effect* and the *market power effect* of Preferential Trade Agreements.

Introduction

Two central issues in the context of discriminatory trade agreements such as Preferential Trade Agreements (PTAs) are their static welfare impact and the dynamic time path question. The former issue deals with whether PTAs are trade creating or trade diverting and the welfare implication of this, while the latter issue is concerned with the impact that such agreements may

Acknowledgements: The paper was submitted as Chapter 1 of my Ph.D dissertation at Columbia University, 2003. I would like to thank my thesis supervisor, Professor Kyle Bagwell, for his support and helpful comments. Special thanks to Professors: Atila Abdulkadiroglu, Ronald Findlay, Sanjay Reddy and Rajiv Sethi for their advice and helpful comments. All errors are my own.

have on the wider multilateral trade liberalization. In the context of the dynamic time path question, which is the focus of this paper, the literature has highlighted three broad themes. The first theme focuses on how the internal political support for multilateral trade liberalization is affected by regional integration (Levy, 1997)[1]. The second theme deals with how regional agreements alter the bargaining outcomes across countries in multilateral trade negotiations (Ludema, 1993). The third theme considers the impact of such agreements on the enforcement of multilateral trade cooperation (Bagwell and Staiger, 1997a, 1997b, 1999b). These studies have shown that regional agreements are a mixed bag for multilateralism. That is, while in some cases they further the cause of multilateralism, in others, they undermine it. Or, in Bhagwati's (1991) terminology, PTAs could be either building blocs or stumbling blocs for the wider multilateral cooperative efforts.

The stumbling-building blocs literature has focused primarily on the magnitude of trade liberalization rather than on efficiency. It does not provide an answer as to why discriminatory agreements that are, for example, stumbling blocs may reduce national or world efficiency. In fact, this problem is quite pervasive and applies in general to the Most Favored Nation (MFN) rule of the World Trade Organization (WTO). As Caplin and Krishna (1998, pp. 281-282) note:

> "There is a simple observation which illustrates the difficulties in providing a general bargaining-theoretic rationale for MFN. There is a grand utility possibility frontier available to countries using all the commercial trading instruments at their disposal, such as tariffs. If we view the bargaining process as yielding efficient outcome, as for example with the Nash bargaining solution, then MFN simply limits the tools available to different countries, shifting in the utility possibility frontier. Hence the most positive aspects of MFN can only be illustrated when the bargaining process absent-MFN yields inefficient outcomes."

[1] We use the terms regional integration, regionalism, discriminatory trade agreements and bilateral agreements synonymously with PTAs.

Some attempts have been made recently to address the concerns expressed above. McCalman (1998) explores the efficiency-enhancing role of MFN when bargaining occurs with private information. Bagwell and Staiger (1999c) analyze the commitment value of MFN when negotiations occur over time. They show that MFN, together with reciprocity, can support efficient outcomes that would, absent-MFN, be time-inconsistent and hence not sustainable. Further, Bagwell and Staiger (1999b) consider the enforcement problem in a repeated game framework and demonstrate how discriminatory agreements affect the level of (efficiency enhancing) cooperation that can be sustained at the multilateral level.

Notwithstanding the concerns expressed by Caplin and Krishna, we believe that the stumbling-building blocs literature serves to isolate and demonstrate important issues and results which, when embedded in an environment with inefficient multilateral cooperation, can provide richer insights. While the above-mentioned attempts are noteworthy and are important steps in understanding the welfare implications of regionalism, they are still in their infancy. For this reason, and to keep our model simple, we restrict our attention to the impact of discriminatory liberalization (via Article XXIV) on the incentive of a country to pursue further multilateral trade liberalization in a magnitude-sense rather than efficiency-sense.[2]

Another concern that has occupied the attention of trade theorists and policy makers is that optimal policies may be time-inconsistent. In fact, Kydland and Prescott (1977) have argued that optimal policies are almost always time-inconsistent and the time-consistent policies sub-optimal. A number of papers have discussed this problem in the context of trade policy. Staiger and Tabellini (1987) show that a benevolent government with redistributive goals will have an incentive to reverse its pre-announced tariff and surprise private agents with unanticipated protection. Under perfect foresight, such temptations yield an equilibrium outcome with an excessively high (sub-optimal) level of protection. Maggi and Rodriguez-Clare (1998) consider the interaction between a government and its private agents in the import-competing sector (the lobby) in a Nash bargaining framework. They

[2]See, for example, Levy 1997, for a similar approach.

point out that when the government lacks credibility in setting trade policy it leads to an inefficiently over-sized import-competing sector that can reduce the government's overall welfare relative to free trade, the contributions from the lobby notwithstanding. As one may expect, the presence of such commitment-related inefficiencies creates an incentive for the government to seek suitable credibility-enhancing devices. Staiger and Tabellini consider the choice between committing, *ex ante*, to a tariff-only versus a subsidy-only regime, while Maggi and Rodriguez-Clare focus on free trade agreements as possible remedies. Both these papers focus on "domestic" commitment problems. That is, the credibility problem between the government and its own (domestic) private agents. In contrast to this, Lapan (1988), Bagwell and Staiger (1999c) deal with commitment problems between the governments across countries.

The present paper attempts at establishing a link between domestic commitment problems and the impact of regional agreements on the functioning of the multilateral trading system. We consider a two-goods and three-country model. There are two foreign and one home countries. Throughout the paper our focus is on the home country. Production of each good in the home country requires a sector-specific factor and a common factor which is called labor. That is, labor can be used in either sector. The wage rate of labor and its allocation across the two sectors are determined endogenously through a perfectly competitive labor market. Following the political economy literature, we assume that owners of the specific factor in home's import-competing sector are politically organized and constitute a lobby. The home government maximizes a weighted average of (pure) national welfare and the welfare of the lobby. A critical feature of our model is that the home country is characterized by short-run rigidity in the wage rate and inter-sectoral mobility of labor. That is, we assume that while labor is perfectly mobile in the long-run, it is inter-sectorally immobile in the short-run. Similarly, the wage rate is assumed to be perfectly flexible in the long-run, but rigid in the short-run. This structure is quite standard in the trade literature except that we allow for wage rigidity as well. The motivation for this is discussed in detail in the sections that follow.

The presence of the above-mentioned rigidities implies that while output in each sector and the wage rate are flexible *ex ante* and *before* trade policy is implemented, they are fixed when trade policy is implemented and thereafter (*ex post*). Since a fully anticipated increase in home's tariffs leads to a welfare loss due to the conventional *production distortion* of a tariff, optimal policy takes this cost into account.[3] However, *ex post*, this cost is absent since production is fixed due to the labor-immobility mentioned above. Consequently, the government has an incentive to revise its pre-announced policy and surprise private agents with a higher level of protection. Thus, optimal policy is time-inconsistent. Foreseeing this, private agents employ more resources (labor) in the import-competing sector. The resulting equilibrium features an inefficiently high level of protection and an inefficiently over-sized import-competing sector. This constitutes one source of the commitment-related inefficiencies in our model.[4] The second source of commitment-related inefficiencies in our model is due to the short-run wage rigidity and the structure of the labor market. To see this, we first note that our assumption of perfectly competitive labor market implies that the lobby is *non-monopsonistic*. That is, while its members are able to overcome the free rider problem to influence trade policy, they do not coordinate on their (labor-hiring) decisions in the labor market. This leads to excessive competition between the members of the lobby for scarce labor with the resulting equilibrium wage higher than the monopsonistic-wage. From the government's point of view this "higher" wage constitutes a distortion since it attaches a higher weight to the welfare of the lobby vis-a-vis the laborers. A fully anticipated increase in home's future tariff increases demand for labor in the import-competing sector and thus enhances this distortion. Consequently, some of the benefits of protection are dissipated away from the lobby and towards labor. This reduces the home government's overall (and marginal) benefit from protection. We interpret this "dissipation of rents" as an extra cost that the government encounters when protection is fully anticipated. For brevity, will refer to this extra cost as the *wage distortion* of a fully anticipated increase in future protection. However, *ex post,* this cost is completely absent since the wage rate is rigid in the short-run. Incentive for surprise protection follows from this. This renders the optimal policy time-inconsistent. As we show in the

[3]Throughout the paper the term "optimal policy" will mean the "full commitment optimal policy" or the "*ex ante* optimal policy". This is consistent with the convention in the literature.

[4]This source of commitment problem has been widely discussed in the literature. See, for example, Maggi and Rodriguez-Clare, Staiger and Tabellini.

sections, the time-consistent equilibrium features excessive protection and a distorted wage rate which reduces the home government's overall welfare relative to the full-commitment case. We may note here that if the lobby behaved as a *monopsonistic* buyer in the labor market then the wage rate will be set optimally from the lobby's and the government's point of view. There will be no "distortion" in the wage rate in this case and thus, short-run wage rigidity will cease to be a source of commitment problem. We establish this argument formally in section 3 where we show that a necessary condition for commitment-related inefficiencies to arise from short-run wage rigidity is that the lobby is *non-monopsonistic* in the labor market. A related result here is that, in contrast to the findings in the literature, commitment problems can arise purely due to *wage distortion* even though *production distortion* may be completely absent.

Faced with these commitment-related inefficiencies the home government seeks suitable commitment-enhancing devices to bind its trade policy. As in Maggi and Rodriguez-Clare, we consider the use of free trade agreements as commitment-enhancing devices. Specifically, we assume that the government can commit itself to free trade agreements before private agents make their decisions and that such agreements are fully binding in the future. In this context we focus on two types of (free) trade agreements. That is, the home government can choose either a multilateral (non-discriminatory) free trade agreement or a bilateral (discriminatory) one in the form of a PTA with one of its trading partners. An important feature of such commitment-enhancing trade agreements is that they are trade-liberalizing as they require free trade between the member countries. The focus of the paper is to study whether this trade-liberalizing effect translates into multilateral trade liberalization and whether the bilateral option helps or hinders this process. In other words, when trade liberalization is driven by credibility enhancing concerns then does the PTA-option act as a stumbling or a building bloc for wider multilateral trade liberalization? Before elaborating on the motivation for focusing on free trade agreements as commitment-enhancing devices, it will be helpful to discuss a real world example.

It is argued that for Mexico the most important aspect of joining the North American Free Trade Agreement (NAFTA) is that it will provide credibility to its trade liberalization policy. Expressing this optimism, an executive from a pulp mill in Chihuahua, Mexico, said:

> "Policies in Mexico have always changed when presidents did, but free trade gives a sense or permanence to the very sound policies of this administration."[5]

Such credible trade agreements not only reduce protectionist forces in the short run but generate trade-liberalizing forces in the long run too. For instance, Gould (1992, p.22) states that:

> "The benefit of the proposed free trade agreement between the United States, Canada and Mexico will derive from not only a decrease in tariff and non-tariff barriers but also from a credible commitment that future trade barriers will not be erected. It is the expectation of lasting free trade, in addition to low trade barriers, that will entice long-term investment away from protected import-competing sectors and into the export sectors where a country's comparative advantage lies."

A relevant question to ask here is about the commitment value of WTO-bindings as compared to explicit trade agreements like NAFTA. That is, why couldn't Mexico use its tariff and non-tariff bindings at WTO as an effective commitment device against the domestic protectionist forces? Do regional agreements have a higher commitment value than WTO-bindings?

The statements cited above seem to suggest that this may be the case, at least for Mexico. While GATT/WTO has been successful in lowering trade

[5] Rodolfo Figueroa of Grupo Chihuahua, quoted in "Latin Turnaround", *Wall Street Journal*, May 24, 1991, p1.

barriers through successive rounds of negotiations, however, its members continue to use the safeguard provisions to restrict trade when domestic pressure for protection rises. Use of Antidumping duties, Voluntary Export Restraints and other "grey-area" measures still continues. For instance, Staiger (1995, p. 1538) states:

> "If the extended decline in multilateral tariff levels represents one major feature of post-war trade policy, the move toward "special" protection which has accompanied it represents a second. ... the possibility of temporary reversions to high protection or on occasion even permanent reversal of previous reductions was anticipated to be a part of the natural process of liberalization, and was viewed as an inevitable consequence of low levels of baseline protection envisioned by GATT. But while the safeguards included in GATT rules were meant to keep the inevitable deviations from rigid tariff bindings within GATT rules, in practice the growing use of VERs, VIEs, OMAs, and tariff programs that are tailor made to suit the needs of particular sectors signals a failure of the rules to contain these actions, and has given rise to the term "managed trade"."

Further, referring to the theoretical literature on commitment problems in trade policy, he states (Staiger, 1995, p. 1540):

> "But while these papers all suggest that GATT rules could be helpful as a commitment device when confronting injured sectors, it is in precisely these situations that GATT commitments can be suspended under the safeguard provisions of the agreements."[6]

[6] Staiger refers to Staiger and Tabellini (1987), Matsuyama (1990), Brainard (1994) and Mayer (1994) here.

From the discussion above we may infer that countries may seek options beyond WTO-bindings to enhance the credibility of their trade policies. We refer to these as explicit trade agreements with NAFTA as a prime example. As discussed above, we restrict our attention to those agreements which require complete elimination of all trade barriers between countries that join the agreement.[7] Such agreements may be built, at the outset, with stringent safeguards which would make them difficult to be exploited, *ex post*, for "special-interest" concerns. This point has been noted in the literature. For example, explaining the widespread use of escape clauses and safeguard actions granted to troubled industries, Maggi and Rodriguez-Clare (1998, p. 595) state:

> "Second, our model examines a government's *demand* for commitment to free trade; therefore, an alternative interpretation of the observed exceptions to free-trade agreements is that the international organization may not be strong enough to enforce free trade under all circumstances. It is conceivable, for example, that NAFTA grants fewer exceptions than the WTO because enforcement in NAFTA is more effective."

We would like to mention here that we do not imply that regional agreements or explicit trade agreements are necessarily and always more credible than WTO-bindings. Instead, we intend to restrict our attention to the scenario where a country seeks explicit (free) trade agreements precisely when these are better enforceable than the WTO-bindings. A full analysis of this issue, while interesting in itself is, however, beyond the scope of this paper. Our focus is on the implications for the conduct of trade policy in a situation where explicit trade agreements are desired for their commitment value. Alternatively, we may interpret our model as one involving two broad regime choices for a country: it can either opt for a regime characterized by fully credible explicit free trade agreement with any of its trading partner(s), or

[7] GATT/WTO rules require that a trade agreement outside the usual negotiating rounds must ensure complete elimination of all trade barriers between the countries involved in the agreement.

it may choose to conduct trade negotiations at the WTO negotiating-rounds without any prior commitment about the degree of trade liberalization that would be reached.[8]

Before proceeding further, it will be useful to summarize the basic arguments in the following simple points:

> (a) Government's face credibility problems vis-a-vis their domestic private agents in setting their trade policy. This leads to a welfare loss through a sub-optimally high level of protection. Commitment-enhancing devices can help reduce this inefficiency.
>
> (b) When WTO-bindings do not serve as effective commitment devices a government may seek explicit free trade agreements that are durable and better enforceable. These agreements could be either multilateral (non-discriminatory) or bilateral (discriminatory).
>
> (c) Such trade agreements can help a government foreclose political pressure and reduce the inefficiency mentioned in (a) above.
>
> (d) In light of these observations this paper seeks to address questions of the following kind: If the option of joining a discriminatory agreement like NAFTA were not available to Mexico, then would it commit to world-wide free trade to overcome its commitment problems? Having joined NAFTA, what incentive does Mexico have to pursue multilateral free trade to further bind its trade policy against the domestic protectionist forces? In general, we analyze whether discriminatory trade agreements are stumbling or building blocs when trade agreements are sought primarily to overcome domestic commitment problems.

[8] The literature on trade policy has frequently utilized such distinctions between WTO-bindings and explicit (regional) agreements. For instance, models dealing with enforcement problems in a repeated game framework have typically assumed that regional agreements (with free trade within the bloc) are strictly binding while inter-bloc tariffs must be self-enforcing. See, for example, Bagwell and Staiger (1999b); Bond, Syropoulos and Winters (2001).

The bridge between domestic commitment problem and PTAs being stumbling or building blocs can be easily summarized by first noting that a PTA between home and, for example, country A produces the *market power effect* and the *trade deflection effect*. The former implies that the PTA increases home's *market power* in trade against the non-member country which generates larger term-of-trade related gains from protectionism for home against the non-member country. The *trade deflection effect* arises because a higher tariff by home on its imports from the non-member country shifts its equilibrium import-volume away from the non-member country and towards the member country. With discriminatory tariffs in place, this reduces home's total tariff revenue and thus its (marginal and total) benefit from protectionism.

The basic results and structure of the paper can be summarized using the elements stated above as follows. We first show that commitment problem is necessary for a government to seek a multilateral free trade agreement (*Proposition* 2). Next, we identify sufficient conditions when, absent-PTA option, such a multilateral agreement is preferred over no trade agreement (*Proposition* 2). This holds when commitment inefficiency is sufficiently large relative to benefits (terms-of-trade and political) from an active protectionist policy. These results imply that, absent PTA-option, a worldwide multilateral free trade would result. We next allow for the bilateral option. Our results show that if home forms a PTA with a sufficiently large country then the *trade deflection effect* eliminates its commitment problem (*Proposition* 3). This occurs because, *ex post*, the *trade deflection effect* increases home's cost of revision in its (*ex ante* optimal) tariff sufficiently so that the full commitment optimal protection level is time-consistent too. Simultaneously, the *market power effect* ensures that home's benefit from an active protectionist policy against the non-member country is strictly positive, the higher cost due to the *trade deflection effect* notwithstanding (*Proposition* 4). Consequently, absent-PTA option, home has incentive to liberalize multilaterally; however, once a PTA is formed, this incentive is completely eliminated and an active protectionist policy against the non-member country is instead chosen. Thus, domestic commitment problems render PTAs stumbling blocs for wider multilateral trade liberalization.

The remainder of the paper is organized as follows. In section 1 we introduce the basic elements of the model. In section 2 we derive the time-consistent equilibrium solution for various trading regimes that we consider. The commitment problem and the full commitment optimal policy is derived in section 3. Putting these elements together we state our main results in the same section. In the conclusion we discuss some generalizations of the model and related issues.

Section 1: The model

We assume that there is one home country and two foreign countries. The foreign countries are called A and B. There are two goods labelled X and Y. To simplify the exposition of our findings we suppose that each foreign country trades only with the home country, who imports X from each of its two foreign trading partners in exchange for exports of Y. The home country is thus the only country that has the opportunity to choose between discriminatory and non-discriminatory trade agreements.[9] To focus on the link between domestic commitment problems and demand for trade agreements we assume that the tariffs of the foreign countries are zero. Home country will demand free trade agreement if and only if its commitment-related inefficiency is sufficiently large. In the conclusion we discuss how the qualitative nature of our results can be extended when the foreign countries set their tariffs optimally.

We now introduce the price notations. The home local relative price is denoted as $p \equiv p_x/p_y$, where $p_x(p_y)$ is the local price of good $X(Y)$ in the home country. Similarly, the local relative price in foreign country J is denoted as $p_J \equiv p_x^J/p_y^J$, where $p_x^J(p_y^J)$ is the local price of good $X(Y)$ in foreign country J =A,B. We define the "world" (i.e., untaxed) relative

[9] High transportation costs between the two foreign countries may render trade infeasible between them. While our results do not depend on this assumption, it allows us to keep the model simple and draw sharp results. Similar structure has been used in the literature. See, for example, Bagwell and Staiger, 1999c.

price for trade between home and foreign country J as $p_J^w \equiv p_x^J/p_y$. This is the ratio of exporter prices for trade between home and foreign country J. We will refer to it as home's bilateral terms-of-trade with foreign country J. Home's tariff on its imports from foreign country J is denoted by T_J which is assumed to be a non-negative specific tariff. Throughout we will focus on the non-prohibitive values of T_J. We define its value in terms of good Y as $t_J \equiv T_J/p_y$. Home's good Y is used as a common numeraire throughout the model.[10] We rule out export subsidies.

1.1 *Endowments and Consumer preferences*

The total number (measure) of agents, and equivalently consumers, in the home country is denoted by L and in foreign country J by L_J. Consumers in the world have identical preferences given by the utility function $u(c_{ix}, c_{iy}) \equiv \alpha c_{ix} - (1/2)c_{ix}^2 + c_{iy}$ where $c_{ix}(c_{iy})$ is the consumption level of good $X(Y)$ by the i^{th} consumer and α is a strictly positive parameter.[11] Individual demand function for good X in the home country is given by $c_{ix}(p) \equiv \alpha - p$ and in foreign country J by $c_{ix}(p_J) \equiv \alpha - p_J$.

To derive the remaining structure we first need to define the endowments and production process in the three countries. To this end, we assume that the two foreign countries are pure exchange economies with a fixed endowment of the two goods. Each agent in foreign country J is endowed with x_J units of good X and y_J units of good Y. National income of foreign country J when measured in terms of good Y is equal to $I_J \equiv L_J(y_J + p_J x_J)$. Production in the home country is endogenously determined within a specific factors framework with R, K and L being the three factors of production. We assume that $R(K)$ is a specific factor used in the production of good $X(Y)$ only. L is the common of production in the two sectors and is interpreted to be either labor or as another input. The total amount (measure) of factors R, K and L is exogenously given and denoted by R, K and L, respectively. Let $L_x(L_y)$ denote the amount of labor employed in sector $X(Y)$ in the home country. Assuming constant returns to scale production function in each sector, we have

[10] Since home's tariffs are assumed to be non-prohibitive and foreign tariffs are zero, it follows that the absolute price of good Y will be equal across the three countries in equilibrium. Hence, we can use any country's good Y as our common numeraire.

[11] The symmetry of consumer preferences is not important for our results.

that output of good X in the home country equals $F = F(L_x) \equiv L_x^{b_1} R^{1-b_1}$ and that of good Y is equal to $G = G(L_y) \equiv L_y^{b_2} K^{1-b_2}$. The parameter b_1 is treated as exogenously given with $0 < b_1 < 1$. The same holds for b_2. We may note here that with constant returns to scale in each sector the specific factors are always fully utilized in equilibrium. We have imposed this restriction in the output levels defined above in the $F(.)$ and $G(.)$ functions. Marginal productivity of labor in sector X, Y is given by $F'(L_x) \equiv \partial F(L_x)/\partial L_x$ and $G'(L_y) \equiv \partial G(L_y)/\partial L_y$, respectively. For future use let $F'' \equiv \partial^2 F(.)/\partial L_x^2 < 0$ and $G'' \equiv \partial^2 G(.)/\partial L_y^2 < 0$.

We next introduce the following notations:
$f \equiv F/L$ which is home's average (per home-agent) output of good X
$M_J \equiv$ home's import-volume of good X from foreign country $J = A, B$
$TR \equiv$ home's total tariff revenue measured in terms of good Y which is assumed to be distributed back to home's consumers (agents) in lump-sum fashion.

Indirect utility of the i[th] agent in the home country with income level I_i equals $I_i + s(p)$ where $s(p) \equiv (\alpha - p)^2/2$ and is the agent's (consumer) surplus from the consumption of the non-numeraire good. Home's national income, inclusive of tariff revenue if any and measured in terms of good Y, is equal to $I \equiv G + pF + TR$. National welfare of the home country is the aggregate of (indirect) utility of all its agents and equals $V \equiv I + Ls(p)$.

From the above structure we can compute the aggregate demand, export supply and import demand functions in each country for the two goods as follows:

$C_x(p) \equiv L(\alpha - p)$ is home's aggregate demand for good X
$E_J \equiv L_J(x_J - \alpha + p_J)$ is foreign country $J's$ export supply function
$M \equiv M_A + M_B = L(\alpha - p - f)$ is home's import demand function
$C_{Jx}(p_J) \equiv L_J(\alpha - p_J)$ is foreign country $J's$ aggregate demand for good X
Home's aggregate demand for good Y equals $I - pC_x(p)$ and the same for foreign country J equals $I_J - p_J L_J(\alpha - p_J)$.

1.2 *Relative size of the three countries*

The relative size of each country is defined as the number of agents in the country relative to the whole world. Home's (relative) size equals $\lambda \equiv L/(L + L_A + L_B)$, and similarly for foreign country J this equals $\lambda_J \equiv L_J(L + L_A + L_B)$. There will be no harm in normalizing the total number of agents in the world to unity although it will not be necessary. Throughout the paper we will mean by the "size" of a country as the relative size of the country.

1.3 *Description of the home country and the structure of the game*

Structure of the home country is characterized by a three-stage game which are as follows:

Stage 1

In the first stage the home country's government decides whether to form a Free Trade Agreement (FTA) with both A and B, or a PTA with country A only or no agreement at all.[12] Throughout the model we will refer to these three possibilities as the "FTA regime", "PTA regime" and the "MFN regime", respectively. To focus on the incentive-structure of the home country we assume that if an FTA or PTA offer is made to either A or B (or both, as the case may be) then it will be readily accepted by them. Trading regime (FTA, PTA or MFN regime) chosen in this stage is fully binding in the future as discussed in the Introduction. The outcome of stage 1 is fully observed by all the private agents.

Stage 2

Owners of the specific factors and laborers negotiate over wages and employment in a perfectly competitive environment. Firms (owners of factors R, K) decide how much labor to hire and workers decide in which sector to work. Wages are determined through competitive bidding. We introduce two features which will be critical for our model. (i) Firstly, we assume that labor, while mobile in stage 2, however, is immobile (across sectors) in stage 3. (ii) Wages negotiated and accepted in stage 2 are irreversible in stage 3 in the sense discussed below. These two constitute the two rigidities in our model. Two key points ought to be noted here. Firstly, in the context of our model,

[12] There is no loss of generailty in choosing country A only as home's possible PTA-partner.

it is important that wages are not fixed entirely in terms of good X alone. In our model they are fixed in terms of the numeraire good Y. However, it will not make any qualitative difference to our results if they were fixed in terms of both the goods. Secondly, complete short-run rigidity in labor movement (across sectors) and wages is not essential. Even small levels of rigidities would suffice for all our results.[13]

A simple motivation for the rigidities mentioned above could be that the process of hiring labor and negotiating wages takes time and cannot be done in stage 3 which is a short-run period. Other motivations can be explored. We outline one alternative structure below and argue that the equilibrium reached therein is exactly the same as with (i) and (ii) above.

Consider an alternative structure where labor is inter-sectorally mobile in the long-run (stage 2) but immobile in the short-run (stage 3). This may be due to, for instance, sector specific training costs which are prohibitively high in the short-run but negligible in the long-run. We assume that the market for labor is perfectly competitive and all firms within each sector are symmetric. Labor is hired on a contractual basis. A contract between a firm and a laborer specifies the wage rate and a promise to supply labor for one period of time. We allow firms to hire labor in stage 2 as well as in stage 3. Note that in stage 3 a firm can increase its labor force by extracting laborers from another firm in the same sector only since inter-sectoral movement of labor is ruled out in stage 3. We assume that labor contracts, once agreed upon, cannot be revoked in the future except in the following two situations. The first situation is when the total earnings of a firm (determined in stage 3) are less than its total cost of labor. We assume that in such a scenario the ownership of the specific factor of the firm is transferred to its laborers. The second situation arises from our assumption that if a firm wants to sell its labor force to another firm then it is perfectly free to do so, provided that the laborers are paid the wage rate agreed upon previously. Note that given the symmetry of firms within each sector, inter-sectoral immobility of labor in stage 3, atomistic agents, perfect foresight with no uncertainty

[13] Small levels of rigidities can modeled by simply allowing for slightly higher cost in revising the wage rate and in moving labor across sectors in the short-run as compared to the long-run. We believe this to be consistent with the real world.

or incomplete information, it follows that in equilibrium neither of these two cases would arise. This structure is equivalent to assuming that on accepting a contract, the firm pays an amount to the laborer and in return the ownership of labor (services) is transferred to the firm for one period of time. Hence, any gain (or loss) in the value of labor (after a contract is signed and accepted) accrues fully to the employers.[14] This is a simplifying assumption and all our results will continue to hold even otherwise provided that firms in the import-competing sector share some part of these gains (positive or negative). A sufficient condition for this is that wages are not fixed in terms of good X alone. Hence, our assumption that wages are fixed in terms of good Y alone should be viewed as a simplifying assumption. The equilibrium in the factor market here is characterized by firms hiring all their labor force in stage 2, no inter-sectoral movement of labor in stage 3, given symmetry of firms in each sector there will be no intra-sectoral movement of labor and wages in stage 3 and, firms will find it optimal to use all the labor that they hire in stage 2. This equilibrium is exactly the same as under (i) and (ii) above.

Rigidities in wages (and other prices) have been considered by economists as an important reason for surprise policy shocks rendering optimal policies time-inconsistent. While this issue has occupied center-stage in the macroeconomics literature, it has been largely neglected by the (commitment related) literature in trade policy. Our motivation in incorporating wage rigidities is that wages are critical in determining the distribution of income (or equivalently, the size of "political contributions" as in Grossman and Helpman, 1994, 1995) which plays a central role in determining optimal policies. We show that anticipated and surprise policy shocks have different effects on equilibrium wages and thus on a government's overall welfare. Commitment related inefficiencies and demand for credibility-enhancing trade agreements can arise purely due to wage rigidities even though the conventional *production distortion* of tariffs may be completely absent.[15]

[14] The change in the value of labor may arise due to, for example, unanticipated price shocks in the future.

[15] The literature dealing with commitment problems in trade policy has exclusively focused on short-run production rigidities and the conventional *production distortion* of the tariff as the reason why a government may prefer surprise protection to fully anticipated one. We suggest that wage rigidities and *wage distortion* in the sense discussed in this

Stage 3

In this stage trade policy is implemented, previously accepted labor contracts are fulfilled as discussed above, production is realized, markets for the two goods are cleared and equilibrium established. Consumption occurs and welfare of all agents is realized. We assume that all this occurs simultaneously as the sequential structure here is irrelevant. Markets for the two goods are assumed to be perfectly competitive. Implementation of trade policy in this stage is governed by two simple rules: firstly, trade agreements reached in stage 1 (FTA or PTA) are binding and duly implemented; secondly, home's tariffs on the non-member countries (if any) are set optimally from the home government's point of view and are strictly MFN. Thus, a PTA is the only form of discrimination that is allowed in our model. The first rule is irrelevant for our MFN regime, while the MFN restriction in the second rule is irrelevant for the PTA regime. To determine home's optimal tariff on the non-member countries we introduce the home government's objective function.

Home government's preferences

Following the political economy literature, we model the home government's objective function as politically motivated. That is, we assume that the government maximizes a weighted average of national welfare and the welfare of special interest groups.[16] Specifically, we consider the case when owners of the specific factor R in the import-competing sector are politically organized. We assume that they can overcome the free-rider problem and constitute a lobby to influence their government's trade policy. While we do not explicitly model political contributions offered by the lobby to the government, our approach is equivalent (reduced form) to the one in Grossman and Helpman (1994, 1995). For simplicity, we will assume that the ownership of factor R is highly concentrated so that the total earnings of the lobby is simply the total producer surplus in sector X.[17] We will denote this by $r \equiv pF - wL_x$ where $w \equiv$ equilibrium wage rate determined endogenously in the model.

paper can be equally important if the government is sensitive to the distribution of income.

[16] For discussion of this literature, see Bagwell and Staiger 1999a, 2002 (Chapter 2). The formulation adopted here is analogous to those used by Bagwell and Staiger 2001,2003 and Baldwin 1987.

[17] This assumption is explored in Grossman and Helpman, 1994, pp. 846-847.

Home government's overall welfare equals

$$\Omega \equiv V + \beta r$$

where β is a non-negative parameter (weight) and captures the government's political preference. A higher value of β implies that the government is more concerned about the lobby's welfare relative to (pure) national welfare. Equilibrium values of V, r are derived in the next section. We may note here that the members of the lobby are atomistic in the labor market and do not coordinate their (labor-hiring) decisions. That is, they are assumed to be competitive and not monopsonistic in the labor market. For future reference we will call this the *non-monopsonistic* nature of the lobby.

We now proceed to section 2 to establish the equilibrium of the game and interpret its properties.

Section 2: Solution of the game

We solve the model using backward induction method. It will be convenient to first derive the expressions for equilibrium prices and trading volumes for arbitrary values of t_A, t_B, f, G, w, L_x. We will call this the *trading equilibrium* of the game. The solution will allow us to solve easily for the equilibrium in each trading regime from a common unified structure.

2.1 *Trading equilibrium*

Let $\tau \equiv (t_A, t_B)$ denote an arbitrary vector of home's tariffs. Treat τ, f, G, w, L_x as exogenously given and fixed for the time being. Equilibrium values of the good's prices and trade-volumes can be solved for from the following market clearing and linkage conditions:

$$L(\alpha - p - f) = L_A(x_A - \alpha + p_A) + L_B(x_B - \alpha + p_B)$$
$$p = t_A + p_J^w \text{ and } p_J = p_J^w \;\; J = A, B$$

The left-hand-side (LHS) of the first of the previous two conditions is home's import demand function for good X and the right-hand-side (RHS) of the same is the total export-supply of good X by the two foreign countries. The second condition is the usual linkage condition connecting home's equilibrium local price and the bilateral world prices.

Equilibrium values of the endogenous variables described above are denoted and given by the following functions:

$$p = p(\tau, f) \equiv \alpha - (\lambda f + \lambda_A x_A + \lambda_B x_B) + \lambda_A t_A + \lambda_B t_B \quad ......(1)$$

$$p_J^w = p_J^w(\tau, f) \equiv \alpha - (\lambda f + \lambda_A x_A + \lambda_B x_B) - (1 - \lambda_J)t_J + \lambda_i t_i \;\; i \neq J = A, B$$
$$p_J = p_J(\tau, f) \equiv p_J^w(\tau, f)$$
$$M = M(\tau, f) \equiv L(\lambda_A x_A + \lambda_B x_B - (1 - \lambda)f - \lambda_A t_A - \lambda_B t_B)$$
$$E_J = E_J(\tau, f) \equiv L_J(x_J - (\lambda f + \lambda_A x_A + \lambda_B x_B) - (1 - \lambda_J)t_J + \lambda_i t_i) \;\; i \neq J = A, B$$

$\lambda f + \lambda_A x_A + \lambda_B x_B$ is the average endowment of good X in the world and α is the average world demand for good X when its price is equal to zero in each of the three countries. We assume that the latter is strictly higher than the former so that under free trade all equilibrium prices and strictly positive.

For the remainder of the paper we will use the notation p to mean home's local (relative) equilibrium price as defined by equation (1) above. Similarly, p_J, p_J^w, M, E_J will mean equilibrium values of these variables as defined above by $p_J(.), p_J^w(.), M(.)$ and $E_J(.)$ functions, respectively.

We now proceed to solve for the equilibrium of the stage-game described in section 1.

2.2 *MFN regime solution*

In stage 3 production takes place in the home country, labor contracts (payment of wages) are fulfilled, the home government implements its optimal MFN tariff and markets are cleared. Since production levels of X, Y are completely determined by decisions made in stage 2 of the game, we treat these as fixed here. The solution of stage 3 will be complete by deriving home's *ex post* optimal tariff and computing equilibrium prices and trade-volumes. To this end, home's *ex post* optimal MFN tariff is given by the solution to the following optimization problem:

Maximize $pF + t_A E_A + t_B E_B + Ls(p) + \beta pF$
by choosing t_A, t_B subject to the constraint that $t_A = t_B$, $p = p(.)$, $E_J = E_J(.)$ $J =$A,B as derived above and F treated as exogenously fixed.

The first order optimization condition for the *ex post* optimal MFN tariff, denoted by t, is given by:

$$-tL \sum_{J=A,B} \frac{\partial p(.)}{\partial t_J} + \sum_{J=A,B} E_J + (1+\beta)F \sum_{J=A,B} \frac{\partial p(.)}{\partial t_J} - C_x(p) \sum_{J=A,B} \frac{\partial p(.)}{\partial t_J} = 0$$

.....(2A)

LHS of equation (2A) is the net benefit to the home government from a unit increase in its MFN tariff with F, G, w treated as fixed and all changes evaluated at the margin. It is decomposed into the following familiar terms. The first two terms together capture the change in home's total tariff revenue, the third term is the change in producer surplus in sector X weighted by $1+\beta$ to reflect the government's political preference. The last term is the change in consumer surplus from the consumption of good X in the home country.

It can be easily checked that the second order maximization condition is satisfied. For interior solution conditions see Appendix (A5).

Dividing equation (2A) throughout by L and substituting the solution from the *trading equilibrium* we get the following implicit solution for the home's *ex post* optimal MFN tariff:

$$t = \beta f + [\lambda/(1-\lambda)]M(t,t,f)/L \quad ....(2B)$$

It will be useful to state the explicit solution for the optimal tariff as well. That is, substituting for $M(.)$ from section 2.1 in equation (2B) and rearranging we get the explicit solution for t as:

$$t = t(f) \equiv \frac{\lambda(\lambda_A x_A + \lambda_B x_B)}{1-\lambda^2} + \frac{\beta - \lambda}{1+\lambda} f \quad ....(3)$$

We now proceed to solve for stage 2 of the game which will give us the equilibrium value of f. Some important properties of equations (2B) and (3) will be discussed after that.

In stage 2 of the game factor markets are cleared and equilibrium wage rate and the allocation of labor across the two sectors in the home country is determined. Decisions of private agents in this stage are based on their expectations about the future equilibrium prices. To analyze this let $p^e \equiv$ the expectation that private agents hold (in stage 2) about home's future (stage 3) equilibrium local price. Factor market equilibrium is given by the following no-arbitrage and full employment conditions[18]:

[18] Full employment of factors of production is guaranteed since the factor markets are perfectly competitive and marginal productivity of each factor of production is strictly positive given our simple Cobb-Douglas production functions in the two sectors.

$$L_x + L_y = L,\ p^e F'(L_x) = w = G'(L_y)$$

These conditions imply that:

$$\frac{b_1 R^{1-b_1}}{b_2 K^{1-b_2}} p^e = \frac{L_x^{1-b_1}}{(L - L_x)^{1-b_2}}$$

Note that RHS of the previous equation is strictly and monotonically increasing in L_x, it approaches zero as L_x approaches zero and approaches infinity as L_x approaches L. This implies that, for any given finite and strictly positive value of p^e, the solution value of L_x is well defined, unique, strictly interior, and strictly increasing in p^e.

For convenience, let $L_x(p^e) \equiv$ value of L_x obtained by solving the previous equation. Using this we can state the solution values of L_y, w, f, F, G as functions of p^e, respectively, as follows:

$$L_y(p^e) \equiv L - L_x(p^e)$$
$$w(p^e) \equiv p^e b_1 [R/L_x(p^e)]^{1-b_1}$$
$$f(p^e) \equiv R^{1-b_1}[L_x(p^e)]^{b_1}/L$$
$$F(L_x(p^e)) = Lf(p^e);\ G(L_y(p^e)) = K^{1-b_2}[L_y(p^e)]^{b_2}$$

It can be easily checked that $dw(.)/dp^e > 0$ (with $b_2 > 0$) and $df(.)/dp^e > 0$. Further, for any given $p^e > 0$, $L_x(p^e), L_y(p^e), F(.), G(.)$ are homogenous of degree one in R, K and L. Lastly, $f(p^e)$ and $w(p^e)$ are homogenous of degree zero in R, K and L for any given $p^e > 0$.

Time-consistent equilibrium in the MFN regime

Equilibrium value of p^e is given by the dual condition that, *ex post*, agents' expectations are realized and the government does not have any incentive to surprise private agents. These dual conditions are captured by the following equation which can be used to solve for the unique time-consistent equilibrium value of p^e in the MFN regime:

$$p^e = p(t(f(p^e)), t(f(p^e)), f(p^e))$$

RHS of the previous equation can be read as follows. For any given value of p^e, we first compute $f(p^e)$ from the factor market equilibrium conditions stated above. Using this value of f we compute home's *ex post* optimal tariff, $t(f(p^e))$, from equation (3). Substituting these values of t, f in equation (1) we get RHS of the previous equation which is simply home's equilibrium local price in the MFN regime with its domestic production of X, Y and the MFN tariff set optimally for any arbitrary value of p^e. The equality of this with LHS is to ensure that expectations are realized *ex post*. The value of p^e that satisfies the previous equation is well defined, unique and strictly interior. For details, see Appendix A1.

Let the value of p^e obtained from the previous equation be denoted by p^*_{mfn}. Home's time-consistent MFN tariff is equal to $t^*_{mfn} \equiv t(f(p^*_{mfn}))$. The solution value of t^*_{mfn} is well defined, unique, stable and strictly positive. See Appendix A2 and A5 for details. The time-consistent equilibrium values of all the remaining endogenous variables can be derived by substituting t^*_{mfn}, p^*_{mfn} in the factor market equilibrium and the *trading equilibrium* solution outlined above.

We derived the time-consistent equilibrium above by first locating the fixed point of the $p(.)$ function in p, p^e space. Equivalently, we could start with expectations about the future MFN tariff (instead of the local price) and locate the equilibrium MFN tariff as a fixed point in t, t^e space. We briefly outline this procedure here as it will be useful in interpreting some of the results later in the section. To this end, let $t^e_J \equiv$ expectations that agents hold (in stage 2) about home's future tariff on its imports from foreign country J, with $J =$A,B. The MFN restriction will be imposed shortly. For any arbitrary values of t^e_A, t^e_B, expectations about future equilibrium price in home's local market must satisfy equation (1) in the sense that:

$$p^e = \alpha - (\lambda f(p^e) + \lambda_A x_A + \lambda_B x_B) + \lambda_A t^e_A + \lambda_B t^e_B$$

The previous equation follows from the assumption that agents are rational and use the "correct" model in predicting future (equilibrium) values of the variables, which is equation (1) for home's local price. From the previous equation we can uniquely solve for p^e as a function of t_A^e, t_B^e. Let this solution value be denoted by $p^e(t_A^e, t_B^e)$, which is well defined, interior and unique.[19] Impose the MFN restriction so that $t_A^e = t_B^e = t^e$ and $p^e = p^e(t^e, t^e)$. Home's ex post optimal MFN tariff is equal to $t(f(p^e(t^e, t^e)))$, where $t(.)$ is as given by equation (3). Time consistent MFN tariff is given by the value of t^e that solves: $t^e = t(f(p^e(t^e, t^e)))$ which is the usual fixed-point condition in t, t^e space. It can be easily checked that the solution value of t^e obtained from the previous equation is equal to t_{mfn}^* which was derived earlier. That is, t_{mfn}^* satisfies the fixed-point condition:

$$t_{mfn}^* = t(f(p^e(t_{mfn}^*, t_{mfn}^*)))$$

We may note here that $p^e(.)$ is strictly and monotonically increasing in t^e. That is, $dp^e(.)/dt^e = \dfrac{1-\lambda}{1+\lambda f'} > 0$, $f' \equiv df(p^e)/dp^e$.

Interpretation of results

A convenient property of the solution stated above is that home's *ex post* optimal MFN tariff for any given p^e, the time-consistent MFN tariff, t_{mfn}^*, and, the time-consistent equilibrium values of all prices (local and world), wage rate and the value of f, are each homogenous of degree zero in L, L_A, L_B, R, K. That is, absolute size of the three countries is irrelevant for these variables. The result follows directly from the basic structure of the model with aggregate demand and output levels of X, Y in each country being linear in its absolute size. Note that equilibrium output of X, Y in the home country are linear in its absolute size due to the constant returns to scale production functions for the two goods. This property of the solution is similar to the one in Syropoulos (2001) and allows us to focus on *relative* size of the three countries only while analyzing the time-consistent policy.

[19]For details, see Appendix A3.

The second feature of the solution is that home's *ex post* optimal MFN tariff, $t(f(p^e))$, can be either increasing or decreasing in the size of its import-competing sector captured by the value of f. It can be seen from equation (3) that $t(f)$ is strictly increasing in f if and only if $\beta - \lambda > 0$ and it is strictly decreasing (constant) if and only if $\beta - \lambda < 0$ $(= 0)$. The intuition for this is simple. Keeping home's MFN tariff fixed momentarily, an increase in the size of home's import-competing sector (value of f) reduces its equilibrium import-volume which lowers the marginal benefit to home from its tariff since the tariff now applies to a smaller import-volume base generating smaller gain in tariff revenue. Thus, home's *ex post* optimal tariff falls on this count. At the same time, with a higher value of f, home's equilibrium local price decreases which increases the consumption of good X. Consequently, the loss in consumer surplus form the marginal tariff is now higher which again lowers the *ex post* optimal tariff. The upward push to the *ex post* optimal tariff comes from a larger gain in (weighted) producer surplus from the marginal tariff due to a larger import-competing sector. The net effect of these factors on the *ex post* optimal tariff depends on the value of $\beta - \lambda$ as stated above in the paragraph. We may note here that the relationship between home's *ex post* optimal tariff and private agents' expectations about the future MFN tariff (t^e) is similarly determined. That is, since $p^e(t^e, t^e)$ is strictly increasing in t^e and $f(p^e)$ is strictly increasing in p^e, it follows that $t(.)$ is strictly increasing (decreasing) in t^e if and only if $\beta - \lambda > 0$ (< 0) and invariant if and only if $\beta - \lambda = 0$. This also establishes qualitatively the dynamic link between home's *ex post* optimal tariff and the private agents' expectations about home's future tariff in the MFN regime. The possibility of a negative relationship between the *ex post* optimal tariff and private agents' expectations (value of t^e, and/or $f(p^e)$) here is in sharp contrast to the findings in the literature, both theoretical and empirical. For example, Brainard and Verdier (1997) show that for a small country with endogenous protection, the level of tariff is an increasing function of past tariffs. Maggi and Rodriguez-Clare, Staiger and Tabellini, derive a similar relationship between the level of protection and the size of the lobby (value of f in our model). Further, Grossman and Helpman (1994), Goldberg and Maggi (1999) predict a negative relationship between the optimal tariff and the import penetration ratio in sectors that are politically organized. Our results show that these theoretical and empirical results will not hold for a large country when $\beta - \lambda < 0$. We put forward this finding as a large-country caveat to the small-country results in the literature.

Lastly, it will be useful to define home's *market power* in trade as it will feature frequently in the analysis that follows. Using the results stated above we have the following definition.

Definition: Home's *ex post market power* in trade in the MFN regime is equal to the improvement in its bilateral terms of trade against each foreign country from a unit increase in its MFN tariff, with its production of X, Y held fixed. This equals:

$$-\sum_{J=A,B} \frac{\partial p_A^w(.)}{\partial t_J} = -\sum_{J=A,B} \frac{\partial p_B^w(.)}{\partial t_J} = \lambda$$

As λ becomes smaller the home country becomes "smaller" in the sense that its *marker power* in trade decreases. When λ approaches zero then the home country approaches the text-book small-country case. The reverse holds when λ approaches one. Thus, λ is our measure of home's relative size and also a measure of its *ex post market power* in trade in the MFN regime.[20]

This completes our description of the time-consistent equilibrium in the MFN regime. To summarize, we derived the time-consistent equilibrium and noted some of its important properties. The dynamic relationship between current period (stage 2) decisions and expectations about home's future MFN tariff was discussed.

We now proceed to the solution of the PTA regime. Since the procedure is similar to the one above, we will omit some of the details.

[20] We need to distinguish between *ex-post* and *ex-ante* market power in trade. The latter is defined later in the section and shown to be different from the former.

2.3 *PTA regime solution*

In this regime the home country forms a PTA with country A in stage 1 which is binding for the rest of the game. Home's tariff on country B will be called its *external* tariff and it is determined endogenously. For any arbitrary value of home's *external* tariff, equilibrium prices and trade-volumes are given by setting $t_A = 0$ in the *trading equilibrium* solution outlined in section 2.1 above. A necessary condition for home's trade policy to be "operational" is that its equilibrium import-volume from country B is strictly positive under complete free trade. To ensure this we impose the restriction that $x_B > x_A$.[21] This assumption does not play any other role in our model. We will elaborate on this towards the end of section 3.

The solution for stage 3 of the game will be complete once we derive home's *ex post* optimal *external* tariff. This is given by the same optimization problem as the one for the MFN regime with the modification that we set $t_A = 0$ throughout and t_B is the only choice variable. The first order condition for this optimal tariff is given by:

$$-t_B(L + L_A)\frac{\partial p(.)}{\partial t_B} + E_B(.) + (1+\beta)F\frac{\partial p(.)}{\partial t_B} - C_x(.)\frac{\partial p(.)}{\partial t_B} = 0 \quad ...(4A)$$

The previous equation is derived and evaluated with $t_A = 0$ and f held constant. The value of t_B that solves this equation is home's *ex post* optimal *external* tariff.[22]

LHS of equation (4A) is the net marginal benefit to the home government from a unit increase in it's *external* tariff with all changes evaluated at the margin. It is decomposed into the following terms. The first two terms

[21] This assumption is necessary also for this purpose in the sense that if $x_B < x_A$ then home will import from country A only; further, if $x_B = x_A$ then home imports from country A only when its *external* tariff is strictly positive while under complete free trade the distribution of its import-volume across the two foreign countries is indeterminate. Either way, for the optimal *external* tariff to be strictly positive a necessary and sufficient condition is that $x_B > x_A$.

[22] It can checked that the second order optimization condition is satisfied under the interior solution conditions specified in Appendix A5.

together capture the change in home's tariff revenue from the marginal tariff, the third term is the change in the weighted producer surplus in the import-competing sector from the marginal tariff and, the last term is the change (loss) in consumer surplus from the consumption of the importable arising from the marginal tariff. These terms compare directly with the ones on LHS of equation (2A) which was used to derive home's *ex post* optimal MFN tariff. Before solving for t_B from equation (4A) it will be useful to state two critical effects which distinguish the PTA regime from the MFN regime. Following the literature we call these effects the *trade deflection effect* and the *market power effect* of PTAs and are as follows.

Trade deflection effect

Since trade policy is implemented in the third stage of the game, we treat output of X, Y in each of the three countries as exogenously fixed here. Under free trade between the home country and its PTA-partner, local prices in the two countries must be equal. That is, with $t_A = 0$ we must have $p = p_A$. With all changes evaluated at the margin, a unit increase in home's *external* tariff reduces its equilibrium import-volume from country B due to two reasons:

(i) Firstly, it raises the local price (p) and hence consumption of the importable drops by amount $L\partial p/\partial t_B$. With home's output of X, Y fixed, this implies that home's import-volume from country B decreases by the same amount ($L\partial p/\partial t_B$) which leads to a loss in its tariff revenue (and hence welfare) by amount $t_B L\partial p/\partial t_B$. This is usual *consumption distortion* of the tariff and appears as part of the first term on LHS of equation (4A).

(ii) Secondly, a unit increase in home's *external* tariff shifts home's import-volume away from country B and towards country A. The reason for this is that with $p = p_A$, the rise in the *external* tariff increases equilibrium value of p_A by amount $\partial p_A/\partial t_B = \partial p/\partial t_B > 0$ The export supply function of country A has a positive slope of L_A which implies that home now imports a larger amount from its PTA-partner. Consequently, home's import-volume from country B decreases by amount $L_A\partial p/\partial t_B$. The important point here is that this change in the composition of home's import-volumes across the two foreign countries reduces its tariff revenue (and hence welfare) at the margin by amount $t_B L_A\partial p/\partial t_B$. This is captured in the first term on LHS of equation

(4A) and it tends to lower home's *ex post* optimal *external* tariff. For brevity we will refer to this as the *trade deflection effect* of PTAs. Note that this effect is similar to the *consumption distortion* of the tariff as discussed in (i) above and appears symmetrically with it in equation (4A). The *trade deflection effect* is completely absent in the MFN regime (equation 2A) since, with MFN tariffs, home's total tariff revenue is independent of the distribution of its import-volume across its trading partners.[23]

Market power effect

Continuing to treat production of X, Y as exogenously fixed in the home country, we define home's *ex post market power* in trade in the PTA regime as the improvement in its bilateral terms-of-trade against country B from a unit increase in its *external* tariff. From the solution in section 2.1 we get that this equals $-\partial p_B^w(0, t_B, f)/\partial t_B = \lambda + \lambda_A$. Compared to the MFN regime, home's *ex post market power* in the PTA regime is higher by an amount which equals the size of its PTA-partner. This effect is well-known in the literature and arises because home's import demand function for good X from country B is more elastic in the PTA regime giving it higher *market power*. We note here that an immediate implication of this is that a unit increase in home's *external* tariff has a much smaller impact on its equilibrium local price (equal to λ_B) than a unit increase in its MFN tariff (equal to $\lambda_B + \lambda_A$). We will refer to both these results as the *market power effect* of PTAs in the remainder of the paper.

The *trade deflection effect* and the *market power effect* as identified above will play a key role in establishing our final result regarding the stumbling-bloc feature of PTAs in the presence of domestic commitment problems. We may note here that the *trade deflection effect* defined above is one component of the "tariff complementarity effect" identified in Bagwell and Staiger (1999b, p. 163). They note three reinforcing effects that generate complementarity across tariffs on imports of the same good from different countries. To see these effects in our model, hold home's tariff on country B fixed and lower its tariff on country A. Note that as a result of this, home's consumption of the importable will now increase which implies a greater loss in consumer surplus from its tariff on country B at the margin.

[23] This point is well noted in the literature. See, for example, Bagwell and Staiger, 1999a.

This is captured by the last term on LHS of equation (4A) where $C_x(.)$ rises as t_A is lowered. This is the first of the three effects in Bagwell and Staiger. Their second effect is captured in our model by a lower value of $E_B(.)$ when t_A is lowered so that home's external tariff now applies to a smaller import-volume base generating less tariff revenue. Their third effect is captured by our *trade deflection effect*. [24]

Continuing with the stage 3 solution in the PTA regime, the implicit solution for the *ex post* optimal *external* tariff is given by:

$$t_B = \frac{\beta f - (E_A/L)}{1 + (\lambda_A/\lambda)} + \frac{(\lambda + \lambda_A)/\lambda_B}{1 + (\lambda_A/\lambda)} E_B/L \quad ....(5)$$

The equation is evaluated at $t_A = 0$, f treated as fixed and, derived after substituting $C_x(.) - F = E_A + E_B$.

Comparing equation (5) and (4A) it is easy to check that the term λ_A/λ in the denominator of RHS in equation (5) is due to the *trade deflection effect* of PTAs as discussed above. The difference in the *ex post* incentive structure of the home government in the MFN and PTA regimes can be seen by comparing equation (2A) with equation (5). We note the following differences between the two. (i) Firstly, the *trade deflection effect* is completely absent in equation (2A) unlike in equation (5). The denominator in RHS of equation (2A) is simply 1 which captures the *consumption distortion* of the MFN tariff while the denominator in equation (5) reflects the sum of the *consumption distortion* of the tariff and the *trade deflection effect*.[25] (ii) Secondly, the *market power effect* is reflected in the difference between the coefficients of M/L and E_B/L in equations (2A) and (5), respectively. That is, in the MFN regime home's *ex post* optimal tariff depends on the ratio of its *market power* relative to that of the rest of the world which is captured

[24] For more details on this, see pages 162-163, Bagwell and Staiger, 1999b.

[25] Note that in equation (5) the denominator terms, $1 + \lambda_A/\lambda$, capture the *consumption distortion* and the *trade deflection effect* per unit increase in home's local price from the marginal *external* tariff (that is, after dividing by $\partial p/\partial t_B$), per agent in the home country and per unit of home's *external* tariff.

by $\dfrac{\lambda}{1-\lambda}$ term in equation (2A). In the PTA regime, home's *ex post market power* equals $\lambda + \lambda_A$ and that of country B equals λ_B. The impact of this change is reflected in the numerator term of the coefficient of E_B/L in equation (5). (iii) The third difference is captured by the E_A/L term in equation (5) which is absent in equation (2A). Intuitively, in the PTA regime, a unit increase in home's *external* tariff shifts home's demand for the importable towards its PTA-partner and worsens its bilateral terms-of-trade vis-a-vis the same country. Evaluated at the margin, this implies that the home country now pays more for its imports from country A which (per home-agent) equals E_A/L. This tends to lower home's *ex post* optimal external *tariff*. This effect is completely absent in the MFN regime due to the very definition of non-discriminatory tariffs.

Substituting for E_A, E_B in equation (5) and solving for t_B we get the explicit solution as:

$$t_B = t_B(f) \equiv \Delta_1 + \frac{\lambda}{\lambda + \lambda_A + \lambda_A\lambda_B + (1-\lambda_B)^2}(\beta - \lambda)f \quad .....(6)$$

$$\Delta_1 \equiv \frac{x_B - (1+\lambda)(\lambda_B x_B + \lambda_A x_A)}{\lambda + \lambda_A + \lambda_A\lambda_B + (1-\lambda_B)^2} > 0$$

The previous inequality follows from the interior solution conditions specified in Appendix A5.

Using the value of t_B as given by equation (6), setting $t_A = 0$ and using the *trading equilibrium* solution in section 2.1 we can derive equilibrium values of all the stage-3 endogenous variables for any arbitrarily given value of f.

The solution for stage 2 of the game (factor market equilibrium) and the time-consistent *external* tariff can be easily derived by following exactly the same steps as outlined earlier for the MFN regime. That is, factor market equilibrium, for any given value of p^e, is exactly the same as stated earlier. For the time-consistent *external* tariff, let $t^*_{pta} \equiv$ home's time-consistent *external* tariff. Set $t^e_A = t_A = 0$. We have that t^*_{pta} is given by the following fixed-point condition:

$$t^*_{pta} = t_B(f(p^e(0, t^*_{pta})))$$

The solution value of t^*_{pta} is well defined, unique, stable and strictly interior. For details, see Appendix B1.

Using the value of t^*_{pta} derived above, $p^e(.)$ function and the *trading equilibrium* solution of section 2.1, we can fully determine the time-consistent equilibrium values of all the remaining endogenous variables in the PTA regime. We may note here that the *market power effect* and the *trade deflection effect* are captured by the $t_B(.)$ function and hence are critical in determining the value of t^*_{pta}. This completes our discussion of the time-consistent equilibrium in the MFN and PTA regimes.

We now proceed to the next section where we derive the full commitment optimal tariff in the MFN and PTA regimes, discuss the sources of the commitment problem and establish our final results.

Section 3: PTAs as stumbling blocs

The structure of section 3 is as follows. In section 3.1 we derive the full commitment optimal MFN tariff of the home country. Using the results, we state the sources of the commitment problem. In section 3.2 we do the same for the PTA regime. Section 3.3 is our final section where we show how PTAs can be stumbling blocs due to domestic commitment problem.

3.1 *Full commitment optimal tariff in the MFN regime*

For brevity, we will refer to the full commitment optimal tariff as simply the "optimal" tariff. In the MFN regime, this is defined as the fully anticipated tariff that is announced before private agents make their stage-2 decisions and is fully binding in the future. The solution stated in section 2 can be easily used here by noting the following sequence of moves. Suppose that the home government announces a tariff vector (t_A, t_B). Since this is fully binding in the future, we set $t_J^e = t_J$ for J =A,B. Under *rational expectations*, private agents must predict the future equilibrium local price in the home market to be $q = q(t_A, t_B) \equiv p^e(t_A, t_B)$. As in section 2, the value of q is implicitly given by the following condition:

$$q = \alpha - (\lambda f(q) + \lambda_A x_A + \lambda_B x_B) + \lambda_A t_A + \lambda_B t_B$$

Factor market equilibrium conditions are exactly the same as in the previous section. That is, aggregate domestic production of good X equals $Lf(q)$, equilibrium wage rate equals $w(q)$ and the allocation of labor in sectors X, Y equals $L_x(q)$ and $L_y(q)$, respectively. Equilibrium world prices, trading volumes and welfare levels in the three countries can be derived from these values and the *trading equilibrium* solution of section 2.1. Thus, our solution will be complete once we determine the optimal values of t_A, t_B.

The optimal MFN tariff is given by the following optimization problem:

Maximize $qLf(q) + G(L_y(q)) + tM(t, t, f(q)) + Ls(q) + \beta(qLf(q) - w(q)L_x(q))$ by choosing t, with $q = q(t, t)$, $t \geq 0$.

The first three terms in the maximization problem together equal home's national income, the fourth term is home's total consumer surplus from the consumption of the non-numeraire good and, the last term is equal to the welfare of the lobby (producer surplus in sector X) weighted by the political parameter, β.

The first order condition for the optimal MFN tariff is given by:

$$-tL\sum_{J=A,B}\frac{\partial q(.)}{\partial t_J}+M+[(1+\beta)F-C_x]\sum_{J=A,B}\frac{\partial q(.)}{\partial t_J}+L(\beta r_2-tf')\sum_{J=A,B}\frac{\partial q(.)}{\partial t_J}=0 \quad .....(7)$$

where $r_2L \equiv (p^e f' - w)\dfrac{dL_x(p^e)}{dp^e} - L_x(p^e)\dfrac{dw(p^e)}{dp^e}$; all variables are evaluated at $p = p^e = q(t,t)$ and $t_J = t_J^e = t$ for J =A,B. For interior solution restrictions and the second order maximization condition, see Appendix (A4).

LHS of equation (7) can be interpreted as follows. The first three terms together capture the change in home government's welfare from a unit increase in its MFN tariff and evaluated at the original wage rate and output levels of X, Y. This is the usual sum of changes in home's tariff revenue, consumer surplus and weighted producer surplus from the marginal MFN tariff, evaluated at the original output level of the two goods. It the same as LHS of equation (2A) except that the impact on home's equilibrium price is different in two equations. That is, in equation (7) we have $\sum_{J=A,B}\dfrac{\partial q(.)}{\partial t_J} = \dfrac{1-\lambda}{1+\lambda f'}$ while in equation (2A) we have $\sum_{J=A,B} \partial p(t,t,f)/\partial t_J = 1-\lambda$. Note that the former term is strictly less than the latter term. The intuition for this difference is that when future policy is fully anticipated then production is endogenous to the policy announcement and a higher anticipated tariff increases domestic production of the importable. This implies a smaller increase in home's equilibrium local price than when production is fixed (as in equation (2A)). The last term on LHS in equation (7) captures the *wage distortion* and the *production distortion* of a fully anticipated increase in home's MFN tariff. These two distortions constitute the two sources of commitment problem in our model and will be discussed shortly after stating the solution for the optimal MFN tariff from equation (7).

Let $\hat{t}_{mfn} \equiv$ value of t that solves equation (7). We have:

$$\hat{t}_{mfn} = \frac{\beta(f + r_2)}{1 + f'} + [\lambda/(1-\lambda)]M(.)/L \quad .....(7A)$$

with RHS of the equation evaluated at the optimal MFN tariff, $\hat{t}_{mfn}$; and $p = p^e = q(.)$.

Sources of commitment problem

These can be seen by comparing the expressions for $\hat{t}_{mfn}$ and the time-consistent MFN tariff, t^*_{mfn}. Noting that t^*_{mfn} must be *ex post* optimal, it must satisfy equation (2B) which we restate here for convenience: $t^*_{mfn} = \beta f + [\lambda/(1-\lambda)]M(.)/L$, with all variables evaluated at t^*_{mfn} and $p = p^e = q(.)$. The last two equalities here follow from the fact that, in equilibrium, expectations must be realized and they must be internally consistent in the *rational expectations* sense. Comparing $\hat{t}_{mfn}$ and t^*_{mfn} we can see that if the home government maximizes pure national welfare ($\beta = 0$) then $\hat{t}_{mfn} = t^*_{mfn}$. That is, under the MFN restriction, short-run production and wage rigidities do not lead to commitment problems when the government maximizes pure national welfare. The intuition for this is simple. Home's full commitment pure national welfare maximizing MFN tariff is equal to the reciprocal of the weighted average of the export-supply elasticities of its trading partners, with the weights being the shares of the trading partners in home's total import-volume.[26] Evaluated at $\hat{t}_{mfn}$ (with $\beta = 0$), *ex post*, these elasticities and weights are completely unaffected by the production and wage rigidities in the home country. Consequently, there is no incentive for the home government to surprise private agents and the optimal policy is time-consistent. We may note here that this result does not depend on the linear structure of the model but is a general result.[27] It will be useful to see this result from equation (7) and (2A) which capture, respectively, the *ex ante* and *ex post*

[26]That is, for any value of home's MFN tariff t, $t^e_J = t$ for J =A,B and $p = p^e = q(.)$, we have: $[\lambda/(1-\lambda)]M/L = p^w(e_A E_A/M + e_B E_B/M)^{-1}$, where $p^w \equiv p^w_A = p^w_B$, $e_J \equiv (\partial E_J/\partial p_J)(p_J/E_J)$ for J =A,B is the elasticity of foreign country J's export-supply function. From this it follows that the optimal MFN tariff can be expressed as: $\hat{t}_{mfn}/p^w = (e_A E_A/M + e_B E_B/M)^{-1}$ where $\hat{t}_{mfn}/p^w$ is the optimal tariff expressed as an ad-valorem rate and all variables are evaluated at $\hat{t}_{mfn}$ and $p = p^e = q(.)$. RHS of the expression for $\hat{t}_{mfn}/p^w$ is the reciprocal of the weighted average of export-supply elasticities of home's trading partners as stated above. Note that none of the variables in $\hat{t}_{mfn}/p^w$ equation are affected by production and wage rigidities in the home country.

[27]See the previous footnote.

incentive structures in the MFN regime. Set $\beta = 0$ in both these equations. For $\hat{t}_{mfn} = t^*_{mfn}$ to hold it must be that LHS of equations (7) and (2A) are equal at $\hat{t}_{mfn}$. Since, *ex ante,* production is endogenous, a higher anticipated tariff leads to the usual *production distortion* which is captured by the f' term in equation (7). This distortion is completely absent from equation (2A) since, *ex post*, production is fixed. It follows from this that the LHS of equation (2A) will be higher than that of equation (7). However, this tendency is countered by the fact that, *ex post,* home's *market power* in trade, equal to λ, is strictly less than it's *market power, ex ante.* The latter equals $1 - \sum_{J=A,B} \partial q(.)/\partial t_J = \frac{\lambda(1+f')}{1+\lambda f'}$.[28] *Ex post*, a lower *market power* creates incentive for the home government to *lower* its tariff below the optimal tariff. These two effects are equal and opposite in sign and, together, eliminate any incentive for surprise protection. Thus, commitment problem is completely eliminated.

With the above explanation in place we now focus on the case when $\beta > 0$ and, without loss of generality, on the "politically optimal" tariff which is completely captured and defined by the first term on RHS of equation (7A) when policy is fully anticipated and by the first term on RHS of equation (2A) when it is unanticipated (the *ex post* case).[29] Noting that the time-consistent MFN tariff must satisfy equation (2A), it is straightforward to check by comparing equations (7A) and (2A) that $\hat{t}_{mfn}$ is not time-consistent. The time inconsistency of the optimal MFN tariff is fully captured by r_2 and f' in equation (7A) which are absent (equal to zero) in equation (2A). These terms can be interpreted as follows. A fully anticipated increase in home's future MFN tariff implies that private agents predict (and rationally so) a higher equilibrium price in home's local market (value of p^e). This implies a higher domestic production of good X (since $f' > 0$) and, consequently, lower import-volume for home. The marginal benefit from a unit increase

[28] Home's *ex ante market power* in the MFN regime is defined as the improvement in its bilateral terms-of-trade against each foreign country due to a unit increase in its MFN tariff when production in the home country is endogenous to the increase in the tariff rate. Note that under the MFN restriction there is no difference between home's bilateral and multilateral terms-of-trade.

[29] Politically optimal tariff is the tariff that the home government finds optimal to implement if it were, hypothetically speaking, to neglect the terms-of-trade benefits. For more details, see Bagwell and Stagier, 1999a.

in home's MFN tariff is hence lower because the lower import-volume generates smaller gain in tariff revenue. *Ex ante*, the home government takes this cost into account which is the usual *production distortion* of the tariff and consequently its (full commitment) optimal tariff is lower on this count. *Ex post*, production is fixed and thus this cost is completely absent. With a lower marginal cost of its tariff, *ex post*, the government has an incentive to surprise private agents with a higher tariff. This constitutes one source of the commitment problem in our model. The second source of the commitment problem arises due to the short-run wage rigidity. That is, when future policy is fully anticipated, a small increase in home's future tariff implies that firms in home's import-competing sector like to produce more and hence demand more labor. As a result, more labor is employed in sector X and the equilibrium wage rate increases. The impact of these two changes on the total surplus of the lobby equals $r_2 L = [q(.)dF(.)/dL_x - w]dL_x/dp^e - L_x dw/dp^e$. Note that since firms are perfectly competitive in the labor market, it follows that $p^e dF/dL_x = w$ (factor market equilibrium condition as stated in section 2) and, $q(.) = p^e$ when future policy is fully anticipated. Substituting these equations in the expression for $r_2 L$ we get: $r_2 L = -L_x dw/dp^e < 0$. That is, when future policy is fully anticipated then the first term in the expression for $r_2 L$ is equal to zero. The intuition for this is simple. For any given wage rate, firms hire labor optimally. Consequently, L_x is always at its optimal level with respect to the surplus of the lobby. Thus, any infinitesimal revision in the value of L_x generates a *second order effect* only on the lobby's welfare. The home government's (full commitment) optimal tariff is completely unaffected by the impact of the change in L_x on the lobby's welfare arising from a higher anticipated future tariff. This holds *ex post* also. That is, *ex post,* L_x is fixed and thus the *ex post* optimal tariff is independent of the impact of the change in L_x on the lobby's welfare. We are now left with $-L_x dw/dp^e$ term only in the expression for r_2. *Ex ante,* this is strictly negative because, unlike L_x, perfect competition in the labor market implies that w is not set optimally (non-monopsonistic wage) from the lobby's point of view. A small increase in the wage rate due to an infinitesimally small increase in anticipated future tariff produces a *first order* welfare loss to the lobby and thus lowers the government's marginal political benefit from protection. This is simply what we earlier referred to as the "dissipation of rents" away from the intended beneficiaries (the lobby) and towards the unintended beneficiaries (the laborers). The (*ex ante*) optimal tariff takes this cost into account but, *ex post,* this cost is absent due to wage rigidity. This difference in the incentive structure, *ex ante* and *ex post*, creates incentive for surprise protection and the commitment problem follows from this. From the structure of the argument above, it can be easily seen that if the lobby were a *monopsonistic* buyer in the labor market then, for any value of t^e, p^e, it will hire labor to the point such that L_x, w will be jointly optimal with respect to its total surplus. The first order condition for this is simply that $r_2 L = 0$. Consequently, short-run wage rigidity will cease to be a source of commitment problem for the home government. That is, a fully anticipated (infinitesimally) small increase in home's future MFN tariff will increase the equilibrium value of L_x, w but these changes will have *second order* effects only on the lobby's welfare since L_x, w are set optimally to begin with. Home's optimal tariff will then be independent of these changes. *Ex post*, the incentive structure holds due to the short-run labor immobility and the wage rigidity. We summarize our discussion on the *wage distortion* as a source of commitment problem as follows. A fully anticipated increase in home's future tariff leads to greater demand for labor and a higher equilibrium wage rate. When the lobby is *non-monopsonistic* in the factor market then these changes shift some of the benefit of higher protection away from the lobby and towards labor. This has a first order effect on the lobby's surplus and the government's overall welfare which tends to push the optimal tariff downwards. However, *ex post*, this effect is completely absent due to the short-run rigidities mentioned above. Consequently, *ex post,* the government finds it optimal to revise its tariff upwards rendering the optimal tariff time-inconsistent.[30]

[30]We may note here that the *wage distortion* in the sense described above as a source of commitment problem can exist even if the *production distortion* is completely absent. This can be seen by setting $dL_x/dp^e = 0$ throughout the model and noting that r_2 is still strictly negative. For example, if labor is completely immobile across sectors *ex-ante* and *ex-post* then *production distortion* will be completely absent. However, if a higher tariff is anticipated in the future, then domestic firms in sector X will compete for scarce labor *within* the sector. Although the aggregate labor employed in the sector will be unchanged, however, equilibrium wage will rise. That is, *wage distortion* of a fully anticipated increase in future protection will still exist. The argument establishes that commitment problem can arise due to *wage distortion* alone even when *production distortion* is completely absent. This result is in sharp contrast to the literature where short-run rigidity in intersectoral movement of labor, as opposed to its long-run perfect (costless) mobility, is the *only* reason which renders optimal policy time inconsistent. See, for example, Stagier and Tabellini, Maggi and Rodriguez-Claire. An exception to this is Mitra (2002). However, his focus is on fixed (and sunk) cost of lobbying which is completely different from the *wage distortion* as the source of commitment problem in this paper. We discuss this issue more in the conclusion of the paper.

From the arguments stated above we get the following Proposition.

Proposition 1: In the MFN regime with $\beta > 0$, the time-consistent tariff is strictly higher than the full commitment optimal tariff.
Proof: See Appendix B2.

The proof of the previous Proposition follows directly from the fact that, *ex post*, the home government has an incentive to surprise private agents with a tariff which is strictly higher than the pre-announced (optimal) one and, the time-consistent MFN tariff is stable and unique.

We now proceed to derive the optimal tariff in the PTA regime. Since the procedure is similar to the one for the MFN regime we will omit some of the details.

3.2 *Full commitment optimal tariff in the PTA regime*

Home's optimal *external* tariff is given by the following optimization problem:
Maximize $qLf(q) + G(L_y(q)) + t_B E_B(t_A, t_B, f(q)) + Ls(q) + \beta(qLf(q) - w(q)L_x(q))$
by choosing t_B, with $q = q(t_A, t_B), t_A = 0$, $t_B \geq 0$.
This is the same as the one for the MFN regime with the additional restriction that $t_A = 0$.

The first order condition for the optimal tariff is given by:

$$-t_B(L + L_A)\frac{\partial q}{\partial t_B} + E_B + (1+\beta)F\frac{\partial q}{\partial t_B} - C_x\frac{\partial q}{\partial t_B} + L(\beta r_2 - t_B f')\frac{\partial q}{\partial t_B} = 0$$

Let $\hat{t}_{pta} \equiv$ solution value of t_B that is obtained from the previous equation.[31]
We have:

[31] For interior solution conditions for $\hat{t}_{pta}$ and the second order optimization condition, see Appendix A4.

$$\hat{t}_{pta} = \frac{\beta(f + r_2) - E_A/L}{1 + f' + (\lambda_A/\lambda)} + \frac{(\lambda + \lambda_A)/\lambda_B}{1 + (\lambda_A/\lambda)} E_B/L \quad .....(8)$$

All variables on RHS of equation (8) are evaluated with $t_B = \hat{t}_{pta}$, $t_A = 0$, $p = p^e = q(0, \hat{t}_{pta})$.

Interpretation of equation (8) is similar to that of equation (7A) except that we now have the *market power effect* and the *trade deflection effect* of PTAs as defined earlier. The commitment problem in the PTA regime can be noted by comparing equation (8) with equation (5). Note that the two equations differ due to the *production distortion* (f' in equation (8)) and the *wage distortion* (r_2 in equation (8)). The sources of the commitment problem are thus the same as in the MFN regime. It is important to note that, home's optimal *external* tariff, $\hat{t}_{pta}$, depends critically on the *trade deflection effect* and the *market power effect.* The former is captured by the λ_A/λ term in the denominator of RHS of equation (8), while the latter is captured in the difference between the numerator of the coefficient of E_B/L in equation (8) and coefficient of M/L in equation (7A or/and 2B). This was discussed in detail in section 2 and we omit elaborating on it here. We will discuss some important properties of equation (8) when we establish our final results in the next-subsection.

3.3 *PTAs as stumbling blocs*

We introduce the following notations to study the stumbling-bloc feature of PTAs:

$\Omega_j^{FC} \equiv$ overall welfare of the home government when it can commit itself, *ex ante,* to the optimal tariff in regime $j =$ MFN, PTA.

$\Omega_j^{TC} \equiv$ overall welfare of the government in the time-consistent equilibrium in regime $j =$ MFN, PTA.

$\Omega_{FTA} \equiv$ overall welfare of the government in the FTA regime.

Values of all these welfare levels can be easily computed by using the solution stated above.

Lastly, commitment inefficiency, defined as the loss in home government's overall welfare due to its inability to credibly commit itself, *ex ante*, to the future trade policy, is given by:

$\eta_{mfn} \equiv \Omega^{FC}_{MFN} - \Omega^{TC}_{MFN} \geq 0$ in the MFN regime and,
$\eta_{pta} \equiv \Omega^{FC}_{PTA} - \Omega^{TC}_{PTA} \geq 0$ in the PTA regime.

Definition: PTAs are stumbling blocs for wider multilateral trade liberalization if and only if:

$$\Omega^{FC}_{PTA} - \eta_{pta} > \Omega_{FTA} > \Omega^{FC}_{MFN} - \eta_{mfn} \quad ......(9)$$

The second inequality in (9) ensures that the government prefers to liberalize multilaterally if the PTA-option is not available. Note that a necessary condition for this to hold is that $\eta_{mfn} > 0$ since $\Omega^{FC}_{MFN} > \Omega_{FTA}$. That is, absent-commitment problem, home's optimal MFN tariff, $\hat{t}_{mfn}$, is strictly positive which implies the previous inequality. Clearly, while this feature is based on our assumption that foreign tariffs are zero, however, it allows us to focus on multilateral trade liberalization driven by domestic commitment problem only. In the Introduction of the paper we had referred to this as the "potential trade liberalizing effect of domestic commitment problems". The first inequality in (9) ensures that if the PTA-option is available then the government will prefer it over multilateral trade liberalization. The two inequalities together put forward a positive theory of why countries may seek regional agreements and, further, why such agreements may be stumbling blocs for multilateral trade liberalization.[32]

Condition (9) allows us to relate our paper with the stumbling-building blocs literature. That is, the findings in the literature explain how condition (9) holds with $\eta_{mfn} = \eta_{pta} = 0$. The present paper contributes to this literature by highlighting the role of domestic commitment problems (η_{mfn} and η_{pta} terms in (9)) in explaining the stumbling-bloc feature of PTAs.

[32] It will not make any difference to our results if either one or both the inequalities in (9) were replaced by weak inequalities.

The aim of the remaining section is to derive conditions (parameter restrictions) for which condition (9) holds and to relate them to domestic commitment problem, *trade deflection effect* and the *market power effect* of PTAs. To this end we state the following Proposition.

Proposition 2:
(i) PTAs are stumbling blocs only if the home government faces commitment problem vis-a-vis its domestic private agents. That is, only if $\eta_{mfn} > 0$.
(ii) There exists a critical values of λ, b_1, denoted by $\lambda^*, b_1^* \in (0, 1)$, such that $\forall \lambda \leq \lambda^*$ and $\forall b_1 \geq b_1^*$, the home government prefers an *ex ante* commitment to multilateral free trade (FTA regime) over no commitment (MFN regime). That is, the second inequality in (9) holds.
Proof: See Appendix B3.

Part (i) of the Proposition has already been discussed above. The intuition for part (ii) is simple. When λ is sufficiently small ($\lambda \leq \lambda^*$) then the terms-of-trade related benefit to home from an active trade policy in the MFN regime is relatively small and it approaches zero as λ approaches zero. Similarly, when b_1 is sufficiently large ($b_1 \geq b_1^*$) then the marginal productivity of labor schedule in sector X is relatively flat. Consequently, the *wage distortion* of an anticipated increase in future tariff is sufficiently large. Benefit from higher protection is mostly dissipated away from the lobby and towards labor. With the terms-of-trade related benefits and the political benefits being sufficiently small, home's optimal MFN tariff is close to zero. However, the time-consistent tariff is strictly positive (for political reason) since, *ex post*, *wage distortion* is completely absent. The *production distortion* associated with the time-consistent equilibrium is strictly bounded away from zero implying that an *ex ante* commitment to multilateral free trade strictly dominates no commitment at all. This result is similar to the one in Staiger and Tabellini although for different reasons. That is, Staiger and Tabellini show that fully anticipated increase in future tariff has no effect on the distribution of income while the *production distortion* (in the time-consistent equilibrium) is strictly positive. Since they focus on a small country, the terms-of-trade benefits from protection do not arise.

In order to analyze the nature of the commitment problem in the PTA regime we state the following Proposition.

Proposition 3

If home's PTA-partner is sufficiently large relative to home, then the *trade deflection effect* ensures that commitment inefficiency in the PTA regime is arbitrarily small. That is, there exists a critical value of $L_A/L = \lambda_A/\lambda$ such that for all λ_A/λ greater than this critical value, $\eta_{pta} \approx 0$.

Proof: Home's time-consistent *external* tariff, t^*_{pta}, must be *ex post* optimal which implies that it must satisfy equation (5). The optimal *external* tariff, however, is given by equation (8). Comparing the two equations we note that as $\lambda_A/\lambda \rightarrow \infty$ then RHS of the two equations converge to the same common value. In the proof of *Proposition* 2 we have already shown that p^e is bounded below and this lower bound is strictly positive. Further, it can be checked from the *trading equilibrium* solution in section 2.1 that E_A/L and E_B/L are bounded above and below. Consequently, as $\lambda_A/\lambda \rightarrow \infty$, the first term on RHS of equation (8) and the same of equation (5) converge to zero. This implies that the difference between t^*_{pta} and $\hat{t}_{pta}$ converges to zero.[33] Note that all our functions are continuous. This implies that as the difference between t^*_{pta} and $\hat{t}_{pta}$ converges to zero, $\Omega^{FC}_{PTA} - \Omega^{TC}_{PTA} \equiv \eta_{pta}$ converges to zero. By the definition of limits, the result implies that there exists a critical value of λ_A/λ such that the statement in the *Proposition* holds. Note that this result is derived completely from the fact that the *trade deflection effect*, captured by λ_A/λ in the denominator of the first term on RHS of equations (5) and (8), becomes arbitrarily large which drives these two terms to zero. Thus, *trade deflection effect* **alone** is sufficient to ensure that home's commitment inefficiency in the PTA regime is arbitrarily small if its PTA-partner is sufficiently large relative to itself. *Market power effect* is completely irrelevant here.

Q.E.D.

Comparing the previous two Propositions we note that while *Proposition* 2 highlights conditions under which home would prefer a multilateral free trade agreement over no commitment, *Proposition* 3 states that if home were to instead form a PTA then, under the conditions specified, its commitment problem will be virtually eliminated. Note that the conditions identi-

[33] The limiting values of t^*_{pta} and $\hat{t}_{pta}$ are well defined under appropriate (interior soulution) conditions. See the proof of *Proposition* 4,5 for more details.

fied in the two Propositions are consistent with each other. That is, both these Propositions hold when b_1 is sufficiently large and λ_A/λ is sufficiently large.[34] The next Proposition ties our arguments together to establish conditions under which domestic commitment problems imply that PTAs will be stumbling blocs.

Proposition 4
Domestic commitment problems imply that PTAs will be stumbling blocs when b_1 is sufficiently large and home can form a PTA with a country which is sufficiently large relative to the rest of the world (high value of λ_A). The *trade deflection effect* and the *market power effect* work in tandem with one another in producing this result.
Proof: The Proposition does not impose any restrictions on the value of L/L_B which is home's size relative to that of country B (non-member country). Hence we treat this as a parameter in the proof that follows. The proof utilizes limit-analysis. As home's PTA-partner (country A) becomes arbitrarily large relative to the rest of the world, it follows by definition that $\lambda_A \to 1$, $\lambda \to 0$ and $\lambda_B \to 0$. Thus, $\lambda_A/\lambda \to \infty$. Let $b_1 \geq b_1^*$ which is defined in *Proposition* 2. These features ensure that the conditions identified in Propositions 2 and 3 are satisfied. Consequently, commitment inefficiency in the MFN regime is sufficiently large so that home would prefer the FTA regime over the MFN regime (*Proposition* 2). That is, the second inequality in (9) holds. All we need to show is that the first inequality in (9) holds too at the same time and relate this feature to the *trade deflection effect* and the *market power effect*. To this end note that since conditions in *Proposition* 3 are satisfied here, it follows that as $\lambda_A \to 1$, $\eta_{pta} \to 0$, and $t^*_{pta} \to \hat{t}_{pta}$. Now consider equation (5) which must be satisfied by t^*_{pta}.As $\lambda_A \to 1$ so that $\lambda_A/\lambda \to \infty$, it follows that the first term on RHS of equation (5) converges to zero due to the *trade deflection effect* as explained in *Proposition* 3. The second term on RHS of the same equation equals $\dfrac{(\lambda+\lambda_A)/\lambda_B}{1+(\lambda_A/\lambda)}E_B/L$. Note that λ_A/λ in the denominator of this term is the *trade deflection effect* as discussed earlier. This *effect* drives the value of this expression (and hence t^*_{pta}) to zero as $\lambda_A \to 1$ (or $\lambda_A/\lambda \to \infty$). This tendency is countered by the fact that the numerator term, $(\lambda+\lambda_B)/\lambda_B = [1+(\lambda_A/\lambda)]\lambda/\lambda_B$ tends to

[34] From the definition of λ, λ_A it follows that as λ_A/λ becomes arbitrarily large then λ becomes arbitrarily small.

infinity as $\lambda_A \to 1$. This is simply due to the *market power effect* of PTAs. PTA with country A increases home's *market power* in trade against country B which increases its optimal *external* tariff for terms-of-trade related benefits against country B. The net effect of these two opposite forces is that the term equals: $\frac{(\lambda+\lambda_A)/\lambda_B}{1+(\lambda_A/\lambda)}E_B/L = \frac{\lambda}{\lambda_B}E_B/L$. Substituting for E_B from the *trading equilibrium* solution in section 2.1, we get that as $\lambda_A \to 1$, $\frac{\lambda}{\lambda_B}E_B/L \to \frac{x_B - x_A}{2} > 0$. Putting these arguments together we get that the *market power effect* and the *trade deflection effect* working in tandem ensure that: *limit* $t^*_{pta} =$ *limit* $\hat{t}_B = \frac{x_B - x_A}{2} > 0$. The previous inequality follows form our interior solution condition that home's import-volume from country B is strictly positive (in the limit) in the PTA regime. Note that $(x_B - x_A)/2$ is home's "revenue-maximizing tariff" in the PTA regime. Home's tariff revenue per home-agent in the limit equals $[(x_B - x_A)^2/4]L_B/L > 0$. The rest of the proof is straightforward. Since t^*_{pta} and $\hat{t}_B$ converge to the same value which is strictly positive, it follows that in the limit $\Omega^{FC}_{PTA} = \Omega^{TC}_{PTA} > \Omega_{FTA}$ which ensures that the first inequality in (9) holds. From the definition of limits, it follows that there exists a sufficiently high critical value of λ_A such that the statement in the Proposition holds for all values of λ_A greater than this critical value (given that $b_1 \geq b_1^*$).
Q.E.D.

We summarize the findings in the previous three Propositions in the following way. Domestic commitment problems reduce the benefit from an active trade policy. The paper highlights the role of production and wage distortions in this context. When commitment related inefficiencies are sufficiently severe then the government prefers to bind its trade policy through an *ex ante* commitment to a multilateral free trade agreement (*Proposition* 2). This is the trade liberalizing effect of domestic commitment problems. The result is similar to the one in Maggi and Rodriguez-Clare although they do not focus on trade liberalization. With this holding, we seek to explore if the PTA-option will undermine this trade liberalizing effect or enhance it from the multilateral point of view. Propositions 3 and 4 address this issue. Specifically, if the home country forms a PTA with a sufficiently large country (relative to itself) then the domestic commitment problem is virtually eliminated (*Proposition* 3). The *trade deflection effect* increases home's cost of any *ex post* revision in its *external* tariff so that incentive for surprise protection is eliminated. The time-consistent tariff approaches the full commitment optimal tariff. The stumbling bloc feature of PTAs is established by next showing that home's optimal *external* tariff is strictly bounded away from zero. *Proposition* 4 establishes this. Briefly, the *trade deflection effect* enhances home's cost of protection against country B which, by itself, drives its optimal *external* tariff to zero when its PTA-partner is sufficiently large relative to itself. While this effect alone renders the stumbling bloc feature of PTAs insignificant, however, the *market power effect* counters this by pushing the optimal *external* tariff upwards. The net effect of these two forces is that home's optimal *external* tariff is strictly bounded away from zero and, depending on the parameter values, can be significant. Thus, the *trade deflection effect* and the *market power effect* work in tandem with one another rendering PTAs stumbling blocs in the presence of domestic commitment problems.

The requirement in *Proposition* 4 that home's PTA-partner is sufficiently large can be immediately diluted in the following way. Suppose that the world consists of a large number of countries with each one being relatively small individually. Despite the fact that there is no single large country, the basic results of this paper will continue to hold since home can now form a PTA with a relatively large proportion of foreign countries. The *market power effect* and the *trade deflection effect* identified in the paper will continue to have the same effect in this alternative world-structure. Thus, the value of λ_A in our model can be alternatively interpreted as the size of home's trading bloc.

Before closing section 3 we would like to mention that, in addition to showing how PTAs could be stumbling blocs due to commitment related problems, the paper serves to highlight a broader theme in that it shows that the benefit from international trading rules (e.g., Article XXIV of WTO) may depend critically on the nature of domestic institutions within individual countries. Analyzing the link between the structure of domestic and international policy-making institutions is likely to enrich our understanding of policy related issues.

Conclusion

Domestic commitment problems create incentive for a country to seek multilateral free trade agreements to bind its trade policy. Such agreements will be preferred over an active trade policy when the latter involves large commitment inefficiencies relative to the terms-of-trade related and political gains in the time-consistent equilibrium. However, if a bilateral agreement is formed then commitment inefficiencies are virtually eliminated and, at the same time, the agreement allows the home country to exploit the terms-of-trade related gains against the non-member countries. Consequently, incentive for further (multilateral) trade liberalization is eliminated. The paper identifies conditions when this will happen. We put forward this as another reason why regional agreements are stumbling blocs for wider multilateral trade liberalization. In the conclusion we consider some generalizations of our model and suggest further extensions.

Firstly, to keep the model simple and focus on the domestic commitment problem, we assumed that the tariffs of the two foreign countries were (exogenously) fixed at zero. Clearly, richer models can be explored here by considering the Nash equilibrium in tariff games. We believe that the basic results of our model will still hold. That is, the home country can again form a PTA with a sufficiently large number of countries so that the *market power effect* and the *trade deflection effect* will again be the dominant effects. We may note here that these effects are quite independent of whether a non-member country retaliates or adopts a free trade policy against the home country. However, a related issue here is that the strategic interaction across countries and between the domestic country and its private agents, are likely to be interdependent and exploring these can provide a better understanding of the relationship between domestic and international institutions. Further, the magnitude of the domestic commitment problem in the home country, when the home country (or the trading block to which it belongs) is not too large, depends crucially on the trade policies of the foreign countries. Consequently, externalities can travel across countries through each country's local prices. This can provide another rationale for coordinating trade policies across countries. Prescribing rules for international institutions like the WTO to internalize this externality would be an important issue here.

Secondly, the model highlights two different sources of time inconsistency problem. The *non-monopsonistic* nature of the lobby was one of them. Clearly, this source can be further explored. For instance, it is easy to see that if the lobby had the option to collude (in the factor market) then its optimal collusion will never be perfectly monopsonistic. An infinitesimally small move away from perfect collusion will produce only second order loss to the lobby (through changes in the wage rate and the amount of labor hired), but it will produce first order gains (for the lobby) through its impact on *ex post* optimal tariff. This opens up the possibility of strategic interaction between the lobby and the government and the impact of this on trade policy and welfare can be significant.

Thirdly, we focused on the trade deflection and the market power effects of PTAs. Clearly, other aspects of regional agreement can be explored too. For example, we may explore whether a country seeking to enhance its credibility would choose trade creating or diverting PTAs. These two are likely to have different commitment-enhancing benefits for the home country as they offer different incentive to for *ex post* surprise protection. The issue can be explored if full detail. Another interesting effect which can reinforce the findings of the paper is the "tariff-complementarity effect". That is, once a bilateral agreement is formed the home country will find it optimal to lower its tariff on its imports from the non-member country for reasons other than the trade deflection cost. As Bagwell and Staiger (1999b) note, free trade between the home country and its PTA-partner implies that, with home's external tariff held fixed momentarily, home's import-volume from the non-member country will fall which reduces home's (terms-of-trade related) marginal benefit from protection. This implies that, *ex post*, home will have a smaller incentive to revise its tariff upwards. Whether this effect in itself can reduce the gap between the home's *ex ante* and *ex post* optimal policy in the sense obtained in this paper through the *trade deflection effect* will depend on the underlying economic conditions and can be further explored. In fact, at a broad level there will no harm in interpreting our *trade deflection effect* as form of tariff complementarity effect since the two play a similar role in the context of this paper.

Fourthly, extending the model to non-linear demand functions will be an important and interesting avenue to further explore the nature and magnitude of the effects discussed above.

Appendix

Appendix A1: Existence and uniqueness of the time-consistent equilibrium prediction about home's future equilibrium local (relative) price in the MFN regime:
Proof: This is given by the value of p^e that satisfies the following equation:

$$p^e = p(t(f(p^e)), t(f(p^e)), f(p^e))$$

Using the $p(.)$ function given by equation (1) we can rewrite the previous equation as:

$$p^e = \alpha - (\lambda_A x_A + \lambda_B x_B) + (1-\lambda)t(f(p^e)) - \lambda f(p^e) \quad ......(A1.1)$$

From the factor market equilibrium outlined in section 2 we can see that f is strictly bounded above by $(R/L)^{1-b_1}$ and bounded below by zero. Under the interior solution condition that α is sufficiently high so that world demand for good X is higher than its supply when p is arbitrarily close to zero, it follows that RHS of equation (A1.1) is strictly positive $\forall p^e > 0$. Consequently, we can find sufficiently small (close to zero) value of p^e such that LHS of equation (A1.1) is strictly smaller than its RHS. Similarly, since f is bounded above, it follows that we can find a sufficiently high value of p^e such that LHS of equation (A1.1) is strictly higher than its RHS. Existence of the solution value of p^e follows from this and the fact that LHS and RHS of equation (A1.1) are continuous in p^e.

Uniqueness: Substituting for $t(f(p^e))$ from equation (3) in RHS of equation (A1.1) we can see that there are two possible cases which are as follows: (i) when RHS is non-increasing in p^e; that is, $(1-\lambda)\beta - 2\lambda \leq 0$, and (ii) when RHS is strictly increasing in p^e; that is, $(1-\lambda)\beta - 2\lambda > 0$. In case (i) uniqueness of the solution is assured since in equation (A1.1) LHS is strictly increasing while RHS is non-increasing in p^e. For case (ii) note that the second derivative of RHS with respect to p^e is strictly negative due to declining marginal productivity of labor in sector X. Given this concavity of RHS, the uniqueness of the solution follows directly from the fact that there exists an interval in the neighborhood of zero such that for any value of p^e

in this interval, LHS of equation (A1.1) is strictly smaller than RHS of the same equation, and that the latter is strictly positive (due to the interior solution condition stated in the proof above).
Q.E.D.

Appendix A2: Existence, uniqueness and stability of the time-consistent tariff (t^*_{mfn}) in the MFN regime.
Proof: The time-consistent equilibrium tariff, t^*_{mfn}, is equal to the value of t that satisfies the following condition:

$$t = t(f(p^e(t,t))) \quad ......\text{(A2.1)}$$

Consider RHS of the previous equation. This is uniquely determined for any given value of t as proved in Appendix A3 below. Further, $f(p^e(t^e,t^e))$ is bounded above and below as discussed in Appendix A1 above. From equation (3) in section 2 it then follows that RHS of equation (A2.1) is bounded above and below. The lower bound is strictly positive under the (interior solution) condition that $f < \lambda f + \lambda_A x_A + \lambda_B x_B$ which follows from the interior solution conditions specified in Appendix A5, part (i). The previous inequality ensures that home is the natural importer of good X. That is, under complete free trade home imports good X from each foreign country. With a strictly positive lower bound on the value of RHS of equation (A2.1) and the fact that RHS of the same equation is well defined and unique, it follows that there exists a sufficiently small value of t such that RHS value is strictly greater than the LHS. Similarly, since RHS of the equation is bounded above it follows that there exists a sufficiently large value of t such that RHS value is smaller than LHS value. Existence of the time-consistent tariff follows from this and the fact that RHS and LHS of the equation (A2.1) are continuous in t.
Uniqueness: We have two possible cases: when $\beta - \lambda \leq 0$ and when $\beta - \lambda > 0$. In the former case uniqueness of the time-consistent equilibrium follows trivially since RHS of equation (A2.1) is non-increasing in t while LHS of the same equation is strictly increasing in t. Now consider the latter case when $\beta - \lambda > 0$. Uniqueness is ensured here by the fact that RHS is strictly concave in t, strictly monotonically increasing in t and, as discussed above, for arbitrarily small (close to zero and positive) values of t RHS is strictly positive and strictly greater than LHS.

Stability: When $\beta - \lambda > 0$ then the stability of the (unique) time-consistent tariff follows directly from the fact that RHS of equation (A2.1) is strictly concave in t. When $\beta - \lambda \leq 0$ it can be seen that stability is ensured by the fact that the absolute value of the first derivative of RHS is strictly less than one. That is, differentiating RHS of equation (A2.1) with respect to t we get that the absolute value of this derivative is equal to $\frac{(1-\lambda)(\lambda-\beta)f'}{(1+\lambda f')(1+\lambda)} < 1$. This ensures the stability of the time-consistent MFN tariff and hence the equilibrium values of all the endogenous variables in the MFN regime.
Q.E.D.

Appendix A3: Existence and uniqueness of $p^e(t^e_A, t^e_B)$ function.
As stated in section 2, p^e is given by the solution to the following equation:

$$p^e + \lambda f(p^e) = \alpha - \lambda_A x_A - \lambda_B x_B + \lambda_A t^e_A + \lambda_B t^e_B$$

RHS of the previous equation is well defined and its value is unique for any arbitrary given values of t^e_A, t^e_B. The value of LHS of the previous equation is strictly and monotonically increasing in p^e. This implies that the solution, if it exists, must be unique. Under the interior solution condition that α is sufficiently high so that world demand for good X is higher than the world supply of good X when local prices in all the three countries are zero, it follows that the LHS is strictly less than the RHS for arbitrarily small values of p^e. Note that from the factor market equilibrium conditions stated in section 2, $f(p^e)$ approaches zero as p^e approaches zero. Next note that RHS is bounded above for any given values of t^e_A, t^e_B. Thus, we can find a value of p^e such that LHS of the previous equation is strictly greater than the LHS of the same. These properties, together with the fact that the LHS is continuous in p^e, ensure the existence of the solution value of $p^e(t^e_A, t^e_B)$, for all t^e_A, t^e_B. Stability of the solution value is ensured by the fact that for arbitrarily small value of p^e, LHS of the previous equation is smaller than RHS of the same equation, both are positive and that the LHS is strictly increasing and strictly concave in p^e.
Q.E.D.

Appendix A4: Interior solution conditions for $\hat{t}_{mfn}$ and $\hat{t}_{pta}$

(i) $\hat{t}_{mfn}$: Consider equation (7). Note that $f + r_2 > 0$ under the assumption that $b_1 < 1$. Set $t = 0$ in LHS of equation (7). It is straightforward to check that with $f + r_2 > 0$, LHS of equation (7) is strictly positive at $t = 0$. This implies that $\hat{t}_{mfn}$ must be strictly positive. Note that when $b_1 = 1$ then $f + r_2 = 0$ and LHS of equation (7) is still strictly positive at $t = 0$ since $\lambda > 0$.

Now consider a prohibitive MFN tariff when production and wage rate is endogenous to the tariff announcement. By definition, at the prohibitive tariff, $M = 0$. Imposing this restriction it is easy to see that LHS of equation (7) is strictly negative provided that β is not too large. We assume that this holds throughout the model and is formally stated in Appendix (A5), part (ii) below. With this holding, $\hat{t}_{mfn}$ is strictly interior and it must satisfy equation (7). Under these restrictions, the second order maximization condition is guaranteed to hold (locally) at a value of t that satisfies equation (7).

(ii) $\hat{t}_{pta}$: Consider the first order condition for $\hat{t}_{pta}$ specified in section 3.2 and reproduced below:

$$t_B(L + L_A)\frac{\partial q}{\partial t_B} + E_B + (1 + \beta)F\frac{\partial q}{\partial t_B} - C_x\frac{\partial q}{\partial t_B} + L(\beta r_2 - t_B f')\frac{\partial q}{\partial t_B} = 0$$

At $t_B = 0$, LHS of the previous equation is strictly positive since $F + Lr_2 > 0$ as stated in part (i) above and the remaining terms are strictly positive under the interior solution conditions stated in Appendix (A5), part (i) below. This implies that $\hat{t}_{pta}$ must be strictly positive. Further, it is evident from the first order condition above that if β is not too large then $\hat{t}_{pta}$ will be non-prohibitive since at a prohibitive external tariff, LHS of the first order condition specified above is strictly negative. The restriction on the value of β is consistent with the interior solution condition specified in Appendix (A5), part (ii) below. With this holding, the second order maximization condition is guaranteed to hold at a value of t_B that satisfies the first order condition.

Q.E.D.

Appendix A5: Interior solution conditions

All our results are based on the assumption that the home country always imports a strictly positive amount from each foreign country. Although our basic results can be derived under weaker conditions too, however, the assumption simplifies the exposition of our findings. The following conditions

are sufficient for strictly positive import-volumes in the sense stated above for any given finite values of the parameters in the model:
(i) $x_B > x_A > \lambda f^* + \lambda_A x_A + \lambda_B x_B$ where $f^* \equiv F(L)/L$. This condition ensures that home's import-volume from each foreign country is strictly positive under complete free trade. Further, it ensures that $t(f), t_B(f)$ are strictly positive $\forall f$.
(ii) We assume that β is not too large so that home's politically optimal tariff is non-prohibitive. It can be easily checked that for sufficiently small values of β, $t(f), t_B(f)$ are non-prohibitive $\forall f$; the same holds for $\hat{t}_{mfn}, \hat{t}_{pta}, t^*_{mfn}$ and t^*_{pta}.

Appendix B1: Existence, uniqueness and stability of the time-consistent tariff in the PTA regime.
Proof: The proof of this follows the same logic as for the MFN regime in Appendix A2. The only difference is that the interior solution conditions here, while consistent with the ones for the MFN regime, are slightly more strict. That is, we need to invoke the assumption made in section 3 that $x_B > x_A$. This assumption ensures that home's import-volume from country B (the non-member country) is strictly positive in the PTA regime when its external tariff is arbitrarily small. This restriction along with the condition that $f < \lambda f + \lambda_A x_A + \lambda_B x_B$ (discussed in Appendix A2) ensure that home's *ex post* optimal *external* tariff in the PTA regime (given by equation (6) in section 2) is strictly positive; $\Delta_1 > 0$ for all finite values of the parameters in the model. The rest of the proof can now be seen by following exactly the same steps as outlined for the proof in Appendix A2 and using the time-consistency condition that $t^e_B = t_B(f(p^e(0, t^e_B)))$.
Q.E.D.

Appendix B2: In the MFN regime, the time-consistent tariff is strictly higher than the full commitment optimal tariff.
Proof: This follows directly from the fact that the home government has an incentive to surprise private agents with a strictly higher tariff *ex post* since $r_2 < 0$, $f' > 0$ and that the time-consistent MFN tariff is stable and unique. Notice that $\frac{dr}{dp^e} \equiv r_2 L = (qF - w)\frac{dL_x}{dp^e} - L_x \frac{dw}{dp^e}$, $w = p^e F'$ and $p^e = q$ when future tariffs are fully anticipated. From this it follows that the first term in

the equation for dr/dp^e is equal to zero and the second term is strictly non-zero implying that $r_2 < 0$. For more details on the stability and uniqueness of the time-consistent tariff in the MFN regime, see Appendix A2.
Q.E.D.

Appendix B3: *Proposition* 2:
(i) A necessary condition for PTAs to be stumbling blocs is that the home government faces commitment problem vis-a-vis its domestic private agents. That is, $\eta_{mfn} > 0$.
(ii) There exists a critical values of λ, b_1 denoted by $\lambda^*, b_1^* \in (0,1)$ such that $\forall \lambda \leq \lambda^*$ and $b_1 \geq b_1^*$, the home government prefers an *ex ante* commitment to multilateral free trade (FTA regime) over no commitment (MFN regime). That is, the second inequality in (9) holds.
Proof: (i) Note that in equation (7A), $f + r_2 > 0$ given that $b_1 < 1$. For $\beta, \lambda > 0$ we have that $\hat{t}_{mfn} > 0$. This implies that $\Omega_{MFN}^{FC} > \Omega_{FTA}$. For PTAs to be stumbling blocs the second inequality in (9) must hold. This will hold only if $\eta_{mfn} > 0$, since $\Omega_{MFN}^{FC} > \Omega_{FTA}$.
(ii) We establish the proof of this using limit-analysis.
Limit-analysis: We consider the limit-values of the endogenous variables when $b_1 \to 1$ and $\lambda \to 0$. For the latter we will assume that L, K, R are fixed and parametrically vary the value of $L_A + L_B$. Thus, as $L_A + L_B \to \infty$ we get $\lambda \to 0$. Although not necessary for the proof here, it will be useful to treat L_B as fixed and vary L_A only. This will allow us to see easily that the conditions required for *Proposition* 2 and *Proposition* 3 to hold are met simultaneously. The proof involves the following steps:
(a) From the equation for p^e in section 2.2, we note that p^e is bounded below. We assume that α is sufficiently high so that the solution values of p^e and other endogenous variables in the limit are strictly interior. This will be pointed out as we proceed. (b) From the factor market equilibrium conditions in section 2 we have that: $\frac{b_1 R^{1-b_1}}{b_2 K^{1-b_2}} p^e = \frac{L_x^{1-b_1}}{(L - L_x)^{1-b_2}}$ when $0 < L_x < L$ and $L_x(p^e) = 0$ if and only if $\frac{b_1 R^{1-b_1}}{b_2 K^{1-b_2}} p^e < \frac{L_x^{1-b_1}}{(L - L_x)^{1-b_2}}$. Now assume that α is sufficiently high so that the lower bound for the equilibrium value of p^e is strictly greater than $b_2(K/L)^{1-b_2}$. Now take limits on both sides of the equation: $\frac{b_1 R^{1-b_1}}{b_2 K^{1-b_2}} p^e = \frac{L_x^{1-b_1}}{(L - L_x)^{1-b_2}}$. Specifically, we get: *limit*

$(L - L_x)^{1-b_2} = b_2 K^{1-b_2} (limit\ p^e)^{-1}$. It is straightforward to see that the limiting value of L_x must be strictly positive under the restriction that α is sufficiently high as stated above. This implies that $f(p^e)$ converges to a strictly positive value. (c) Now consider the expression for r_2 stated in section 3. Impose the (time-consistent and full commitment) equilibrium conditions that $w = p^e F'$ and $p = p^e = q(.)$. Taking limits on both sides of the expression for $r_2 L$ we get that $limit\ r_2 = -limit\ f(p^e)$.
The rest of the proof follows simply from these limit-values. Taking limits on both sides of equation (7A) we can see that $\hat{t}_{mfn} \to 0$. From part (c) above, the first term on RHS of equation (7A) converges to zero and the second term (pure welfare maximizing tariff) converges to zero also, as λ approaches zero. Since all our functions are continuous, it follows that as $\hat{t}_{mfn} \to 0$, $\Omega^{FC}_{MFN} \to \Omega_{FTA}$. Since the time-consistent MFN tariff must be *ex post* optimal, it follows that it must satisfy equation (3). Taking limits on both sides of equation (3) we get that $t^*_{mfn} \to \beta\ limit\ f(p^e) > 0$, where the inequality follows from the results in part (b) of the proof. The difference between the limiting values of $\hat{t}_{mfn}$ and t^*_{mfn} implies that *limit* $(\Omega^{FC}_{MFN} - \Omega^{TC}_{MFN}) = limit\ \eta_{mfn} > 0$. With our results that $\Omega^{FC}_{MFN} \to \Omega_{FTA}$ and $limit\ \eta_{mfn} > 0$, the second inequality in (9) is guaranteed to hold in the limit. From the definition of limits, it follows that $\exists\ \lambda^*, b_1^*$ such that part (ii) of the Proposition is satisfied.
Q.E.D.

References

[1] Bagwell, K. and Staiger, R.W. "A Theory of Managed Trade", *American Economic Review*, September 1990, 80(4), pp. 779-795.

[2] Bagwell, K. and Staiger, R.W. "Multilateral Tariff Cooperation During the Formation of Customs Unions", *Journal of International Economics*, February 1997a, 42, pp. 91-123.

[3] Bagwell, K. and Staiger, R.W. "Multilateral Tariff Cooperation During the Formation of Free Trade Areas", *International Economic Review*, May 1997b, 38(2), pp. 291-319.

[4] Bagwell, K. and Staiger, R.W. "An Economic Theory of GATT", *American Economic Review*, March 1999a, 89(1), pp. 215-248.

[5] Bagwell, K and Staiger, R.W. "Regionalism and Multilateral Tariff Cooperation", in John Piggot and Alan Woodland, eds., **International Trade Policy and the Pacific Rim**, 1999b, Macmillan: London, pp. 157-185.

[6] Bagwell, K and Staiger, R.W. "Multilateral Trade Negotiations, Bilateral Opportunism and the Rules of GATT", *National Bureau of Economic Research*, Working Paper No. 7071, April 1999c. Forthcoming, *Journal of International Economics*.

[7] Bagwell, K and Staiger, R.W. "Reciprocity, Non-discrimination and Preferential Agreements in the Multilateral Trading System", *European Journal of Political Economy*, 2001, 17, pp. 281-325.

[8] Bagwell, K and Staiger, R.W. **The Economics of World Trading System**, 2002, Cambridge: MIT Press.

[9] Bagwell, K and Staiger, R.W. "The Case of Auctioning Countermeasures in the WTO", *Working Paper*, March 2003.

[10] Baldwin, R. "Politically Realistic Objective Functions and Trade Policy", *Economic Letters*, 1987, 24, pp. 287-90.

[11] Bhagwati, J. **The World Trading System at Risk**, Princeton University Press, Princeton, New Jersey, 1991.

[12] Bhagwati, J. "Regionalism and Multilateralism: An Overview", in A. Panagariya and J. De Melo eds., **New Dimensions in Regional Integration**, 1993, Washington, D.C., World Bank.

[13] Bond, Eric W., and Park, Jee-Hyeong. "Gradualism in Trade Agreements with Asymmetric Countries", *Review of Economic Studies*, April 2002, 69(2), pp. 379-406.

[14] Bond, Eric W., Syropoulos, C., and Winters, Alan L. "Deepening of Regional Integration and Multilateral Trade Agreements", *Journal of International Economics*, 2001, 53, pp. 335-361.

[15] Brainard, S.L. "Last One Out Wins: Trade Policy in an International Exit Game", *International Economic Review*, 1994, 35, pp. 151-172.

[16] Brainard, S.L. and Verdier, T. "The Political Economy of Declining Industries: Senescent Industry Collapse Revisited", *Journal of International Economics*, 1997, 42, pp. 221-237.

[17] Caplin, A. and Krishna, K. "Tariffs and the Most-Favored-Nation Clause: A Game Theoretic Approach", *Seoul Journal of Economics*, 1988, 1, pp. 267-289.

[18] Dixit, A. and Norman, V. **Theory of International Trade: A Dual General Equilibrium Approach**, Cambridge University Press, 1980.

[19] Findlay, Ronald. and Wellisz, Stainslaw. "Endogenous Tariffs, the Political Economy of Trade Restrictions, and Welfare", in J. N. Bhagwati ed., **Import Competition and Response**, 1982, Chicago, University of Chicago, pp. 223-234.

[20] Goldberg, Pinelopi K., and Maggi, Giovanni. "Protection for Sale: An Empirical Investigation", *American Economic Review*, December 1999, 89(5), pp. 1135-55.

[21] Grossman, Gene M. and Helpman, Elhanan. "Protection for Sale", *American Economic Review*, September 1994, 84(4), pp. 833-850.

[22] Grossman, Gene M. and Helpman, Elhanan. "Trade Wars and Trade Talks", *Journal of Political Economy*, 1995, 103(4), pp. 675-708.

[23] Gould, D.M. "Free Trade Agreements and the Credibility of Trade Reforms", *Economic Review- Federal Reserve Bank of Dallas*, 1992, First Quarter, pp. 17-27.

[24] Hull, Cordell. **The Memoirs of Cordell Hull**, 1948, Hodder and Stoughton: London.

[25] Keenan, J. and Riezman, R. "Do Big Countries Win Tariff Wars?", *International Economic Review*, February 1988, 29, pp. 81-85.

[26] Krishna, Pravin. "Regionalism and Multilateralism: A Political Economy Approach", *Quarterly Journal of Economics*, February 1998, pp. 227-251.

[27] Kydland, Finn E. and Prescott, Edward. "Rules Rather than Discretion: The Inconsistency of Optimal Plans", *Journal of Political Economy*, June 1977, 85(3), pp. 473-491.

[28] Lapan, Harvey E. "The Optimal Tariff, Production Lags, and Time Consistency", *American Economic Review*, June 1988, 78(3), pp. 395-401.
[29] Levy, Philip I. "A Political-Economic Analysis of Free-Trade Agreements", *American Economic Review*, September 1997, 87(4), pp. 506-519.
[30] Ludema, R.D. "On the Value of Preferential Trade Agreements in Multilateral Negotiations", *Mimeo*, University of Western Ontario, 1993.
[31] Maggi, G and Rodriguez-Clare, A. "The Value of Trade Agreements in the Presence of Political Pressures", *Journal of Political Economy*, June 1998, 106(3), pp. 574-601.
[32] Matsuyama, K. "Perfect Equilibria in a Trade Liberalization Game", *American Economic Review*, June 1990, 80(3), pp. 480-492.
[33] Mayer, Wolfgang. "Theoretical Considerations on Negotiated Tariff Adjustment", *Oxford Economic Papers*, March 1981, 33(1), pp. 135-53.
[34] Mayer, Wolfgang. "Optimal Pursuit of Safeguard Actions Over Time", in A.V. Deardorff and R.M. Stern eds., **Analytical and Negotiating Issues in the Global Trading System**, 1994, University of Michigan Press, Ann Arbor.
[35] McCalman, P. "Multilateral Trade Negotiations and the Most-Favored-Nation Clause", Ph.D Dissertation, 1998, The University of Wisconsin.
[36] McLaren, John. "Size, Sunk Costs, and Judge Bowker's Objection to Free Trade", *American Economic Review*, June 1997, 87(3), pp. 400-420.
[37] Mitra, D. "Endogenous Political Organizations and the Value of Trade Agreements", *Journal of International Economics*, August 2002, 57(2), pp. 473-485.
[38] Richardson, Martin. "Endogenous Protection and Trade Diversion", *Journal of International Economics*, 1993, 34, pp. 309-324.
[39] Richardson, Martin. "Tariff Revenue Competition in a Free Trade Area", *European Economic Review*, 1995, 39, pp. 1429-1437.
[40] Staiger, R.W and Tabellini, G. "Discretionary Trade Policy and Excessive Protection", *American Economic Review*, 1987, 77(5), pp. 823-837.
[41] Staiger, R.W. "International Rules and Institutions for Trade Policy," in Gene M. Grossman and Kenneth Rogoff, eds., **Handbook of International Economics**, Vol. 3, Amsterdam: North Holland, 1995, pp. 1495-1551.
[42] Staiger, R.W and Tabellini, G. "Do GATT Rules Help Governments Make Domestic Commitment", *Economics and Politics*, July 1999, 11(2), pp. 109-44.
[43] Syropoulos, Constantinos. "Optimum Tariffs and Retaliation Revisited:

How Country Size Matters", *Mimeo*, Florida International University, September, 2001.