A Political-Economy Theory of Trade Agreements*

Giovanni Maggi
Princeton University and NBER

Andrés Rodríguez-Clare
Pennsylvania State University and NBER

March 2007

Abstract

This paper presents a theory of trade agreements where "politics" play an central role. This stands in contrast with the standard theory, where even politically-motivated governments sign trade agreements only to deal with terms-of-trade externalities. We develop a model where governments may be motivated to sign a trade agreement both by the presence of standard terms-of-trade externalities and by the desire to commit vis-à-vis domestic industrial lobbies. The model is rich in implications. In particular, it predicts that trade agreements result in deeper trade liberalization when governments are more politically motivated (provided capital mobility is sufficiently high) and when capital can move more freely across sectors. Also, governments tend to prefer a commitment in the form of tariff ceilings rather than exact tariff levels. In a fully dynamic specification of the model, trade liberalization occurs in two stages: an immediate slashing of tariffs and a subsequent gradual reduction of tariffs. The immediate tariff cut is a reflection of the terms-of-trade motive for the agreement, while the domestic-commitment motive is reflected in the gradual phase of trade liberalization. Finally, the speed of trade liberalization is higher when capital is more mobile across sectors.

Keywords: trade agreements, domestic commitment, terms of trade, lobbying for protection, capital mobility.
JEL classification: F13, D72

*We thank Kyle Bagwell, Elhanan Helpman, Robert Staiger, Marcelo Olarreaga and seminar participants at UC Berkeley, ECARES, Ente Luigi Einaudi di Roma, Federal Reserve Bank of New York, FGV Rio de Janeiro, Harvard University, IADB, Penn State University, Southern Methodist University, Syracuse University, University of Texas at Austin, University of Virginia, University of Wisconsin, Vanderbilt University and the 2005 NBER Summer Institute for very useful comments. We also thank three anonymous referees for very useful comments and suggestions. Richard Chiburis provided outstanding research assistance. Giovanni Maggi acknowledges financial support from the National Science Foundation (SES-0351586).

1 Introduction

The history of trade liberalization after World War II is intimately related with the creation and expansion of the GATT (now WTO), and with the signing of countless bilateral and regional trade agreements. Clearly, there are strong forces pushing countries to sign international trade agreements, and it is important for economists and political scientists to understand what these forces are. Why do countries engage in trade agreements? What determines the extent and form of liberalization that takes place in such agreements?

The standard theory of trade agreements dates back to Johnson (1954), who argued that, in the absence of trade agreements, countries would attempt to exploit their international market power by taxing trade, and the resulting equilibrium – a trade war – would be inefficient for all countries involved. International trade agreements can be seen as a way to prevent such a trade war. This idea was later formalized in modern game-theoretic terms by Mayer (1981).

Grossman and Helpman (1995a) and Bagwell and Staiger (1999) have extended this framework to settings where governments are subject to political pressures. In these models, as Bagwell and Staiger emphasize, even politically-motivated governments engage in trade agreements *only* to correct for terms of trade externalities. Thus, "politics" does not affect the motivation to engage in trade agreements.

In this paper we present a theory where politics is very much at the center of trade agreements. In particular, we consider a model where trade agreements help governments to deal with a time-inconsistency problem in their interaction with domestic lobbies. Maggi and Rodríguez-Clare (1998) showed how such a time-inconsistency problem may emerge in a small-open economy when capital is fixed in the short run but mobile in the long run. The present paper builds on this idea to develop a fuller theory of trade agreements.[1]

We start by reviewing the logic behind the domestic-commitment problem that is at the basis of our theory. This logic is easily illustrated for the case of a small economy. According to the modern political-economy theory of trade policy, it is not clear why a small-country government would want to "tie its hands" and give up its ability to grant protection. For example, in Grossman and Helpman (1994), lobbies compensate the government for the distortions associated with trade policy, and hence there is no reason why the government would want to commit

[1] For a real-world illustration of the possible domestic-commitment role of trade agreements, we refer the reader to Fernando Salas and Jaime Zabludovsky (2004), who argue that the main benefit of NAFTA for Mexico has been the strengthening of its commitment to maintain an open trade and investment regime.

not to grant protection. In fact, if the government is able to extract rents from the political process it is strictly better off in the political equilibrium than under free trade. But this may no longer be true when one takes into account that capital can move across sectors. This is because, given the expectation of protection in a sector, there will be excessive investment in that sector. Since this happens before the government and lobbies negotiate over protection, the government is not compensated for this "long-run" distortion. This allocation distortion is the essence of the domestic-commitment problem.

Next we explain how we develop a political-economy theory of trade agreements based on the domestic-commitment problem just described. We consider two large countries whose respective governments are subject to pressures from import-competing lobbies, and where capital is fixed in the short run but mobile in the long run. In this setting, the noncooperative equilibrium entails two types of inefficiency: a domestic time-inconsistency problem and a prisoner's dilemma arising from the terms-of-trade externality. Starting from the noncooperative equilibrium, the two countries get a chance to sign a trade agreement. After the agreement is signed, investors can re-allocate their capital (subject to possible frictions), and then governments choose trade policies subject to the constraints set by the agreement. A key parameter of the model is the degree to which capital can move across sectors. This parameter captures the importance of the domestic-commitment problem: if capital mobility is zero, the model collapses to the standard one where the only motive for trade agreements is the terms-of-trade externality; as capital mobility increases, the domestic-commitment problem becomes more severe.

We distinguish between "ex-ante lobbying," which influences the selection of the trade agreement, and "ex-post lobbying", which influences the choice of trade policies subject to the constraints set by the agreement. Of course, the notion of ex-post lobbying is meaningful only if the agreement leaves some *discretion* in the governments' choice of trade policies after the agreement is signed. This is the case, for example, if the agreement imposes a tariff ceiling, so that a government is free to choose a tariff below the ceiling level. This way of thinking about trade agreements is a significant departure from the existing models, where agreements leave no discretion to governments, and which therefore cannot make a meaningful distinction between ex-ante and ex-post lobbying.[2]

Another novel feature of our model relative to the existing literature is that it integrates both

[2] A notable exception is Ornelas (2005), who explores the importance of ex-ante lobbying for the welfare implications of regional free-trade agreements.

motives for trade agreements, namely terms-of-trade externalities and domestic-commitment problems. As we will show, these two motives interact in nontrivial ways, giving rise to interesting empirical predictions.

In the first part of the paper we present a simple two-period model that, in spite of its highly stylized nature, is capable of conveying most of our main points, and is useful to relate our results to the existing literature. As we argue later, the two-period model can be broadly viewed as a reduced form of a fuller dynamic specification.

The main results of the basic model are three. First, we show that the degree of capital mobility is a key determinant of the extent of trade liberalization. In particular, we find that trade liberalization is deeper when capital is more mobile across sectors. To understand this result, consider the extreme case in which capital can be freely reallocated after the agreement has been signed. In this case, the import-competing lobbies suffer no loss from trade liberalization, since capital can exit the affected sectors and avoid any losses associated with lower domestic prices. With imperfect capital mobility, however, trade liberalization does generate losses for import-competing lobbies, so these lobbies will resist trade liberalization. Although our model generates only a comparative-statics result, it nevertheless suggests a cross-sectional empirical prediction: we should observe deeper trade liberalization in sectors where capital is more mobile. We are not aware of any empirical work exploring the link between factor mobility and trade liberalization, but casual observations seem to be in line with our model's prediction: for example, trade liberalization has been very limited in the agricultural sector, which is intensive in resources that are not very mobile (e.g. land).

Second, the model generates interesting results regarding the impact of "politics" on trade liberalization. We find that, if the domestic-commitment motive for the trade agreement is strong enough, trade liberalization is deeper when governments are more politically motivated (in the sense that they care more about political contributions). This contrasts with the standard theory of trade agreements, where trade liberalization tends to be *less* deep when governments are more politically motivated (the reason being that stronger political motivations lead to a lower trade volume in the non-cooperative equilibrium, hence to a weaker terms-of-trade externality, which calls for a smaller reduction in tariffs). The difference in predictions arises from the fact that in our model the domestic-commitment motive for a trade agreement is *directly* determined by the presence of politics, whereas in the standard theory politics affects trade agreements only *indirectly* through trade volumes.

Third, the model can explain why trade agreements typically specify tariff ceilings rather than exact tariff levels. Tariff ceilings and exact tariff commitments have very different implications. With exact tariff commitments, lobbying effectively ends at the time of the agreement, since the agreement leaves no discretion for governments to choose tariffs in the future. With tariff ceilings, on the other hand, governments retain the option of setting tariffs below their maximum levels, and this will invite lobbying and contributions also after the agreement is signed. We show that tariff ceilings are preferred to exact tariff commitments. The broad intuition is that, if the commitment takes the form of tariff ceilings, lobbies will be induced to pay contributions ex post, and this will lower the net return to capital, thus mitigating the overinvestment problem. Interestingly, then, keeping the lobbying game alive can help reduce the distortions caused by lobbying itself. We also emphasize that our model may help explain why trade agreements are incomplete contracts, without relying on the traditional causes for this, such as contracting costs or nonverifiable information.[3]

In the second part of the paper we analyze a full continuous-time specification of the model. In addition to providing dynamic foundations for the two-period model, this specification generates some important insights concerning the dynamics of trade liberalization.

We consider a scenario in which the trade agreement constrains the future path of tariffs, and show that the optimal agreement is made of two components: an immediate slashing of tariffs relative to their noncooperative levels, and a subsequent phase of gradual tariff reduction. The immediate drop in tariffs is due to the terms-of-trade motive for the trade agreement, while the domestic-commitment motive is reflected in the gradual component of trade liberalization. We also find that the *speed* of trade liberalization is higher when capital is more mobile.

Gradualism in our model emerges due to the interaction between frictions in capital mobility and ex-ante lobbying by capital owners. On their own, governments would want to implement free trade immediately, so it is costly for lobbies to "convince" them otherwise. The lobbies, which are composed of capital owners currently "stuck" in the import-competing sectors, are willing to offer contributions in order to keep some protection in the near future, but not to

[3] In the literature there are two papers that offer alternative explanations for the use of tariff ceilings in trade agreements. Horn, Maggi and Staiger (2005) examine the optimal structure of trade agreements in the presence of contracting costs. They show that, in order to save on contracting costs, it may be optimal to specify rigid (i.e. noncontingent) tariff ceilings. Bagwell and Staiger (2005) propose a model where tariff ceilings are motivated by the presence of privately observed – and therefore nonverifiable – shocks in the political pressures faced by governments. The explanation for tariff ceilings proposed in the present paper is quite different from those proposed in the above two papers, since it does not rely on the presence of contracting costs or verification problems. We will further discuss the relationship with those papers at the end of section 2.3.

keep protection far into the future, since by then all the capital will have become "unstuck." The result is gradual trade liberalization. This explanation, based on domestic commitment problems and imperfect capital mobility, is novel and – we feel – empirically plausible.

In our basic model, the trade agreement comes as a surprise to investors. The model can be easily extended to consider the case in which investors can foresee the trade agreement. In this case, we show that the qualitative results are the same as in the case of a surprise agreement, except that the re-allocation of capital occurs partly before the agreement is signed and partly afterwards. Thus, the domestic-commitment problem is solved in two phases: first, the (credible) announcement of an agreement, with the consequent reaction of investors; and second, the implementation of the trade agreement itself.

We want to emphasize that most of our insights follow from our structural modeling of the lobbying game, in which interest groups and governments exchange contributions for trade protection. If we modeled political pressures with a reduced-form approach, by assuming that governments attach a higher weight to producer surplus than to the other components of welfare, and we kept lobbies and contributions in the background, we would lose most of our results. In particular, one might be tempted to model the domestic-commitment problem by assuming that there is a divergence between ex-ante and ex-post government objectives (e.g. at the stage of signing the agreement governments maximize welfare, while ex-post they maximize a combination of welfare and industry profits). This reduced-form setup would *not* be equivalent to our structural setup: for example, in the reduced-form setup there would be no role for tariff ceilings, and there would be no gradualism in trade liberalization.

This paper is related to two literatures: first, the literature on trade agreements motivated by terms-of-trade externalities (see the papers cited at the beginning of this introduction); and second, the literature on trade agreements motivated by domestic-commitment problems. In this second group, Maggi and Rodríguez-Clare (1998) and Mitra (2002) have highlighted the role of politics in creating demand for commitment, while Staiger and Tabellini (1987) have focused on purely economic considerations. However, these three papers focus on a single small economy and do not attempt a full-fledged analysis of trade agreements. One important disadvantage of a small-country model is that it does not allow one to study the interaction between the terms-of-trade and domestic-commitment motives for a trade agreement.[4]

[4] We also note that in Maggi and Rodríguez-Clare (1998) the government is only allowed to choose between two extreme options, namely free trade or no commitment at all. If we want to study what determines the *extent* of trade liberalization, we need to allow governments to commit to intermediate levels of trade protection

A recent paper that considers a two-country model of trade agreements in the presence of domestic commitment problems is Conconi and Perroni (2005). They consider a self-enforcing agreement between a large country and a small country, where the only motive for a trade agreement is a domestic commitment issue that affects the small country.[5] In contrast, our model integrates both motives for trade agreements, namely terms-of-trade externalities and domestic-commitment problems. Another important difference is that they take a reduced-form approach where there is a divergence between ex-ante and ex-post objectives of the governments. As we pointed out above, this approach is not equivalent to our structural approach where lobbying and contributions are explicitly modeled; most of our points could not be made with a reduced-form approach. In any event, Conconi and Perroni's paper makes very different points from ours, as they focus on the implications of the self-enforcement constraints and argue that they can explain the granting of temporary Special and Differential treatment to developing countries in the WTO.

Also related to our paper is the literature on gradual trade liberalization. This literature includes Staiger (1995), Deveraux (1997), Furusawa and Lai (1999), Chisik (2003), Bond and Park (2004), Conconi and Perroni (2005) and Lockwood and Zissimos (2005). In these papers, gradual trade liberalization is explained as a consequence of the self-enforcing nature of the agreements. Indeed, in these models trade liberalization would occur at once if agreements were perfectly enforceable, or if players were sufficiently patient.[6] In our model, on the other hand, gradualism emerges even though agreements are perfectly enforceable; as we remarked above, gradualism in our model is a consequence of the interaction between frictions in capital mobility and lobbying by capital owners.

The paper is organized as follows: section 2 presents the basic two-period model; section 3 presents the full continous-time model; and section 4 offers some concluding remarks.

– which we do in the present paper. Moreover, that paper does not consider the possibility that lobbies might influence the shaping of the trade agreement ("ex-ante" lobbying), which plays an important role in the present paper. Finally, in this paper we allow for imperfect capital mobility, whereas Maggi and Rodríguez-Clare (1998) only consider the case of perfect capital mobility.

[5] Conconi and Perroni (2004) consider a model of self-enforcing international agreements between two large countries where there is both a domestic commitment problem and an international externality. This paper is different from ours in that it analyzes issues of self-enforcement in a model with very little structure and thus offers no implications for the extent of trade liberalization brought about by trade agreements, which is the focus of our present paper.

[6] One exception is Furusawa and Lai (1999), where trade liberalization may be gradual even with perfect enforceability if there are market imperfections, such as job-search externalities, that lead to excessively fast exit under immediate liberalization (as in Mussa, 1986). Our explanation of gradualism does not rely on such market imperfections.

2 A two-period model

In this section we present a very stylized two-period model that allows us to convey most of our main points in a relatively familiar setting. In the next section we will examine a fully dynamic model where time is continuous and the horizon is infinite, and there it will become clear that the two-period model can be broadly viewed as a reduced form of the full dynamic specification. Presenting the two-period model before moving on to the full dynamic specification has two further benefits: it allows for an easy comparison with the results of previous models, particularly Grossman and Helpman (1995a), Bagwell and Staiger (2001), and our previous work (Maggi and Rodríguez-Clare, 1998); and it is pedagogically useful, because understanding the results of the continuous-time model will be easier and more intuitive after seeing the simpler two-period model.

There are two countries, Home (H) and Foreign (F), and three goods: one numeraire good, denoted by N, and two manufacturing goods, denoted by M_1 and M_2. In both countries preferences are given by $U = c_N + \sum_{i=1}^{2} u(c_i)$, where c_i denotes consumption of good M_i. We assume $u(c_i) = vc_i - c_i^2/2$ (where v is a positive parameter), so that the demand function for good M_i is $d(p_i) = v - p_i$. The consumer surplus associated with good M_i is $s(p_i) = u(d(p_i)) - p_i d(p_i)$.

There are two types of capital, type 1 and type 2. The M_1 good is produced one-for-one from type-1 capital, and the M_2 good is produced one-for-one from type-2 capital. Each country is endowed with one unit of each type of capital. The only difference between the two countries is in the technology to produce the N good: in country H, the N good is produced one-for-one from type-1 capital, while in country F, the N good is produced one-for-one from type-2 capital. Given these assumptions, Home is a natural importer of good M_1 and Foreign is a natural importer of good M_2.[7] The reason we chose this particular production structure is that it generates a simple symmetric setup where, in each country, capital mobility is relevant only between the import-competing sector and the numeraire sector. This in turn ensures that in each country the domestic-commitment motive for trade agreements concerns the import-competing sector but not the export sector, a feature that simplifies the analysis considerably.

Home chooses a specific tariff t on imports of M_1 and Foreign chooses a specific tariff t^* on

[7] More precisely, under free trade Home imports a nonnegative amount of good M_1 and Foreign imports a nonnegative amount of good M_2. Note that this is true for any given allocation of capital. Without imposing further conditions it is possible that there is no trade in equilibrium, but below we will assume a parameter condition that prevents this possibility.

imports of M_2. Thus, if tariffs are not prohibitive, the domestic price of good M_1 in Home is given by $p_1 = p_1^* + t$. Similarly, the domestic price of good M_2 in Foreign is $p_2^* = p_2 + t^*$.[8]

Let x (x^*) denote the level of capital allocated to sector M_1 (M_2) in country H (F). Welfare (i.e., utility of the representative agent) is given by factor income plus tariff revenue plus consumer surplus. Thus, welfare in Home and Foreign, respectively, is given by:

$$\begin{aligned} W &= (1-x) + (p_1 x + t m_1 + s_1) + (p_2 + s_2) \\ W^* &= (1-x^*) + (p_2^* x^* + t^* m_2^* + s_2^*) + (p_1^* + s_1^*) \end{aligned}$$

where m_i (m_i^*) denotes Home (Foreign) imports of good i and s_i (s_i^*) represents Home (Foreign) consumer surplus derived from good i. Note the separability between sectors M_1 and M_2. Specifically, note that we can express W as the sum of two components: the first one, $(1-x) + (p_1 x + t m_1 + s_1)$, depends on t and x; and the second one, $p_2 + s_2$, depends on t^* and x^*. The same separability applies to foreign welfare. Together with symmetry, this separability implies that we can focus on sector M_1; the equilibrium in sector M_2 will be its mirror image. Thus, to simplify notation, we drop the subscript 1 from now on, and simply refer to sector M_1 as the "manufacturing" sector.

The international market clearing condition for manufacturing is $d(p) + d(p^*) = x + 1$. This yields

$$\begin{aligned} p^*(t,x) &= v - \frac{1}{2}(x+1+t) \\ p(t,x) &= v - \frac{1}{2}(x+1-t) \end{aligned}$$

where we emphasize the dependence of equilibrium prices on the tariff and the capital allocation in the home country.

Letting $m = d(p) - x$ denote imports of manufactures by Home, then $m(t,x) = \frac{1}{2}(\Delta x - t)$, where $\Delta x \equiv 1 - x$ is the difference in supply between the two countries. Given this notation,

[8] In this paper we do not consider export subsidies and taxes. If the agreement takes the traditional form of exact tariff and subsidy commitments, this restriction is innocuous, because only net protection (i.e. the difference between import tariff and export subsidy in a given sector) matters for the optimal agreement, therefore t and t^* can be reinterpreted in terms of net protection in the two sectors. If the agreement takes the form of tariff and subsidy ceilings, on the other hand, not only net protection but also the levels of import tariffs and export subsidies matter, and this makes the analysis substantially more complex. Maggi and Rodriguez-Clare (2006) study optimal agreements when both import and export instruments are allowed but there is no capital mobility. We also note that assuming away export instruments is relatively common in the existing literature on trade agreements: see for example Grossman and Helpman (1995b), Krishna (1998), Maggi (1999) and Ornelas (2004).

welfare in Home is:

$$W(t,x) = (1-x) + p(t,x)x + tm(t,x) + s(t,x) + [\cdot]$$

where $[\cdot]$ does not depend on t and x. Analogously, Foreign welfare is

$$W^*(t,x) = p^*(t,x) + s^*(t,x) + [\cdot]$$

We now turn to the political side of the model. We assume that, in each country, the capital owners in the import-competing sector get organized as a lobby and offer contributions to their government in exchange for protection.[9] We model the interaction between lobby and government in a similar way as Grossman and Helpman (1994). We assume that the political structure is symmetric in the two countries, so we can focus on the Home country.

The government's objective function is $U^G = aW + C$, where C denotes contributions from the import-competing lobby. The parameter a captures (inversely) the importance of political considerations in the government's objective: when a is lower, "politics" are more important.

The lobby maximizes total returns to capital net of contributions, $U^L = px - C$.[10] The lobby collects contributions in proportion to the amount of capital in the manufacturing sector, thus total contributions are given by $C = cx$, where c is the contribution per unit of capital.

2.1 The noncooperative equilibrium

The timing of the non-cooperative game is the following. In the first stage, investors allocate their capital. The value of x summarizes the choices of investors in this stage. In the second stage, the government and the import competing lobby in each country bargain efficiently over tariff and contributions. For simplicity we assume that the lobby has all the bargaining power (in a later section we will discuss how results would change if the government had some bargaining power). An equivalent assumption would be that the lobby makes a take-it-or-leave-it offer to the government that consists of a tariff level and a contribution level.To determine the subgame perfect equilibria of the game we proceed by backward induction, starting with

[9] We are implicitly assuming that the export sector and the numeraire sector are not able to get organized. This is a simple lobby structure that generates trade protection in the political equilibrium.

[10] This is a shortcut. To be more precise, we should specify the lobby's objective as the aggregate well-being of its lobby members, but this would give rise to the same results. Letting α be the fraction of the population that owns some capital in the import-competing sector, the lobby's objective is $px + \alpha(tm + s) - C$, so the joint surplus of government and lobby is proportional to $\frac{a+\alpha}{1-\alpha}W + px$, an expression that has the same qualitative structure as the one we derive below.

the determination of equilibrium tariffs and contributions given the allocation of capital. This is the equilibrium of the subgame, or the "short-run" equilibrium.

Given the assumption of efficient bargaining, the government (G) and the lobby (L) in the Home country choose t to maximize their joint surplus:

$$J^{SR}(t,x) = aW(t,x) + p(t,x)x$$

This yields

$$t = t^J(x) \equiv (1/3)(\Delta x + 2x/a)$$

The noncooperative tariff t^J can be decomposed in two parts. The component $\Delta x/3$ captures the incentive to distort terms of trade: when the supply difference Δx is bigger, the volume of imports is larger, and hence this incentive is stronger. The component $2x/3a$ captures the political influence exerted by the lobby. This component is more important when the sector is larger (x is higher) and when the government's valuation of contributions relative to welfare is higher (a is lower). We let the national welfare maximizing tariff (given x) be denoted by

$$t^W(x) \equiv \lim_{a\to\infty} t^J(x) = \Delta x/3$$

For future reference, we define $c(t,x)$ as the contributions per unit of capital such that G is just willing to impose tariff t, or in other words, such that G is kept at its reservation utility given tariff t. In the absence of contributions, G would choose the welfare maximizing tariff given x, that is $t^W(x)$, so G's reservation utility is $W(t^W(x),x)$. Since the short-run equilibrium tariff given x cannot be below $t^W(x)$, we only need to focus on the case $t \geq t^W(x)$. For the government to choose a tariff $t \geq t^W(x)$, total contributions would have to be equal to

$$a\left[W(t^W(x),x) - W(t,x)\right] = (3a/8)\left(t - t^W(x)\right)^2$$

Thus, the function

$$c(t,x) \equiv (3a/8x)\left(t - t^W(x)\right)^2$$

determines the contributions per unit of capital necessary to induce the government to choose tariff $t \geq t^W(x)$. Note that we need to define this function only for $t \geq t^W(x)$, since the lobby would never pay the government to impose a tariff lower than the government would choose on its own.

We now move back one step, to examine the "long-run" non-cooperative equilibrium, where x is endogenous and is determined according to investors' expectations about future protection in the absence of a trade agreement. Before we proceed, however, it is useful to derive the free trade long-run equilibrium. Suppose that in this equilibrium Home produces both the N good and the M good (in a moment we will impose a parameter condition that ensures this). Then the domestic (and international) price of the M good must be equal to one. Thus the free trade allocation of capital, x^{ft}, is determined by the condition $p(0, x^{ft}) = 1$, or $v - \frac{1}{2}(x^{ft} + 1) = 1$. To ensure that $0 < x^{ft} < 1$ we impose the condition $3/2 < v < 2$.[11] We maintain this assumption throughout the rest of the paper. Note that, because of the symmetry of the model, under free trade there is no trade in the numeraire sector.

We can now turn to the long-run equilibrium. The equilibrium conditions are:

$$\begin{aligned} t &= t^J(x) \\ p(t, x) - c(t, x) &= 1 \end{aligned} \tag{1}$$

The second condition requires that the return to capital net of contributions be equal in the import-competing sector and in the numeraire sector. This equal-returns condition implicitly defines a curve in (t, x) space that we label $x^{er}(t)$.[12]

We let $(\hat{t}, \hat{x})$ denote a solution to the above system. Also, we let $\left(t^W, x^W\right)$ denote the intersection of the curves $t^W(x)$ and $x^{er}(t)$. Note that this is the long-run noncooperative equilibrium in the benchmark case of welfare-maximizing governments (i.e., $(t^W, x^W) \to (\hat{t}, \hat{x})$ as $a \to \infty$). The proof of the following proposition, together with all the other proofs of the paper, can be found in Appendix.

Proposition 1 *If $a > (6v - 7)/6(2 - v)$ there exists a unique long-run noncooperative equilibrium. In this equilibrium each country imposes a positive but non-prohibitive tariff $\hat{t}$. The equilibrium tariff $\hat{t}$ is decreasing in a, and approaches t^W as $a \to \infty$.*

Figure 1 illustrates the long-run noncooperative equilibrium. In the figure, the $t^J(x)$ curve is increasing, but nothing would change if it were decreasing. To understand the shape of the $x^{er}(t)$ curve, note that since the lobby has all the bargaining power and extracts all the joint surplus, $t^J(x)$ maximizes the net returns to capital in the M sector (i.e. $p - c$). Given concavity of J^{SR} in t, this implies that $p - c$ is increasing in t below the $t^J(x)$ curve and is decreasing in

[11] If $v \leq 3/2$ then $x^{ft} = 0$, so there is no import-competing industry, and if $v \geq 2$ then $x^{ft} = 1$, so there is no trade at all.

[12] Since we defined the function $c(t, x)$ only for $t \geq t^W(x)$, the curve $x^{er}(t)$ is defined only in the region $t \geq t^W(x)$.

t above this curve. Under the condition assumed in the proposition, entry into the M sector has the intuitive effect of reducing net returns to capital there (i.e., $p-c$ is decreasing in x). It follows that, under this condition, the equal-returns curve $x^{er}(t)$ is increasing below the $t^J(x)$ curve and decreasing above it, with a slope of zero at the $(\hat{t}, \hat{x})$ point. In the rest of the paper we maintain the assumption $a > (6v-7)/6(2-v)$, which ensures the existence and uniqueness of the long-run equilibrium.

Not surprisingly, for any positive but finite level of a, the non-cooperative tariff is higher than the national welfare-maximizing tariff: $\hat{t} > t^W$. Also, from inspection of Figure 1, it is clear that the noncooperative equilibrium allocation $\hat{x}$ exceeds the allocation that would result in the absence of politics (i.e. when $a \to \infty$), that is $\hat{x} > x^W$. As we will show formally in a later section, this excess of $\hat{x}$ above x^W represents an overinvestment problem, or a "long-run" distortion associated with the government's lack of commitment vis-à-vis domestic investors. Each government is compensated by its lobby for the short-run distortion associated with protection (i.e. the consumption distortion given x), but is not compensated for the long-run allocation distortion. For this reason a government may value a commitment to a lower level of the tariff. This is the heart of the domestic-commitment motive for trade agreements, which operates alongside the standard terms-of-trade motive.[13] We are now ready to examine the optimal agreement.

2.2 The optimal trade agreement

We suppose that, before capital is allocated, the two governments can sign a trade agreement. In Maggi and Rodríguez-Clare (1998) we assumed that lobbies do not influence the selection of the trade agreement, i.e. there is no ex-ante lobbying. Here we allow for ex-ante lobbying by assuming that the agreement maximizes the ex-ante joint surplus of the two governments and the two lobbies.[14]

[13] The reader might wonder whether our results rely on the assumption that the supply of the M good is fixed (at $1+x$) in the short run. If supply were responsive to prices also in the short run (which would be the case if we introduced a fully mobile factor – e.g. labor – in the model), the short-run distortion associated with trade protection would include also a production distortion (misallocation of labor), not just a consumption distortion, but the main qualitative results of the model would be unlikely to change.

[14] This efficient-agreement approach can be justified as equivalent to a more structural game between governments and lobbies. One possibility would be to consider a game along the lines of Grossman and Helpman's (1995) "Trade Talks" model. Suppose that (1) each lobby offers a contribution schedule to its own government; and (2) governments bargain efficiently (with symmetric bargaining powers) given the contribution schedules. One can show that the equilibrium outcome of this game maximizes the joint surplus of governments and lobbies. Note that, given the symmetry of the model, there is no need to have international transfers for this

The agreement maximizes the following objective:

$$\Psi = U^G + U^{G^*} + U^L + U^{L^*} \tag{2}$$

where U^G, U^{G^*}, U^L and U^{L^*} denote the second-stage payoffs of the governments and lobbies as viewed from the ex-ante stage. In section 2.4 we will discuss how results would change in the absence of ex-ante lobbying (i.e. if the agreement maximizes $U^G + U^{G^*}$).

The trade agreement is assumed to be perfectly enforceable. In the concluding section we will discuss how the insights of our model might extend to a setting of self-enforcing agreements.

We assume that the inherited level of x at the agreement stage is equal to $\widehat{x}$, the long-run equilibrium allocation in the absence of an agreement. The interpretation is that the commitment opportunity comes as a surprise to the private sector. In section 3.1 we show that this element of surprise is not essential to our main results: if the agreement is fully anticipated by investors, the qualitative results are essentially preserved.

Following the agreement, but before trade policy is determined, each capital owner gets a chance to reallocate her capital with probability $z \in [0, 1]$. Assuming that capital-owners are "small" and that the opportunity to reallocate their capital is independent across capital owners, this implies that a fraction z of the capital in sector M has the opportunity to exit. The parameter z captures the ease with which capital can be reallocated. The case $z = 0$ captures the case in which capital is "stuck" in the M sector, whereas the case $z = 1$ captures a situation in which capital is perfectly mobile in the long run but fixed in the short run. With a slight abuse of terminology, from now on we refer to the case $z = 1$ simply as "perfect capital mobility", and to the case $z < 1$ as "imperfect capital mobility".

To recapitulate, the timing of the model is as follows: (1) the agreement is selected; (2) capital is reallocated (when feasible); and (3) given the capital allocation and the constraints (if any) imposed by the agreement, each government-lobby pair chooses a tariff.

Again, given separability and symmetry across the two manufacturing sectors, we can analyze them independently. Thus, just as in the previous sections, we can focus on sector M_1 (omitting subscripts) and find the optimal agreement by maximizing the joint surplus of the two governments and Home's import-competing lobby in this sector.

We will consider two forms of agreement: agreements that specify tariff ceilings, that is constraints of the type $t \leq \bar{t}$, and agreements that specify exact tariff levels, that is constraints

result to hold; but in a more general asymmetric situation, transfers would be needed in order to justify a joint-surplus-maximizing approach.

of the type $t = \bar{t}$. The main difference between these two types of agreement is that in the case of exact tariff commitments the lobby will not have to pay contributions to obtain protection ex-post, since such protection is effectively part of the agreement. Under tariff ceilings, on the other hand, the government can credibly threaten to impose its unilateral best tariff $t^W(x)$ (if $t^W(x) < \bar{t}$). Thus, the lobby would have to compensate the government for deviating from this tariff, and there would be positive contributions ex-post. In what follows we characterize the optimal agreement with tariff ceilings. In section 2.3 we will show that tariff ceilings weakly dominate exact tariff commitments.

It is instructive to start by characterizing and contrasting the two benchmark cases of perfect capital mobility ($z = 1$) and fixed capital ($z = 0$), and then consider the more general case of imperfect capital mobility.

Perfect capital mobility

Recall that the optimal agreement maximizes the ex-ante joint surplus of the two governments and the importing lobby in each sector. Given that $\widehat{x}$ is the inherited allocation of capital, this objective function can be written as:

$$\Psi(t, x) = aW(t, x) + aW^*(t, x) + xp(t, x) + (\widehat{x} - x) \tag{3}$$

This expression is valid only for $x \leq \widehat{x}$, but we do not need to consider the alternative case $x > \widehat{x}$, because this can never hold in equilibrium. To gain a better understanding about this objective function, note that $\Psi = J^{SR} + aW^* + (\widehat{x} - x)$. There are two extra terms relative to the short-run objective J^{SR}: the term aW^*, which takes into account terms of trade externalities, and the term $(\widehat{x} - x)$, which captures the rents of those lobby members that will move to the N sector in the following period.

The analysis proceeds by backward induction, in three steps: (i) we solve for the equilibrium tariff and contribution as functions of the tariff ceiling and the capital allocation, $t(\bar{t}, x)$ and $c(\bar{t}, x)$;[15] (ii) we solve for the equilibrium allocation as a function of the tariff ceiling, $x(\bar{t})$; and (iii) we express the ex-ante objective Ψ as a reduced-form function of the ceiling $\bar{t}$ and then find the optimal ceiling.

To derive $t(\bar{t}, x)$, notice that this is the tariff that maximizes $J^{SR}(t, x)$ subject to the constraint $t \leq \bar{t}$, and recall that $J^{SR}(t, x)$ is concave in t and maximized at $t^J(x)$. Therefore, if

[15] Note that we are using the same notation $c()$ as for the contribution schedule in the noncooperative equilibrium, even though this is not the same function. This is an abuse of notation, but the reader can distinguish the two functions because the first argument is t in one case and $\bar{t}$ in the other.

$\bar{t} \geq t^J(x)$ the tariff ceiling is not binding, hence $t(\bar{t}, x) = t^J(x)$; and if $\bar{t} < t^J(x)$ the ceiling is binding, so $t(\bar{t}, x) = \bar{t}$. Summarizing, $t(\bar{t}, x) = \min\{\bar{t}, t^J(x)\}$. Note that there is no loss of generality in focusing on agreements in which ceilings are not redundant, i.e. $\bar{t} \leq t^J(x)$. Thus, from now on we simplify notation by simply using $\bar{t}$ rather than $t(\bar{t}, x)$.

Turning to $c(\bar{t}, x)$, the key observation is that, if $\bar{t} > t^W(x)$, then the home government will get contributions, because its outside option in the negotiation with the lobby is given by the tariff $t^W(x)$, and the lobby has to compensate G for raising the tariff towards the ceiling $\bar{t}$; on the other hand, if $\bar{t} < t^W(x)$ no contributions will be forthcoming, because G has no credible threat. Thus

$$c(\bar{t}, x) = \begin{cases} (3a/8x)\left(\bar{t} - t^W(x)\right)^2 \text{ if } \bar{t} \geq t^W(x) \\ 0 \text{ if } \bar{t} < t^W(x) \end{cases}$$

The next step is to derive the equilibrium allocation conditional on $\bar{t}$. Clearly, if $\bar{t} > \hat{t}$ then the tariff ceiling is not binding, and the equilibrium will be given by $(\hat{t}, \hat{x})$, just as if there were no agreement. On the other hand, if $\bar{t} \leq \hat{t}$ then the equilibrium allocation is implicitly defined by the equal-returns condition

$$p(\bar{t}, x) - c(\bar{t}, x) = 1$$

We let $x^{er}(\bar{t})$ denote the solution in x to the above equation for $\bar{t} \leq \hat{t}$.[16] Figure 2 illustrates the curve $x^{er}(\bar{t})$. Below the $t^W(x)$ curve, this is a line with slope one (because in this region the condition that defines it is $p(\bar{t}, x) = 1$), and between the curves $t^W(x)$ and $t^J(x)$ it coincides with the equal-returns curve in the absence of agreements (which is depicted in Figure 1).

We can now move back one more stage in our backward induction analysis and derive the optimal trade agreement. Clearly, the optimal tariff ceiling is the one that maximizes $\Psi(\bar{t}, x^{er}(\bar{t}))$ for $\bar{t} \leq \hat{t}$. The next result shows that $\Psi(\bar{t}, x^{er}(\bar{t}))$ is maximized at free trade:

Proposition 2 *In the case of perfect capital mobility, the optimal agreement is* $\bar{t}^A = 0$ *(free trade) for all* a.

When capital is perfectly mobile, the optimal agreement is free trade even in the presence of ex-ante lobbying. Intuitively, if capital is mobile, the lobby anticipates that there cannot be any rents in the ex-post stage, and hence is not willing to pay anything to compensate the government for the long run distortions associated with protection. This will of course no longer be true when capital is imperfectly mobile, as we will show below.

[16] Again, the notation $x^{er}(\bar{t})$ is slightly abused because this is not the same function as $x^{er}(t)$, the equal-returns condition in the absence of agreements, which was defined only for $t \geq t^W$.

In this model there are two motives for a trade agreement: the standard terms-of-trade (TOT) externality and the domestic-commitment problem. We can disentangle the two motives with the following thought experiment. We consider a hypothetical scenario in which the home government can commit *unilaterally* (subject to the lobby's pressures), and characterize the tariff ceiling that would be chosen in this case ($\bar{t}^{DC}$). The movement from $\hat{t}$ to $\bar{t}^{DC}$ can be thought of as the component of trade liberalization that is due to the domestic-commitment motive, while the movement from $\bar{t}^{DC}$ to $\bar{t}^{A} = 0$ can be thought of as the component due to the TOT motive.

We can also view this thought experiment from a slightly different perspective. A trade agreement may provide governments with the credibility to make unilateral commitments, not only the opportunity to negotiate reciprocal commitments. Thus we can think of the benefits from a trade agreement as stemming from two sources: first, a country's *membership* in the agreement, which allows a country to commit unilaterally, thereby solving its credibility problem in the domestic arena; and second, the negotiation of reciprocal commitments, which takes care of TOT externalities.[17] In this perspective, what we do here can be interpreted as disentangling the role of membership from the role of negotiated tariff reductions.

Formally, we consider the following timing: the home government and the lobby choose a tariff ceiling without negotiating with the foreign government; then capital is allocated, and then the home government and the lobby choose the tariff given the ceiling and the capital allocation. The objective is the same as in the previous case except that foreign welfare is not taken into account. So $\bar{t}^{DC}$ maximizes

$$J(\bar{t}, x) \equiv aW(\bar{t}, x) + xp(\bar{t}, x) + (\widehat{x} - x)$$

subject to $x = x^{er}(\bar{t})$. The following result characterizes $\bar{t}^{DC}$:

Proposition 3 *In the case of perfect capital mobility, if a government can commit unilaterally it will choose* $\bar{t}^{DC} = t^{W}$.

This result is illustrated in Figure 2. Note that the TOT component of the agreement, i.e. the difference $\bar{t}^{DC} - \bar{t}^{A}$, is just given by t^{W}, which is the tariff that optimally exploits a country's

[17] We emphasize however that this decomposition into gains from membership and gains from negotiated commitments is purely conceptual, and may not have an empirically observable counterpart. When a group of countries gets together to form a trade agreement, to the extent that the agreement is motivated by both TOT and domestic-commitment considerations, the agreed-upon tariff cuts are likely to incorporate both motivations.

monopoly power over TOT (taking into account the endogeneity of the allocation x). Thus, the TOT component of the agreement removes the economically-motivated part of trade protection, while the domestic-commitment component of the agreement removes the politically-motivated part of trade protection.[18]

It is important to note that the TOT component of the agreement is independent of politics, i.e. t^W is not affected by a (straightforward algebra reveals that $t^W = 1 - \frac{v}{2}$). On the other hand, the domestic-commitment component of the agreement, $\hat{t} - \bar{t}^{DC}$, is larger when politics are more important (a is lower).[19]

Fixed capital

In the previous section we derived the optimal agreement for the case in which capital can be freely reallocated after the agreement is signed. We now consider the opposite extreme, in which capital cannot be re-allocated at all ($z = 0$). It is useful to consider this benchmark case for two reasons: first, the more general case of imperfect capital mobility ($z \in [0, 1]$) will be easier to understand after looking at the extreme cases $z = 1$ and $z = 0$; and second, the case of fixed capital will serve to establish the link between this model and the "standard" models in which trade agreements are motivated only by TOT externalities.

If capital is fixed at some level x, the optimal tariff ceiling is simply the one that maximizes $\Psi(\bar{t}, x)$.[20] Letting $t^{\Psi}(x) \equiv \arg\max_{\bar{t}} \Psi(\bar{t}, x)$, it is easy to show that $t^{\Psi}(x) = \frac{x}{a}$ (see Figure 2). Note that $t^{\Psi}(x)$ lies uniformly below $t^J(x)$ for $x \leq \hat{x}$, thus the extent of trade liberalization is given by $t^J(x) - t^{\Psi}(x) > 0$.

Next we want to decompose the optimal agreement into its domestic-commitment and

[18] A natural question is whether there exists a domestic-commitment motive for trade agreements even in the absence of politics (i.e. if governments maximize welfare). In our model the answer is no. To see this note that $\hat{t}$ approaches t^W as $a \to \infty$, which implies that there is no need for domestic commitment. Intuitively, with no politics, the government sets the tariff to maximize national welfare given x (i.e., $t = t^W(x)$), and as a consequence the investors' allocation decisions are efficient, yielding the (unilateral) optimal point (t^W, x^W). We note that this is a consequence of the supply structure assumed in our model, and does not hold more generally (for example, in Lapan, 1988, there is a time-inconsistency problem in tariff setting even though the government maximizes welfare). We regard this feature of our model as a useful simplification that allows us to focus sharply on politics as a determinant of the domestic commitment motive for trade agreements.

[19] The feature that the TOT component of the agreement is independent of politics is a consequence of perfect capital mobility, and does not hold when capital is imperfectly mobile, as will become clear in the next sections. But the case of perfect capital mobility is an important benchmark to keep in mind, because it helps to understand why the predictions of our model regarding the impact of politics on trade liberalization differ from those of the standard TOT theory.

[20] This follows immediately from the fact that the equilibrium tariff is equal to the ceiling level ($t = \bar{t}$). Note also that, with fixed capital, tariff ceilings are equivalent to exact tariff commitments, so the same points can be made by looking at either form of commitment.

TOT components. Following the methodology described before, we consider the domestic-commitment benchmark when x is fixed. The optimal tariff ceiling in this case maximizes $J(\bar{t}, x)$ for given x. Clearly, this objective is maximized by $t^J(x)$, that is, the optimum involves no agreement at all. We can conclude that, when x is fixed, the domestic-commitment component of the agreement is nil, and the whole tariff cut is coming from the TOT component. A domestic-commitment motive for trade agreements is present only if capital is mobile.

At this point it is useful to relate this case of fixed capital with the standard TOT story, and more specifically with Grossman and Helpman's (GH) (1995a) model. Note that our model with fixed capital is essentially a simplified version of GH's model. To see this, note that $\Psi(\bar{t}, x)$ reduces to

$$\Psi(\bar{t}, x) = aW(\bar{t}, x) + aW^*(\bar{t}, x) + xp(\bar{t}, x) + (\cdot)$$

where we omit the term in $(\cdot)$ because it is constant in $\bar{t}$. This is the joint surplus of the two governments and the lobby. As in GH's model, the optimal agreement maximizes this joint surplus.

Next consider the impact of the political parameter a on the extent of trade liberalization. It is easy to verify that the agreed-upon tariff cut is given by $t^J(x) - t^{\Psi}(x) = m(t^J(x), x)$. Thus the tariff cut is deeper when the noncooperative import volume is higher. This is intuitive: when imports are larger, the TOT externality is more important, and hence the trade agreement will cut the tariff by a larger amount. The observation that trade liberalization is increasing in the import volume has a straightforward implication for the comparative-statics effect of changes in a: when politics are more important (a is lower), the noncooperative tariff t^J is higher, hence the import volume is lower, and as a consequence the agreed-upon tariff cut is less deep.[21]

This prediction regarding the impact of "politics" on the extent of trade liberalization is a fairly general feature of models where trade agreements are motivated only by TOT externalities, as emphasized by Bagwell and Staiger (2001). This contrasts sharply with our earlier finding in the case of perfect capital mobility, where we found that the extent of trade liberalization is decreasing in a. This result points to an important insight: *when the domestic-commitment motive for a trade agreement is important, the impact of "politics" on the extent*

[21] How can this result be reconciled with the observation made earlier, that with perfect capital mobility the TOT component of the tariff cut (t^W) is independent of a? The key is to observe that, with perfect capital mobility, the TOT component of the tariff cut is also proportional to the import volume $m(t, x)$, but evaluated at the point (t^W, x^W), not at the noncooperative equilibrium. Since $m(t^W, x^W)$ is independent of a, so is the TOT component of the tariff cut.

of trade liberalization has the opposite sign as the one predicted by the standard TOT theory.

Imperfect capital mobility

We are now in a position to characterize the optimal agreement for the more general case $z \in [0,1]$. Let us start by considering the equilibrium conditional on a given tariff binding $\bar{t}$. To develop intuition, suppose that $\bar{t} < \hat{t}$ and z is small. From the analysis of the previous section, we know that if capital were perfectly mobile, the equilibrium allocation would be the one that equalizes returns given $\bar{t}$, that is $x^{er}(\bar{t}) < \hat{x}$. But if z is small, capital will not be able to exit the import-competing sector in sufficient amount to equalize net returns to capital across sectors. The allocation will then be $x_z \equiv (1-z)\hat{x}$ and the rate of return will be higher in the N sector. In general, the equilibrium allocation conditional on $\bar{t}$ is $\max\{x^{er}(\bar{t}), x_z\} \equiv \tilde{x}^{er}(\bar{t})$ if $\bar{t} \le \hat{t}$ and $\hat{x}$ otherwise. This is simply the equal-returns curve truncated at x_z.

This result implies that the optimal agreement is the one that maximizes $\Psi(\bar{t}, \tilde{x}^{er}(\bar{t}))$ for $\bar{t} \le \hat{t}$. Assuming that investors are risk neutral, what matters for the lobby is only the total expected future returns for the lobby members, which are given by $x(p-c)+(\widehat{x}-x)$. Thus, the same expression we had for $\Psi(\bar{t}, x)$ with perfect mobility is valid also with imperfect mobility. The key is that the parameter z enters the problem only through its effect on x.

Recall from the previous analysis that, if $z=0$, the optimal tariff ceiling is given by $t^{\Psi}(\hat{x})$, and if $z=1$, the optimal tariff ceiling is zero. The next proposition "connects the dots" between these two extremes. Let $t^{er}(x)$ be the inverse of $x^{er}(\bar{t})$ (in the relevant region $x^{er}(\bar{t})$ is increasing, so its inverse exists).

Proposition 4 *(i) The optimal tariff ceiling is given by*

$$\bar{t}^A = \begin{cases} \min(t^{er}(x_z), t^{\Psi}(x_z)) \textit{ for } x_z \ge x^{ft} \\ 0 \textit{ for } x_z < x^{ft} \end{cases}$$

(ii) The tariff cut $\hat{t} - \bar{t}^A$ is (weakly) increasing in z.

(iii) The tariff cut $\hat{t} - \bar{t}^A$ is increasing in a for sufficiently low values of z and decreasing in a for sufficiently high values of z.

Figure 3 illustrates how the optimal agreement point A depends on z. Consider two values of z, say z' and z''. For point z' the agreement is given by A', located on the $t^{\Psi}(x)$ curve, whereas for point z'' the agreement is given by point A'', located on the equal-returns curve. Thus, as z increases from zero, point A travels along the $t^{\Psi}(x)$ schedule until it hits the equal-returns curve $t^{er}(x)$, and then travels down along the $t^{er}(x)$ curve until it reaches the free trade point (this path is marked in bold in Figure 3). Note that both $t^{\Psi}(x)$ and $t^{er}(x)$ are increasing in x, so the optimal tariff binding decreases as z increases. As a consequence, the tariff cut $\hat{t} - \bar{t}^A$ increases with z, as stated in point (ii) of the proposition.

This result suggests an empirical prediction: trade agreements should lead to deeper trade liberalization in sectors where factors of production are more mobile. Although our basic model cannot yield cross-sectoral predictions because there is a single organized sector, it would not be hard to write a multi-sector model that delivers a genuinely cross-sectoral prediction along these lines.[22]

Point (iii) of Proposition 4 focuses on the impact of the political parameter a on the extent of trade liberalization. Recall from the previous analysis that, if $z = 0$, the tariff cut is deeper when politics are less important (a is higher). By continuity, this is the case also if z is sufficiently small. On the other hand, we saw that, if $z = 1$, the tariff cut is deeper when politics are *more* important (a is lower). Again, by continuity this is the case whenever z is sufficiently high. Thus the model highlights that, if the domestic-commitment motive is important enough, the prediction of the standard TOT model – that trade liberalization is deeper when politics are less important – gets reversed.

Finally we want to decompose the optimal agreement into its domestic-commitment and TOT components. We saw earlier that, if $z = 0$, the optimal domestic-commitment point is the same as the noncooperative equilibrium (there is no domestic-commitment motive for a trade agreement), while for $z = 1$ the optimal domestic-commitment point is (t^W, x^W). The next result connects the dots between these two extreme cases and shows that, as z increases, the optimal domestic-commitment point travels from the noncooperative point $(\hat{t}, \hat{x})$ to point (t^W, x^W) along the equal-returns curve:

Proposition 5 *If a government can commit unilaterally, it will choose*

$$\bar{t}^{DC}(z) = \begin{cases} t^{er}(x_z) \text{ for } x_z \geq x^W \\ t^W \text{ for } x_z < x^W \end{cases}$$

[22] Our model suggests also a couple of predictions of a cross-country nature. If one thinks of adjustment policies (e.g. trade adjustment assistance programs) as increasing the degree of resource mobility (z), the model suggests that trade agreements should lead to deeper tariff cuts when they involve countries with stronger adjustment policies. By a similar logic, the model suggests that trade agreements should lead to less deep tariff cuts when they involve countries with more rigid labor markets.

The domestic-commitment component of the agreed-upon tariff cut, $\hat{t} - \bar{t}^{DC}(z)$, is clearly increasing in z. What can we say about the effect of z on the TOT component, $\bar{t}^{DC}(z) - \bar{t}^{A}(z)$? In general the answer is ambiguous, but notice that for small z the TOT component of the tariff cut *decreases* with z. To see this, consider a small increase in z from zero. Then $\bar{t}^{DC}(z)$ goes down with infinite slope, while $\bar{t}^{A}(z)$ goes down with finite slope, therefore $\bar{t}^{DC}(z) - \bar{t}^{A}(z)$ decreases. Thus we can say that the liberalization-deepening effect of factor mobility is entirely due to the domestic-commitment motive, at least for z relatively small.

2.3 Tariff ceilings versus exact tariff commitments

Thus far we have focused on agreements that impose tariff ceilings. In this section we compare tariff ceilings with exact tariff commitments (ETC), i.e. equality constraints of the form $t = \bar{t}$. The key difference is that, if the agreement takes the form of an ETC, the lobbying game effectively ends with the agreement, since governments are left with no discretion and hence there is no scope for lobbying *ex post*; if, on the other hand, the agreement takes the form of a tariff ceiling, then lobbying may not end with the agreement. In this section we will argue that tariff ceilings are generally preferable to ETCs. The reason, as will become clear shortly, is that *keeping the lobbying game alive may reduce the very distortions created by lobbying.*

Let us start with an intuitive comparison between tariff ceilings and ETCs. Tariff ceilings may induce ex-post contributions, while ETCs do not. In our transferable-utility setting, ex-post contributions *per se* are a "wash" from the point of view of the joint surplus, but they nevertheless play an important role, because they affect the net returns to capital and hence its allocation, which is relevant for the joint surplus. Formally, c does not enter Ψ directly, but it affects Ψ through x. This effect on the allocation may be beneficial because there is an overinvestment problem in the M sector. We now explore this intuition more formally.

As a preliminary observation, note that in the two extreme cases of fixed capital ($z = 0$) and perfectly mobile capital ($z = 1$) celings are equivalent to ETCs: if $z = 1$, we know that the optimal ceiling is $\bar{t}^{A} = 0$, which is clearly equivalent to imposing an ETC at $t = 0$; and if $z = 0$, the optimal ceiling and the optimal ETC are both equal to $t^{\Psi}(\hat{x})$. But things may be different in the case of imperfect capital mobility ($0 < z < 1$), to which we turn next.

The analysis of ETCs is similar to that of ceilings, except that, since there are no ex-post contributions, the equal-returns curve for ETCs is simply defined by $p(t, x) = 1$, or $x - x^{ft} = t$. Notice that for $x < x^{W}$ the equal-returns curve for ETCs coincides with the equal-returns curve for ceilings, and for $x > x^W$ the former lies below the latter. Intuitively, tariff ceilings induce ex-post contributions and hence a higher price is needed to equalize returns across the two sectors. In general, given an exact commitment at level t, the equilibrium allocation is given by the line $x - x^{ft} = t$ truncated on the left at $(1-z)\hat{x}$ and on the right at $\hat{x} + z(1-\hat{x})$ (which is the level of x that obtains when a share z of the capital stock in the N sector moves to the M sector). In Figure 4, this truncated equal-returns curve for ETCs is labeled E_z^E (to keep the figure simple we draw this E_z^E curve only for $x \leq \hat{x}$; considering the region $x > \hat{x}$ is not essential for the analysis). The optimal ETC is the one that maximizes the objective $\Psi(t, x)$ in (3) subject to the constraint that (x, t) lie on curve E_z^E.

We can now compare graphically the problem of finding the optimal ceiling with that of finding the optimal ETC. Recall that the optimal ceiling maximizes $\Psi(\bar{t}, x)$ subject to the constraint that $(x, \bar{t})$ lie on the truncated equal-returns curve for ceilings, which in Figure 4 is labeled E_z^C. A key step is to draw the map of iso-Ψ curves in (t, x) space. It is direct to verify that each iso-Ψ curve is concave and has a peak on the $t^\psi(x)$ curve, as shown in Figure 4. Moreover, the value of Ψ increases as we move toward the left in the map of iso-Ψ curves.[23] It follows that the optimal ceiling is given by the point of tangency between an iso-Ψ curve and the E_z^C curve, while the optimal ETC is given by the point of tangency between an iso-ψ curve and the E_z^E curve. By graphical inspection, then, a ceiling can achieve a weakly higher level of Ψ than an ETC, because the E_z^C curve lies weakly to the left of curve E_z^E. Thus we can say that tariff ceilings perform at least as well as ETCs.

Next we establish that there exists a parameter region in which ceilings are *strictly* better than ETCs. Given the shape of the iso-Ψ curves, this is the case if the optimal ceiling lies strictly above the E_z^E curve, because in this case the outcome implemented by the optimal ceiling cannot be implemented by an ETC. If the $t^\psi(x)$ curve passes above point W (which is the case if and only if $a < \frac{3v-4}{2-v}$), it is easy to identify a range of z for which this is the case.

[23] This can be shown by recalling that $t^\psi(x)$ is the tariff level that maximizes Ψ for given x, and checking that (i) Ψ is jointly concave in (x, t), which implies that the iso-Ψ curves are concave; and (ii) in the relevant range $\Psi_x(t^\psi(x), x) < 0$, which implies that iso-Ψ curves more to the left are associated with higher values of Ψ.Also, for completeness, the statements made in the text require a qualification. In the case of ceilings we only needed to consider the case $x \leq \hat{x}$, and we defined the objective Ψ only for this case. In the case of ETCs, one has to consider also the possibility of entry in the M sector after the agreement $(x > \hat{x})$, in which case Ψ will include only the returns to capital currently in the M sector (i.e. the returns to the current lobby members), not capital that will enter the M sector after the agreement. A consequence of this observation is that, in the region $x > \hat{x}$, the $t^\psi(x)$ curve is constant and equal to $\hat{x}/a$. It is easy to establish, however, that the qualitative properties of the iso-Ψ curves higlighted above hold also in the region $x > \hat{x}$, and hence our arguments are not affected by considerations of entry.

Consider a strictly positive value of z such that $x_z > x^W$ and $t^\psi(x_z)$ lies above the E_z^C curve, as in Figure 4. By graphical inspection it is clear that for this value of z – and for an interval of z around it – the optimal ceiling lies above the E_z^E curve, and hence it is strictly superior to any ETC.[24]

The analysis thus confirms the intuitive argument made at the beginning of this section, but with a qualification. While it is true that *for a given tariff level* a ceiling induces a weakly lower level of x than an ETC, and hence it is correct to say that ceilings mitigate the overinvestment problem, the optimal allocation is actually the same in the two cases ($x = x_z$). Thus, a more precise statement is that ceilings are preferable to ETCs because they allow governments and lobbies to implement tariff levels closer to the "static" optimum $t^\psi(x)$ while still inducing maximal exit from the M sector. The following proposition records the main result of this section:

Proposition 6 *Tariff ceilings perform at least as well as exact tariff commitments. Moreover, if $a < \frac{3v-4}{2-v}$ there exists an intermediate interval of z for which tariff ceilings are* strictly *preferred to exact tariff commitments.*

Proposition 6 highlights that our model can help explain the use of tariff ceilings, which is pervasive in real trade agreements. Notice that this type of agreement is an incomplete contract, because it leaves some discretion to governments. The reason the optimal agreement may be incomplete (even though in our model there is none of the "usual" causes of contract incompleteness, e.g. nonverifiable information or contracting costs) is that the agreement cannot specify the contributions that the lobby will have to pay in the future.[25] If the agreement could specify both tariffs and contributions, a complete contract would be optimal. But since the contract cannot specify contributions, it may be optimal to leave the contract partially incomplete also in the other dimension, that is tariffs.[26]

[24] Note that the condition $a < \frac{3v-4}{2-v}$ is compatible with the condition that we assumed for the existence and uniqueness of the noncooperative equilibrium ($a > \frac{6v-7}{6(2-v)}$), because $\frac{3v-4}{2-v} > \frac{6v-7}{6(2-v)}$ for all relevant values of v; thus we have indeed identified a parameter region where ceilings are strictly preferable to ETCs.

[25] We believe the assumption that the agreement cannot specify future contributions is not only realistic, but can also be justified from a more fundamental perspective. To the extent that there is some capital mobility, today's lobby members may not be the same as tomorrow's. For this reason it is not very plausible that current lobby members would be able to commit future lobby members to pay a given amount of contributions. Enforcing a commitment of this kind seems more problematic than enforcing a commitment for future governments to respect a trade policy commitment.

[26] One might object that there is an alternative type of agreement that can accomplish the same objective as

We note that we have stacked the deck against tariff ceilings, by assuming that governments have zero bargaining power vis-à-vis domestic lobbies. If governments have some bargaining power, ex-post contributions will be higher, hence net returns to capital in sector M would be lower, and the overinvestment problem would be further mitigated.

A question might be raised as to whether the empirical evidence is consistent with the prediction that the applied tariff is equal to the ceiling level, or in GATT-WTO jargon, that there is no "binding overhang". In the GATT-WTO experience, developed countries tend to keep tariffs at the ceiling levels, while LDCs tend to apply tariffs well below the ceilings. Indeed, according to Bchir et al. (2005), the share of (non-agricultural) products with zero binding overhang in developed countries is 85%, while the respective share for developing countries is 35%. This is probably an indication that there are other important factors that play a role in reality but are not captured in our model, for example uncertainty and contracting imperfections (as in Bagwell and Staiger, 2005, or Horn, Maggi and Staiger, 2006). In this paper we do not claim to have a complete theory of tariff ceilings, but we make the more limited point that our model can *contribute* to explain the use of tariff ceilings.[27]

2.4 Ex-ante lobbying and relative bargaining powers

In the basic model we assumed that lobbies can influence the trade agreement in a similar way as they can influence ex-post trade policies, and we focused on the benchmark case in which governments have no bargaining power. In this section we discuss briefly how the results of the model would change if these assumptions were modified.

First we consider the role of ex-ante lobbying. The case in which lobbies can influence the shaping of the trade agreement is a natural benchmark to consider, but one can also think of situations in which lobbies are not involved in the negotiation of the trade agreement. For example, the institutions of a country might be such that trade negotiators are relatively

a tariff ceiling, namely one that specifies an exact tariff level and a tax on capital (or some other policy that lowers the net returns to capital) in the organized sector. But note that this type of agreement would constrain not just trade policies but also *domestic* policies, hence it is likely to involve higher contracting costs (i.e. costs of negotiating, writing and verifying/enforcing the agreement) relative to one that imposes only tariff ceilings. The intuitive appeal of tariff ceilings is that they can mitigate the overinvestment problem by leaving *more* – rather than less – discretion to governments.

[27] Furthermore, it is not clear that the two models mentioned above are fully consistent with the available evidence: if the only reason for the use of tariff ceilings were the presence of uncertainty and contracting imperfections, we should observe that, for a given product in a given country, the applied tariff is sometimes at the ceiling level and sometimes below it – but there is no evidence of this behavior in reality.

insulated from lobbying pressures. If there is no ex-ante lobbying, then the optimal agreement maximizes $U^G + U^{G^*}$ rather than $U^G + U^{G^*} + U^L$. How would the results change in this case?

With no ex-ante lobbying, the objective function can be written as

$$aW(\bar{t}, x) + C + aW^*(\bar{t}, x) = aW(\min\{t^W(x), \bar{t}\}, x) + aW^*(\bar{t}, x), \tag{4}$$

where we have used the fact that the Home government's reservation utility given ceiling $\bar{t}$ is given by $aW(\min\{t^W(x), \bar{t}\}, x)$. For a given z, the optimal tariff ceiling is the one that maximizes (4) subject to the constraint that $(x, \bar{t})$ lie on the truncated equal-returns curve, E_z^C. It is not hard to see that the solution of this problem is $\bar{t} = 0$ for all z.[28] Thus we can conclude that, in the absence of ex-ante lobbying, the optimal agreement entails free trade for any degree of capital mobility. It is important to note, however, that this extreme result depends on the extreme assumption that lobbies have all the bargaining power. As we argue next, this may no longer be the case if governments have bargaining power.

Next we discuss the role of the assumption that governments have no bargaining power vis-à-vis their domestic lobbies. This is a convenient assumption because in this case the government does not derive any rents from the political process, and hence it has a strong desire to foreclose domestic political pressures. If the government has some bargaining power, however, the domestic-commitment motive for a trade agreement is weaker. The question then is, how do our results change when we drop the assumption that the lobby has all the bargaining power?

The main change in results is that *ceteris paribus* trade liberalization tends to be less deep, and the parameter region for which the optimal agreement entails free trade is smaller. This is true both in the presence and in the absence of ex-ante lobbying. In particular, in the case of no ex-ante lobbying (as well as in the case of ex-ante lobbying with perfect capital mobility) there exists a parameter region where the trade agreement does not completely eliminate trade protection.

Intuitively, when governments have more bargaining power, the noncooperative equilibrium allocation is less distorted, because ex-post contributions are higher and hence net returns to capital in the M sector are lower, and as a consequence the domestic commitment motive for trade liberalization is less important. In addition to this effect, government barganing power

[28] To see this, notice that for any given x the objective is maximized at $\bar{t} = 0$. Since $W(0, x) + W^*(0, x)$ is decreasing in x for $x > x^{ft}$, it follows that the optimal ceiling is always $\bar{t} = 0$.

tends to reduce trade liberalization through another, more direct channel, which applies in the case of no ex-ante lobbying: when governments have bargaining power, they can extract rents from the ex-post lobbying game, therefore they are more reluctant to liberalize trade because doing so reduces those rents.[29]

To illustrate this in the simplest way, we consider the opposite case as in our baseline model, namely the case in which the government has all the bargaining power. For some x consider the government and lobby negotiating a tariff above $t^W(x)$. The rents obtained by the lobby would be given by:

$$x\left[p(t,x) - p(t^W(x),x)\right]$$

Since the government has full bargaining power, it captures all these rents in the form of contributions. Thus,

$$c(t,x) = \begin{cases} p(t,x) - p(t^W(x),x) \text{ for } t \geq t^W(x) \\ 0 \text{ for } t < t^W(x) \end{cases}$$

The net profit per unit of capital in sector M is then $p(t,x)$ for $t < t^W(x)$ and $p(t^W(x),x)$ for $t \geq t^W(x)$. This implies that the equal-returns (or ER) curve now becomes vertical at point (t^W, x^W), so that $x^{er}(t) = x^W$ for $t \geq t^W$, and that the long-run non-cooperative equilibrium - which, as before, is given by the intersection of $t^J(x)$ and $t^{er}(x)$ - is given by $\hat{x} = x^W$, $\hat{t} = t^J(x^W)$.

Consider now what happens when governments can sign a trade agreement. The first step is to write down the ex-ante joint surplus of the two governments and the lobby. Letting δ be an indicator variable that is equal to one if there is ex-ante lobbying and zero if not, we obtain:

$$\Psi(\bar{t},x) = \begin{cases} a[W(\bar{t},x) + W^*(\bar{t},x)] + xp(\bar{t},x) + [\cdot] \text{ for } \bar{t} \geq t^W(x) \\ a[W(\bar{t},x) + W^*(\bar{t},x)] + \delta xp(\bar{t},x) \text{ for } \bar{t} < t^W(x) \end{cases}$$

where the terms $[\cdot]$ do not depend on $\bar{t}$. Given $z \in [0,1]$, the optimal agreement in our model is the point that maximizes $\Psi(\bar{t},x)$ along the truncated equal-returns curve.[30]

[29] It is legitimate to ask whether our comparative-statics results are preserved when governments have bargaining power, and in particular whether it is still the case that (i) trade liberalization is deeper when capital is more mobile, and (ii) with high capital mobility, lower a implies deeper liberalization. It is easy to verify that result (i) holds even if governments have all the bargaining power. As for result (ii), it may or may not hold if governments have high bargaining power, depending on the configuration of parameters. But the result is still robust in an important sense. As the previous analysis makes clear, the domestic-commitment motive for a trade agreement is stronger when capital is more mobile and when governments have lower bargaining power. Thus the robust aspect of result (ii) is that, if the domestic-commitment motive for a trade agreement is strong enough, a lower level of a implies deeper trade liberalization.

[30] To be specific, this is the ER curve truncated at $x_z = x^W(1-z)$. If z is such that $(1-z)x^W \leq x^{ft}$, then in the relevant region this is just the ER curve; and in the case of fixed capital ($z = 0$), this is just the vertical line at $x = x^W$.

Let us first focus on the case in which there is ex-ante lobbying. In this case, letting $t^{\Psi}(x)$ denote the value of $\bar{t}$ that maximizes $\Psi(\bar{t}, x)$ given x, we have $t^{\Psi}(x) = a/x$. We need to consider two possibilities, depending on whether $t^{\Psi}(x)$ passes below or above point (t^W, x^W). If it passes below, then the analysis is basically the same as before (i.e., with the government having no bargaining power), and a result analogous to Proposition 4 obtains. In particular, when capital mobility is sufficiently high, the agreement entails free trade.

Consider now the case in which $t^{\Psi}(x)$ passes above the (t^W, x^W) point. It is simpler to start with the case of full capital mobility, $z = 1$. By definition of $t^{\Psi}(x)$, $\Psi(\bar{t}, x^W)$ increases as $\bar{t}$ falls from $t^J(x^W)$ to $t^{\Psi}(x^W)$ but then decreases as $\bar{t}$ continues to fall to t^W. On the other hand, we already know that $\Psi(\bar{t}, x)$ increases as we move along the ER curve from the (t^W, x^W) point towards free trade. Thus, there are two local maxima: $t^{\Psi}(x^W)$ and $t = 0$. Depending on parameters, either one may be the best agreement. In particular this depends on the height of the optimal terms-of-trade tariff t^W. If t^W is low (which is the case when trade volume is low, which in turn happens when v is relatively high), then $t^{\Psi}(x^W)$ is close to $t^J(x^W)$ and the (t^W, x^W) point is close to the free trade point. This implies that the decline in Ψ as $\bar{t}$ falls from $t^{\Psi}(x^W)$ to t^W exceeds the increase in welfare as we move from (t^W, x^W) to the free trade point, hence $t^{\Psi}(x^W)$ is better than free trade.

Now consider the more general case $z \in [0, 1]$. How does the optimal agreement vary as z increases from zero? Since $t^{\Psi}(x^W) > t^W$, then the optimal agreement remains at $t^{\Psi}(x^W)$ for all z if this point is better than free trade. If not, then there will be a sufficiently high level z at which the best agreement switches discontinuously from $t^{\Psi}(x^W)$ to $t^{er}(x^W(1-z))$, following this curve until the free trade point as z increases further.

Next we analyze the case in which there is no ex-ante lobbying. To make our points, it suffices to focus on the case of fixed capital ($z = 0$). As above, we have to distinguish between the cases in which the curve $t^{\Psi}(x)$ passes above or below the (t^W, x^W) point. If it passes below, then the optimal agreement entails free trade. To see this, note that (i) the point (t^W, x^W) dominates any other point (t, x^W) with $t \in (t^W, \hat{t}]$, and (ii) the point $(0, x^W)$ dominates the point (t^W, x^W), since at both points contributions are zero and $t = 0$ maximizes $W + W^*$ for any x. Finally, since $W(0, x) + W^*(0, x)$ increases as we move towards $x = x^{ft}$, then the best agreement entails $t = 0$ and x^{ft}.

Consider now the case in which the curve $t^{\Psi}(x)$ passes above (t^W, x^W). Now there are two candidates for an optimal agreement: $t = 0$ and $\bar{t} = t^{\Psi}(x^W)$. The basic trade-off is that, in the case of free trade, efficiency is maximized but there are no contributions, whereas in the case $\bar{t} = t^{\Psi}(x^W)$ efficiency is not maximal but the government gets some rents. The optimal agreement may be one or the other, depending on parameters.

In the more general case $z \in [0, 1]$, it is not hard to show that the optimal tariff cut is weakly increasing in z. Also note that, in the case of perfect capital mobility ($z = 1$) the cases of ex-ante lobbying and no ex-ante lobbying are equivalent, because even if the lobby can participate in the negotiation of the agreement, it will not influence the agreement, because it derives no benefits from protection in the long run.

To summarize the points of this section: (i) the presence of ex-ante lobbying tends to reduce the extent of trade liberalization, compared with the case of no ex-ante lobbying, and (ii) trade liberalization tends to be less deep when governments have bargaining power vis-à-vis their domestic lobbies.

3 The full dynamic model

In this section we consider a continuous-time specification of the model, where the agreement can determine the whole future path of the tariff ceiling. This analysis is important for two reasons. First, this will provide dynamic foundations for the reduced-form model considered in the previous section, and will indicate how the results of that section should be interpreted. Second, the full dynamic model allows us to address two sets of questions that cannot be adequately addressed within the two-period model: (i) Does the optimal agreement entail instantaneous liberalization, gradual liberalization, or a combination of the two? If trade liberalization has a gradual component, what is the nature of the optimal tariff path, and what determines the speed of liberalization? (ii) In the baseline model we assumed that the agreement comes as a surprise to investors; how do the model's predictions change if investors can foresee the agreement?

Consider the same model as above, but now suppose that time is continuous, denoted by s. We assume that at each point in time each capital-owner in sector M gets a chance to exit the sector with a (constant and exogenous) instantaneous probability z; the arrival of the opportunity to exit is independent across investors. On the other hand, to simplify the exposition and the derivation of the main results, we assume that entry into the M sector is free: that is, owners of capital in the N sector can freely move to the M sector if they wish. A more symmetric specification, where there is friction in capital mobility also from the N sector to the M sector, would deliver the same results, but the analysis would be more cumbersome.[31]

We assume that governments and capital-owners discount the future at a common rate $\rho > 0$. At each point in time the lobby (which is composed of the capital-owners currently in sector M) makes a take-it-or-leave-it offer (t, c) to the government. We focus on Markov equilibria, that is equilibria where players' strategies depend on the history only through the state variable x.[32]

Consider first the noncooperative equilibrium. It is direct to show that the steady state of this equilibrium entails $x = \hat{x}$, $t = \hat{t}$ and $c = c(\hat{t}, \hat{x})$, just as in the long-run equilibrium of the two-period model. Intuitively, if $x = \hat{x}$ and $t = \hat{t}$ the net returns to capital are equalized across sectors, hence there is no incentive for capital-owners to move; the lobby cannot do better than offering $\left(\hat{t}, c(\hat{t}, \hat{x})\right)$ to the government, and the government accepts this offer because it cannot do better on its own.

Next we consider the optimal trade agreement. We start by considering a scenario where the agreement comes as a surprise to investors, and assume that when the agreement opportunity arises (at time $s = 0$) the world is sitting at the steady state of the noncooperative equilibrium. In section 3.1 we will analyze the case in which the agreement is anticipated by investors.

We assume that the agreement is chosen to maximize the ex-ante joint surplus $U^G + U^{G^*} + U^L + U^{L^*}$, where U^j is interpreted as player j's payoff in PDV terms. The agreement determines a (fully enforceable) future path for the tariff ceiling, $\bar{t}(s)$. (We consider ceilings rather than ETCs because, just as in the previous section, ceilings perform at least as well as ETCs.) After the agreement has been signed, at each point in time the lobby makes a take-it-or-leave-it offer (t, c) to the government, of course taking into account that t is constrained by the agreement.

The first step of the analysis is to derive the equilibrium paths of x, t and c for a *given* path of the tariff ceiling $\bar{t}$. We will omit the time argument s whenever this does not cause confusion. First note that the equilibrium tariff given x is simply $t = \min\{\bar{t}, t^J(x)\}$. If $\bar{t} \leq t^J(x)$ the tariff ceiling is not redundant and hence the tariff is given by $\bar{t}$, otherwise the tariff is given by $t^J(x)$.

[31] The problem is that without free entry into the M sector, our way of modeling capital mobility would lead to a discontinuity in the law of motion of x. The rate of change dx/dt would go from $-zx$ when the value of capital in the N sector is higher than in the M sector, to $z(1-x)$ in the opposite situation. This makes the optimal control problem harder to solve, but it can be shown that the results hold also in this case.

[32] Our restriction to Markov equilibria rules out "reputational" equilibria, in which the domestic commitment problem could potentially be solved without the need for an international agreement if players are sufficiently patient. In the conclusion we discuss how our results are likely to extend to a setting where such reputational equilibria are allowed.

The associated contributions are given by $c(t, x)$, just as in the static model.

To characterize the equilibrium path for x, let V^M (V^N) be the value of a unit of capital in the M (N) sector. Since there is free entry into the M sector, then in equilibrium it must be that $V^M \leq V^N$. Moreover, the following no-arbitrage conditions must hold:

$$\rho V^M = z(V^N - V^M) + \dot{V}^M + p - c \tag{5}$$

$$\rho V^N = \dot{V}^N + 1 \tag{6}$$

To understand the first of these no-arbitrage conditions, note that the flow return to a unit of capital in the M sector (the RHS of (5)) is composed of three terms: the expected capital gain from moving to the N sector, $z(V^N - V^M)$; the change in the value of capital in the M sector, $\dot{V}^M$; and the instantaneous profits or "dividends," $p - c$. In equilibrium this flow return must be equal to the opportunity cost of holding an asset of value V^M, given by ρV^M. The second no-arbitrage condition (6) is similar, except that because of free capital mobility from N to M there cannot be capital gains from moving from N to M in equilibrium.

Combining the no-arbitrage equations (5) and (6) and letting $y \equiv V^M - V^N$, we obtain:

$$\dot{y} = (\rho + z)y - (p - c - 1) \tag{7}$$

Letting $g(t, x) \equiv p(t, x) - c(t, x) - 1$, integrating and imposing the condition $y(s) \to 0$ as $s \to \infty$,[33] we obtain:

$$y(s) = \int_s^\infty e^{-(\rho+z)(v-s)} g(t(v), x(v)) dv \leq 0 \text{ for all } s \tag{8}$$

It follows from the above discussion that, given a path for the maximum tariff $\bar{t}(s)$, the equilibrium conditions for $t(s), x(s)$ and $y(s)$ are the following:

$$t(s) = \min\{\bar{t}(s), t^J(x(s))\} \tag{i}$$

$$y(s) \text{ satisfies (8) and } y(s) \leq 0 \text{ for all } s \tag{ii}$$

$$\dot{x}(s) = -zx(s) \text{ if } y(s) < 0 \text{ and } \dot{x}(s) \geq -zx(s) \text{ if } y(s) = 0, \text{ with } x(0) = \hat{x} \tag{iii}$$

Condition $y(s) \leq 0$ in (ii) is a consequence of the assumption that there is free entry into the M sector. Condition (iii) simply states that if $y < 0$ then capital leaves the M sector as fast

[33] This is a condition that there be no "bubbles" in the asset market. We could replace this by the weaker condition that y converges to a finite value as $s \to \infty$.

as possible, whereas if $y = 0$ then any reallocation is an equilibrium as long as it satisfies the physical restriction that capital cannot leave the M sector faster than at rate $-zx$.

We can now derive the optimal agreement $\bar{t}(s)$. The objective function is

$$\int_0^{\infty} e^{-\rho s}\Psi(t(s), x(s))ds, \tag{9}$$

$$\text{where } \Psi(t,x) \equiv aW(t,x) + aW^*(t,x) + x(p(t,x) - 1) + \hat{x}$$

We say that a *plan* $(t(s), x(s), y(s))$ is *implementable* if there is an agreement $\bar{t}(s)$ such that $(t(s), x(s), y(s))$ is an equilibrium, i.e. satisfies conditions (i)-(iii). We will look for the plan that maximizes (9) in the set of implementable plans, and then we will identify the agreement $\bar{t}(s)$ that implements this plan. To turn this maximization into a more standard optimal control problem, we let $u = \dot{x}$ and note that any implementable plan must satisfy conditions (i) and (ii) together with the following "relaxed" condition:

$$u(s) + zx(s) \geq 0 \text{ for all } s \geq 0, \text{ and } x(0) = \hat{x} \tag{iii'}$$

Condition (8), (i),(ii) and (iii') are *necessary* for implementability. Our approach is to maximize the objective (9) subject to these necessary conditions for implementability, and then verify that the solution satisfies all implementability conditions. If this is the case, then we have found the optimal plan.

We need the problem to be concave, so that we can apply sufficiency conditions from optimal control theory. A simple sufficient condition for this is $a \geq \frac{v-1}{2-v}$; we will maintain this assumption for the remainder of the section.[34]

We are now ready to state the main result of this section. Let $x^z(s)$ denote the path of x obtained when capital exits the M sector as fast as possible until the free trade allocation is reached: $x^z(s) \equiv \max\{\hat{x}e^{-zs}, x^{ft}\}$. Also, as in the previous section, $t^{er}(x)$ is implicitly defined by $g(t,x) = 0$, and we let $t^{\Psi}(x) \equiv \arg\max_t \Psi(t,x)$.

Proposition 7 *The optimal agreement entails four phases:*
(i) an instantaneous drop in the tariff from $\hat{t}$ to $t^{\Psi}(\hat{x})$;
(ii) a first gradual liberalization phase in which $t(s) = t^{\Psi}(x^z(s))$, and $y(s) < 0$;
(iii) a second gradual liberalization phase in which $t(s) = t^{er}(x^z(s))$, and $y(s) = 0$;

[34] We note that the condition in the text is stronger than the one assumed in the static model (Proposition 1), since $\frac{v-1}{2-v} > \frac{6v-7}{6(2-v)}$ for all relevant values of v.

(iv) a steady state in which the tariff is zero.
The optimal path for the allocation is given by $x(s) = x^z(s)$ for all s.

This proposition states that the optimal trade agreement entails a discrete tariff cut at time zero, with the tariff dropping from $\hat{t}$ to $t^{\Psi}(\hat{x})$, which is then followed by gradual trade liberalization and exit of capital from the M sector. This gradual trade liberalization is characterized by two phases. In the first phase, the tariff is given by the optimal "static" tariff $t^{\Psi}(x)$ as a function of the evolving capital allocation, whereas in the second stage the tariff is the one that equalizes net returns across sectors (given the allocation $x^z(s)$). Note that in the first phase capital-owners in the M sector want to leave as fast as possible, since the returns to capital in that sector are lower than in the N sector ($y(s) < 0$); in the second phase capital-owners are indifferent as to where to locate their capital ($y(s) = 0$), but the government induces exit at the fastest possible rate. After a period of adjustment, the capital stock reaches the free trade allocation, and the tariff stays at zero thereafter.

An interesting implication of Proposition 7 is that the length of time it takes for the tariff to reach its steady-state value of zero is decreasing in z. Thus, if we define the speed of liberalization as the inverse of this length of time, we have the following corollary:

Corollary 1 *The speed of trade liberalization increases with the degree of capital mobility (z).*

Corollary 1 suggests an interesting prediction of the model: the tariff paths established in trade agreements should entail faster liberalization for sectors where exit can proceed at a faster pace. An interesting open question is whether this prediction is consistent with empirical observations.[35]

We can now come back to the question of how one should interpret the results of the reduced-form static model analyzed in section 2, and in particular the comparative-statics results concerning the degree of capital mobility (z) and the political parameter a (see Proposition 4). Consider first the result that trade liberalization is deeper when capital is more mobile.

[35] A natural question that might arise at this point concerns the empirical plausibility of the prediction that the steady-state level of the tariff is zero. While this seems fairly consistent with reality in the case of regional trade agreements, such as NAFTA or Mercosur, it is less consistent with the observed history of the GATT-WTO. However it would not be hard to enrich the model in such a way that the steady-state level of the tariff is strictly positive. For example, the result relies on the assumption that all the capital in the M sector can eventually be reallocated to the N sector. But if part of the capital cannot be reallocated even in the long run, then the steady-state tariff level may not be zero. Also, if governments have bargaining power vis-à-vis domestic lobbies, the trade agreement may not bring tariffs to zero (see section 2.4).

In light of the above analysis, one interpretation of this result is that the tariff cut evaluated between the time of the agreement and any given point in time T (i.e. $\hat{t} - t(T)$) is larger when capital is more mobile; or alternatively, as Corollary 1 states, that the tariff falls faster when capital is more mobile. Similarly, the result concerning the impact of a can be interpreted as saying that the tariff cut evaluated between the time of the agreement and any given point in time increases with a when z is sufficiently small but decreases with a when z is sufficiently high.[36]

As in the previous section, we want to understand the role of the TOT motive and of the domestic-commitment motive in the determination of the optimal trade-liberalization path. Following a similar approach as in the previous section, we consider the hypothetical case in which, starting from the non-cooperative equilibrium $(\widehat{t}, \widehat{x})$, each government gets a chance to unilaterally commit to a future path for the tariff ceiling, but without a trade agreement. In this case, the objective function that the government maximizes is the same as in (9), except that it does not include foreign welfare. Let us denote the resulting path as $\bar{t}^{DC}(s)$. We can think of $\widehat{t} - t^{DC}(s)$ as the domestic commitment component of the trade liberalization path, with the remainder, that is $\bar{t}^{DC}(s) - t(s)$, being the TOT component.

It is straightforward to show that the optimal domestic-commitment path entails $t^{DC}(s) = t^{er}(x^z(s))$ until the point (t^W, x^W) is reached. One direct implication is that trade liberalization associated with the domestic commitment motive takes place gradually, with no discrete tariff reduction. The reason for this is that the government can achieve the desired reallocation of capital towards the N sector without any reduction in returns to capital in the M sector by following the path $t^{er}(x^z(s))$ from $(\widehat{t}, \widehat{x})$ to (t^W, x^W). This implies that the discrete tariff drop that follows the signing of the trade agreement (from $\widehat{t}$ to $t^{\Psi}(\hat{x})$) is entirely associated with the TOT motive, while the domestic-commitment motive is reflected in the gradual component of trade liberalization. More generally, this suggests that the gradual component of trade liberalization should be more important, relative to the instantaneous component, when the domestic-commitment motive is more important relative to the TOT motive.[37]

[36] To see this notice that, if z is close to zero, for any fixed $T > 0$ the tariff cut $\hat{t} - t(T)$ is close to $\hat{t} - t^{\psi}(\hat{x})$, which is increasing in a; and if z is close to one, $\hat{t} - t(T)$ is close to $\hat{t}$, which is decreasing in a.

[37] An interesting question is whether the TOT component of the trade agreement is gradual or not. The answer is that it depends. Recall that the TOT motive is responsible for (i) an instantaneous tariff drop from $\widehat{t}$ to $t^{\Psi}(\hat{x})$, and (ii) a reduction in the steady-state tariff from t^W to zero. Thus, if $\widehat{t} - t^{\Psi}(\hat{x}) < t^W$, we can say that the TOT component of the tariff cut is partly instantaneous and partly gradual, but if $\widehat{t} - t^{\Psi}(\hat{x}) > t^W$, then the TOT component of the tariff cut *overshoots* its steady-state level, hence we can say that this component is "anti-gradual". It is easy to find parameter values for which each of the two cases described above ($\widehat{t} - t^{\Psi}(\hat{x})$

3.1 Anticipated agreement

We have assumed so far that the trade agreement arrives as a complete surprise to capital-owners. Of course this is not very realistic: in practice, trade agreements never come as a surprise to investors, if nothing else because it takes time to negotiate, ratify and execute a trade agreement. In this section we explore how the model's predictions change if the agreement is partially or fully anticipated by capital-owners.

Formally, we assume that at time s_0 investors learn that at time s_1 a trade agreement will take effect. We allow for an arbitrary lag between s_0 and s_1. If $s_1 - s_0$ is sufficiently high that there is no capital loss to owners of capital in the M sector at time s_0, we can interpret this as the case in which the agreement is fully anticipated.

Consistently with the previous section, we assume that the agreement maximizes the joint surplus of governments and lobbies in PDV terms at time s_1. We also assume that governments have no bargaining power, and more specifically, contributions are such that the governments' joint payoff ($U^G + U^{G^*}$) is the same as they would obtain in the absence of lobbying.

Let us first examine the problem at an intuitive level. We can reason by backward induction. At time s_1 the situation is similar to the case of surprise agreement, except that the level of x may be different from $\hat{x}$. In particular, letting $x_1 \equiv x(s_1)$, the agreed-upon path for the tariff ceiling is the same as in Proposition 7 except that the initial level of x is given by x_1 rather than $\hat{x}$. Formally, if $x^z(s; x_1) \equiv \max\{x_1 e^{-z(s-s_1)}, x^{ft}\}$ denotes the path of maximal exit starting from x_1 until the free-trade level x^{ft} is reached, the path of the tariff ceiling is $\bar{t}(s; x_1) \equiv \min\{t^\Psi(x^z(s; x_1)), t^{er}(x^z(s; x_1))\}$.

Before moving back to the pre-agreement phase, we need to determine the "ex-ante" contributions paid by the lobby at s_1. As we remarked above, these contributions must give the two governments exactly the joint payoff that they would get in the absence of ex-ante lobbying. It is not hard to show that the optimal agreement in the absence of ex-ante lobbying entails free trade at all times (as in Section 2.4). As a consequence, ex-ante contributions at s_1 are zero for $x_1 = x^{ft}$ and are increasing in x_1.

We can now examine what happens before time s_1. As a preliminary observation, notice that for all $s < s_1$ the tariff must be equal to $t^J(x)$. Consider first the case in which $s_1 - s_0$ is small, so that the agreement is announced with a short notice. In this case, at time s_0 the

higher or lower than t^W) obtains.

value of capital in the M sector drops below that in the N sector ($y < 0$), both because of the ex-ante contributions that will be paid at s_1, and because for a period after s_1 the tariff will be below $t^{er}(x)$. In this case, then, at time s_0 capital will start exiting the M sector at maximum speed, and it will continue to do so after the agreement is signed, until it reaches the free-trade allocation. Note that in the period from s_0 to s_1 capital-owners in the M sector derive some rents ($p - c > 1$) because $t^J(x)$ is higher than $t^{er}(x)$ for $x < \hat{x}$, but since $s_1 - s_0$ is small this cannot compensate the losses caused by the agreement at s_1 and beyond; this confirms that investors indeed exit sector M between s_0 and s_1.

Now consider increasing $s_1 - s_0$. Then the capital loss in the M sector generated by the announcement of the future agreement decreases. This happens because of discounting, because there is more time to adjust before the agreement takes place (so that x_1 gets closer to x^{ft}) and because, as mentioned above, the pre-agreement period entails $p - c > 1$. If $s_1 - s_0$ is sufficiently high, the announcement of the future agreement generates no capital loss in the M sector, hence in this case there is no immediate exit at s_0. Capital will initially stay where it is, and at some point in time between s_0 and s_1 it will start exiting. Interestingly, the process of capital reallocation will *not* be completed by time s_1, but will continue for some time thereafter, even though the agreement is fully anticipated. Intuitively, it cannot be an equilibrium for x to reach x^{ft} on or before s_1: if this were the case, ex-ante contributions would be zero (as we argued above), and in the pre-agreement period net returns would be higher in the M sector, hence capital would *enter* the M sector before the agreement is signed.

The following proposition confirms the intuition we just developed:

Proposition 8 *An equilibrium exists and must have the following properties:*
(i) Between s_0 and some point in time $\hat{s}$, there is a stationary phase with $x = \hat{x}$ and $t = \hat{t}$. This time interval $(s_0, \hat{s})$ is non-empty if and only if the agreement is anticipated sufficiently in advance, i.e. $s_1 - s_0$ is sufficiently large;
(ii) Between $\hat{s}$ and s_1, the tariff is $t(s) = t^J(x(s))$, and capital exits the M sector at maximal speed, but does not reach level x^{ft};
(iii) From time s_1 on, the agreement comes into effect, the tariff ceiling follows the path $\bar{t}(s, x_1)$, and capital continues to exit at maximal speed until it reaches x^{ft}.

The results of Proposition 8 have a couple of important implications. First, if the agreement is anticipated there is *some* reallocation of capital in advance of the agreement; this means that part of the overinvestment problem is solved before the agreement actually takes place. Second, even if the agreement is fully anticipated, the allocation is still distorted at the time of the agreement, and hence the investors' anticipation of the agreement is not by itself sufficient to solve the domestic-commitment problem. Thus, the domestic-commitment problem is solved in part by the (credible) announcement of an agreement, and in part by the signing of the agreement itself. Finally, we note that the agreement itself is essentially the same as the one characterized in Proposition 7, except that the level of capital in sector M at the time of the agreement is lower than $\widehat{x}$; thus we can say that the qualitative results of the model hold also in the case of fully anticipated agreement.

4 Conclusion

In this concluding section we briefly discuss three possible extensions of the model.

First, we assumed that international agreements are perfectly enforceable, while there are no domestic commitment mechanisms. An alternative approach would be to assume that there are no exogenous enforcement mechanisms of any kind, so that both domestic and international agreements must be self-enforced through "punishment" strategies in a repeated game. The question that arises in this case is the following: if domestic punishments are not enough to solve the domestic commitment problem, is it the case that international punishments can help governments live up to their domestic commitments?

A more precise way to formulate the above question is the following. Consider an infinitely repeated version of our model, and compare two punishment strategies: (i) a purely domestic punishment strategy where, if a country's tariff deviates from its equilibrium level, the capital allocation and the tariff in that country revert to their long-run noncooperative levels; (ii) an international punishment strategy where, if a country's tariff deviates from its equilibrium level, the capital allocation and the tariff in *both* countries revert to their long-run noncooperative levels. (One could consider more severe international punishment strategies, for example a reversion to autarky; this would only strengthen the argument we are making here.) Suppose the optimal international agreement with perfect enforcement entails tariff t^A, and the optimal domestic-commitment tariff in the absence of international agreements is $t^{DC} > t^A$. Now suppose that a purely domestic punishment strategy is not enough to sustain t^{DC}, so that the domestic commitment problem remains partially unsolved. We then ask: can an international punishment strategy sustain t^A? If this is the case, then we can say that the international punishment strategy helps governments fully solve their domestic commitment problems. We conjecture that there will be a region of parameters for which this is the case. However, a rigorous examination of self-enforcing trade agreements will have to await further research.

Second, in our model a government cannot receive contributions from foreign capital-owners. How would foreign lobbying affect the trade agreement? A thorough examination of this question would require changing the production structure of the model, because we assumed that foreign capital is fixed. Since the interaction between lobbying and capital mobility is central to our arguments, we would need to modify the model so that foreign capital can move across sectors.[38] But one can reasonably speculate on what results would emerge in a model of this type. Suppose that the government's objective is given by $aW + C_d + bC_f$, where C_d and C_f are domestic and foreign contributions (as in Gawande, Krishna and Robbins, 2006); and the parameter $b > 0$ captures the government's valuation of foreign contributions relative to domestic contributions.

Consider our basic two-period model, and focus first on the benchmark case of fixed capital. In this case, since foreign exporters benefit from reducing tariffs, the presence of foreign lobbying will reduce both the noncooperative tariff level (t^J) and the tariff level set by the agreement (t^ψ), and if b is sufficiently high these tariff levels will turn into import subsidies. Note that these implications of foreign lobbying are not specific to our model, and apply also to the "standard" models of trade agreements a' la Grossman-Helpman (1995). Next consider the opposite benchmark case of perfect capital mobility. If entry fully dissipates rents in the foreign country just as in the domestic country, then the optimal agreement would entail free trade, by the same logic as in our basic model. Thus a reasonable conjecture is that introducing foreign lobbying in our model would decrease the tariff level set by the agreement, but this effect would be smaller the more mobile is capital.

Third, our model has the feature that trade protection in the M sector does not affect the returns to capital in the long run (when capital is perfectly mobile). This is due to the fact that the returns to capital in the N sector are independent of the capital allocation. An important consequence of this feature is that the lobby in the M sector attaches no value to trade protection in the long run. It is natural to ask how our results would extend to environments

[38] This would not be a minor change because, for returns to capital to be equalized across sectors in the long run (in both countries), we would need to introduce some curvature in the model. If the production structure is linear as in the current model, all capital will concentrate in one sector, at least in one country.

where Stolper-Samuelson effects are present, so that protection in the M sector would affect the returns to capital even in the long run. This would be the case, for example, if the N sector uses capital and labor. Even if this is the case, however, the impact of trade protection on the returns to capital will be stronger in the short run than in the long run, and this is the main driving force behind our results. Also note that, if sector M is small relative to the whole economy, trade protection in this sector will have only a small impact on the economy-wide returns to capital, hence in this case our results would hold with little change.[39] Finally, the assumption that trade protection in the M sector does not affect the long-run return to capital could be justified also from a trade-and-growth perspective. Consider an environment in which capital can be accumulated over time: even if increasing the domestic price of good M had a substantial impact on the economy-wide returns to capital, in the long run this would trigger a change in investment that would bring back the return to capital to its steady-state level, which is pinned down by the discount rate.

[39] How should one think about the size of sector M? One could simply think of sector M as a narrowly defined good, in which case the sector would be small in relation to the whole economy almost by definition. Alternatively, and perhaps more accurately, one could think of sector M as a subset of goods whose producers are organized to lobby together for protection. Even in this case, however, given the difficulties associated with collective action among different producers, it is reasonable to assume that sector M is a relatively narrow subset of goods. We also note that in the existing empirical literature on the political economy of trade policy it is commonly assumed that the relevant level of aggregation for organized lobbies is the 3-digit SIC level; a sector defined at this level is fairly small relative to the whole economy.

Appendix

Proof of Proposition 1:

We start by deriving $\hat{x}$. Plugging $t^J(x)$ into the equal-returns condition and solving in x we find the unique solution:

$$\hat{x} = 2a\left(\frac{3v-4}{4a-1}\right)$$

Note that the condition on a assumed in the Proposition implies that $a > 1/4$ and hence the denominator of the previous expression for $\hat{x}$ is positive, and also the numerator is positive given that we have assumed that $v > 3/2$. Also, one can show that the condition assumed in the proposition ensures $\hat{x} < 1$.

The equilibrium tariff is given by

$$\hat{t} = t^J(\hat{x}) = \frac{1}{3}[1 + \frac{4-2a}{4a-1}(3v-4)]$$

Differentiating with respect to a,

$$\frac{d\hat{t}}{da} = -\frac{7}{3} \cdot \frac{3v-4}{(4a-1)^2} < 0$$

which shows the last part of the claim.

We need to show that imports are positive at the equilibrium we just found. This requires $1-\hat{x} > \hat{t}$. Plugging in the values for $\hat{x}$ and $\hat{t}$ we get $a > (6v-7)/6(2-v)$, which is the condition assumed in the proposition.

For future reference we also show that the $x^{er}(x)$ curve is upward (downward) sloping below (above) the $t^J(x)$ curve. Differentiation shows

$$\frac{dx^{er}(t)}{dt} = \frac{1/2 - (a/4x)(3t^{er}(x) - \Delta x)}{1/2 + (a/12x)(3t^{er}(x) - \Delta x) - c(t^{er}(x), x)/x} \tag{10}$$

The condition on a in the proposition implies that $a > 1/3$, which in turn implies that the denominator is positive. On the other hand, it is easy to show that the numerator is positive if $t < t^J(x)$ and negative if $t > t^J(x)$. **Q.E.D.**

Proof of Proposition 2: See the proof of the more general Proposition 4**.**

Proof of Proposition 3: See the proof of the more general Proposition 5.

Proof of Proposition 4: The first step is to write a convenient expression for the objective $\Psi(\bar{t}, x^{er}(\bar{t}))$. Adding and subtracting total contributions $C = xc$ in expression (3), and recalling that the net rate of return to capital must be equal to one, we can write

$$\Psi(\bar{t}, x^{er}(\bar{t})) = aW(\bar{t}, x^{er}(\bar{t})) + C(\bar{t}, x^{er}(\bar{t})) + aW^*(\bar{t}, x^{er}(\bar{t})) + \widehat{x}$$

If $\bar{t} \geq t^W(x^{er}(\bar{t}))$, then there are positive contributions and

$$aW(\bar{t}, x^{er}(\bar{t})) + C(\bar{t}, x^{er}(\bar{t})) = aW(t^W(x^{er}(\bar{t})), x^{er}(\bar{t}))$$

On the other hand, if $\bar{t} < t^W(x^{er}(\bar{t}))$, then contributions are zero. Noting that $\bar{t} \geq t^W(x^{er}(\bar{t}))$ if and only if $\bar{t} \geq t^W$, we can write

$$\Psi(\bar{t}, x^{er}(\bar{t})) = \begin{cases} aW(t^W(x^{er}(\bar{t})), x^{er}(\bar{t})) + aW^*(\bar{t}, x^{er}(\bar{t})) + \widehat{x} \text{ if } \bar{t} \geq t^W \\ aW(\bar{t}, x^{er}(\bar{t})) + aW^*(\bar{t}, x^{er}(\bar{t})) + \widehat{x} \text{ if } \bar{t} < t^W \end{cases} \tag{11}$$

The next step is to show that $\Psi(\bar{t}, x^{er}(\bar{t}))$ in (11) is decreasing in $\bar{t}$. If $\bar{t} < t^W$, there are no contributions and $x^{er}(\bar{t})$ is a line with slope one, so the claim is easy to verify. Let us focus on the less straightforward case $\bar{t} \geq t^W$. We want to show that $F(\bar{t}) \equiv W(t^W(x^{er}(\bar{t})), x^{er}(\bar{t})) + W^*(\bar{t}, x^{er}(\bar{t}))$ is decreasing in $\bar{t}$ for $\bar{t} > t^W$. Applying the Envelope Theorem, then:

$$F'(\bar{t}) = W_x(t^W(x^{er}(\bar{t})), x^{er}(\bar{t}))dx^{er}(\bar{t})/d\bar{t} + W_t^*(\bar{t}, x^{er}(\bar{t})) + W_x^*(\bar{t}, x^{er}(\bar{t}))dx^{er}(\bar{t})/d\bar{t}$$

Since $W_t^* < 0$ and $dx^{er}(\bar{t})/dt > 0$, it suffices to show that $W_x(t^W(x^{er}(\bar{t})), x^{er}(\bar{t}))$ and $W_x^*(\bar{t}, x^{er}(\bar{t}))$ are both negative. The second part is obvious given that the only effect of x on W^* is through terms of trade, and an increase in x worsens Foreign's terms of trade. To show the first part, note that it is equivalent to $W_x(t^W(x), x) < 0$ for $x > x^W$. Some simple algebra reveals that:

$$W_x(t^W(x), x) = p(t^W(x), x) - 1 \tag{12}$$

Given the definition of x^W (i.e., $p(t^W(x^W), x^W) = 1$), the claim follows directly.

Now notice that, if $t^\Psi(x_z) < t^{er}(x_z)$, by definition of $t^\Psi(x_z)$ the point $(t^\Psi(x_z), x_z)$ is superior to the point $(t^{er}(x_z), x_z)$, which in turn is superior to all the other points on the curve $\widetilde{x}^{er}(\bar{t})$ for $\bar{t} > t^{er}(x_z)$. On the other hand, if $t^{er}(x_z) < t^\Psi(x_z)$ then all points on the vertical segment of curve $\widetilde{x}^{er}(\bar{t})$ are dominated by the point $(t^{er}(x_z), x_z)$.[40] Hence, in this case, the optimal tariff binding is simply $t^{er}(x_z)$. Of course this is true only as long as $x_z \geq x^{ft}$, otherwise the optimum is free trade.

[40] This follows from the fact that (as can be easily verified) Ψ is concave in t.

To prove that $\hat{t} - \bar{t}^A$ is (weakly) increasing in z, note that (a) $t^{er}(x)$ is increasing in the relevant range; and (b) $t^{\Psi}(x) = x/a$ is increasing in x. As a consequence, $\min(t^{er}(x_z), t^{\Psi}(x_z))$ is increasing in x_z, and hence decreasing in z, which implies the claim.

Point (iii) follows from the observations we made previously, that (a) $\hat{t} - \bar{t}^A$ is increasing in a for $z = 0$ and (b) $\hat{t} - \bar{t}^A$ is decreasing in a for $z = 1$. Since the problem is continuous in z, the claim follows. **Q.E.D.**

Proof of Proposition 5: Consider first the case $z = 1$. Following a similar logic as in the proof of Proposition 4, the objective can be written as

$$G(\bar{t}) \equiv J(\bar{t}, x^{er}(\bar{t})) = \begin{cases} aW(t^W(x^{er}(\bar{t})), x^{er}(\bar{t})) + \widehat{x} \text{ if } \bar{t} \geq t^W \\ aW(\bar{t}, x^{er}(\bar{t})) + \widehat{x} \text{ if } \bar{t} < t^W \end{cases} \tag{13}$$

Consider first the case $\bar{t} < t^W$. Differentiation yields:

$$G'(\bar{t}) = aW_t(\bar{t}, x^{er}(\bar{t})) + aW_x(\bar{t}, x^{er}(\bar{t}))$$

where we have used the fact that $dx^{er}/d\bar{t} = 1$ for $\bar{t} < t^W$. Note that $\bar{t} < t^W$ implies $\bar{t} < t^W(x^{er}(\bar{t}))$ and hence $W_t(\bar{t}, x^{er}(\bar{t})) > 0$, whereas $W_t(t^W, x^{er}(t^W)) = 0$. Turning to the second term above, differentiation yields

$$W_x(\bar{t}, x^{er}(\bar{t})) = (1/2)(1 - (v - 1) - 2\bar{t})$$

It is easy to show that this is equal to zero for $\bar{t} = t^W$, and hence is positive for $\bar{t} < t^W$. Thus, $G'(\bar{t}) > 0$ for $\bar{t} < t^W$. The previous arguments establish also that $G'(t^W) = 0$.

Now consider the case $\bar{t} > t^W$. In this case, differentiation yields (using the Envelope Theorem):

$$G'(\bar{t}) = [aW_x(t^W(x^{er}(\bar{t})), x^{er}(\bar{t}))]dx^{er}/d\bar{t}$$

We have already established that $dx^{er}/d\bar{t} > 0$ (see proof of Proposition 1) and that $W_x(t^W(x^{er}(\bar{t})), x^{er}(\bar{t})) < 0$ for $\bar{t} > t^W$ (see proof of Proposition 4). Therefore $G'(\bar{t}) < 0$ for $\bar{t} > t^W$, and the result follows immediately.

Next consider the case $z < 1$. Noting that $J(\bar{t}, x)$ is maximized by $t^J(x)$ and applying the same logic as in the proof of Proposition 4, one can show that

$$\bar{t}^{DC}(z) = \begin{cases} \min(t^{er}(x_z), t^J(x_z)) \text{ for } x_z \geq x^W \\ t^W \text{ for } x_z < x^W \end{cases}$$

But $\min(t^{er}(x_z), t^J(x_z)) = t^{er}(x_z)$, hence the claim. **Q.E.D.**

Proof of Proposition 7: We start by noting that we can set $t = \bar{t}$ without loss of generality. Using $\beta(s)$ as the Kuhn-Tucker multiplier of the constraint on the control, $u(s) + zx(s) \geq 0$, and $\phi(s)$ as the multiplier function of the pure state constraint, $y(s) \leq 0$, then we have the following Hamiltonian:

$$H = e^{-\rho s}\Psi(t,x) + \beta\left[u + zx\right] - \phi y + \lambda_x u + \lambda_y[(\rho + z)y - g(t,x)]$$

Necessary conditions for optimality are $H_t = H_u = 0$, the Euler equations $\dot{\lambda}_x = -H_x$ and $\dot{\lambda}_y = -H_y$, the constraints $u + zx \geq 0$, $y \leq 0$, and the complementary slackness (CS) conditions: $\beta \geq 0$, $\beta(u + zx) = 0$, and $\phi \geq 0$, $\phi y = 0$.

The condition $H_t = 0$ implies:

$$e^{-\rho s}\Psi_t - \lambda_y g_t = 0 \tag{14}$$

while $H_u = 0$ implies $\beta + \lambda_x = 0$, or

$$\beta = -\lambda_x \tag{15}$$

The Euler equation $\dot{\lambda}_x = -H_x$ yields:

$$\dot{\lambda}_x = -e^{-\rho s}\Psi_x - \beta z + \lambda_y g_x \tag{16}$$

while $\dot{\lambda}_y = -H_y$ yields:

$$\dot{\lambda}_y = \phi - \lambda_y(\rho + z) \tag{17}$$

Our methodology is to guess that the solution is the one stated in the Proposition and verify that it satisfies necessary and sufficient conditions for an optimum. Our conjectured solution entails three phases: the first phase starts at $s = 0$ and entails $t = t^{\Psi}(x^z(s))$; the second phase starts at $s = \widetilde{s}$, where $\widetilde{s}$ is defined implicitly by $t^{\Psi}(x^z(\tilde{s})) = t^{er}(x^z(\tilde{s}))$, and entails $t = t^{er}(x^z(s))$; the third stage starts at $s = s^{ft}$, where s^{ft} is defined as the time at which $e^{-zs}\hat{x}$ reaches x^{ft}, and entails a steady state where $t = 0$ and $x = x^{ft}$.

To check this conjecture, we move backwards, checking first that free trade can be a steady state. From (14), using $W_t + W_t^* = 0$ at $t = 0$, and noting that $g_t = 1$ (since there are no contributions in this case), then we get

$$\lambda_y(s) = e^{-\rho s}x^{ft} \tag{18}$$

This implies $\dot{\lambda}_y = -\rho\lambda_y$. Plugging in (17) yields $\phi(s) = z\lambda_y(s)$. Hence, using (18) we get:

$$\phi(s) = e^{-\rho s}zx^{ft} \tag{19}$$

Note that this clearly satisfies the condition $\phi(s) \geq 0$.

Now, since at the conjectured steady state we have $u = 0$, then $u + zx = zx > 0$, and hence the CS conditions imply $\beta = 0$. Equation (15) then implies that

$$\lambda_x(s) = 0 \tag{20}$$

Plugging this and $\dot{\lambda}_x = 0$ into the Euler equation (16) yields $e^{-\rho s}\Psi_x = \lambda_y g_x$, which is satisfied because at free trade $\Psi_x = x^{ft} g_x$ (since $W_x + W_x^* = 0$ at free trade) and (18) implies $\lambda_y g_x = e^{-\rho s} x^{ft} g_x$.

We can now move backwards to the second phase, $s \in [\widetilde{s}, s^{ft}]$. We will solve for λ_x and λ_y and check that β, ϕ are positive (as required by the CS conditions). Condition (14) can be used to solve for λ_y:

$$\lambda_y = e^{-\rho s}\Psi_t / g_t \tag{21}$$

Plugging this result and $\beta = -\lambda_x$ in (16), and using $dt^{er}/dx = -g_x/g_t$, yields:

$$\dot{\lambda}_x = \lambda_x z - e^{-\rho s}\frac{d\Psi}{dx} \mid_{g=0}$$

Now, since $e^{-\rho s}\frac{d\Psi}{dx} \mid_{g=0}$ evaluated at $x = x^z(s)$ is merely a function of time, we can denote it as $\mu(s)$, and hence we have a differential equation, which can be solved imposing $\lambda_x(s^{ft}) = 0$. This yields:

$$\lambda_x(s) = \int_s^{s^{ft}} \mu(v) e^{-z(v-s)} dv \tag{22}$$

In the proof of Proposition 4 we established that $\Psi(\bar{t}, x^{er}(\bar{t}))$ is decreasing in $\bar{t}$. This implies that $\frac{d\Psi}{dx} \mid_{g=0} < 0$, which in turn implies $\lambda_x(s) < 0$, and hence $\beta(s) > 0$.

The only condition left to check for the second phase is $\phi(s) \geq 0$ for $s \in [\widetilde{s}, s^{ft}]$. From the Euler equation (17), this is true if

$$\phi(s) = \lambda_y'(s) + (\rho + z)\lambda_y(s) \geq 0$$

Letting $t^{er}(s) \equiv t^{er}(x^z(s))$ and $f(s) \equiv \frac{\Psi_t(t^{er}(s), x^z(s))}{g_t(t^{er}(s), x^z(s))}$, and using (21), we have

$$\phi(s) = z\lambda_y(s) + e^{-\rho s} f'(s) \tag{23}$$

We need to consider two cases, corresponding to $x < x^W$ and $x \geq x^W$. If $x < x^W$ then $\bar{t}(x) = t^{er}(x) < t^W(x)$, so there are no contributions. Differentiation shows that $f(s) = x^z(s) - at^{er}(s)$, which given $t^{er}(x) < t^{\Psi}(x) = x/a$ is positive, and hence $\lambda_y > 0$. Using $dt^{er}/dx = 1$, we find:

$$f'(s) = (a-1)zx^z(s)$$

The condition $a \geq \frac{v-1}{2-v}$ together with $v > 3/2$ implies $a > 1$, so we conclude that $\phi(s) \geq 0$ for $s \in [\widetilde{s}, s^{ft}]$ such that $x^z(s) < x^W$.

Now consider the second case, with $x \geq x^W$. Note that equation (23) implies $\phi(s) = e^{-\rho s}(zf(s) + f'(s))$. Differentiation then shows that

$$\phi(s) = \frac{e^{-\rho s}z}{g_t}\left(\Psi_t - x\left(\Psi_{tt}\frac{dt^{er}}{dx} + \Psi_{tx} - \frac{\Psi_t}{g_t}(g_{tt}\frac{dt^{er}}{dx} + g_{tx})\right)\right)$$

Plugging in $t = t^{er}(x)$ and after (a lot of) simplification we get

$$\phi(s) = \frac{e^{-\rho s}za}{6g_t}\frac{a(x(8-3v) - (3v-4)) - 3x(v-1)}{2a(3v-4) - x(4a-1)}$$

Since $x < \hat{x} = \frac{2a(3v-4)}{4a-1}$, the denominator is positive. As for the numerator, since $x \geq x^W = \frac{3v-4}{2}$, then $x(8-3v) - (3v-4) \geq 0$. Then, using $a \geq \frac{v-1}{2-v}$, we get

$$\begin{aligned} a(x(8-3v) - (3v-4)) - 3x(v-1) &\geq \frac{v-1}{2-v}(x(8-3v) - (3v-4) - 3x(2-v)) \\ &= \frac{v-1}{2-v}(2x - (3v-4)) \geq 0 \end{aligned}$$

where the last inequality comes again from $x \geq x^W$. Therefore, $\phi(s) \geq 0$ for all $s \in [\widetilde{s}, s^{ft}]$.

Moving now to the first phase (i.e., $s < \widetilde{s}$), our conjecture $y < 0$ implies by the CS conditions that

$$\phi(s) = 0 \tag{24}$$

Moreover, $t(s) = t^{\Psi}(x^z(s))$ implies $\Psi_t = 0$, and hence from (14) we get $\lambda_y = 0$ and hence $\lambda_y(\widetilde{s}) = 0$. The second Euler equation (17) is trivially satisfied with

$$\lambda_y(s) = 0 \tag{25}$$

and $\phi = 0$. To check the first Euler equation (16) we use $\beta = -\lambda_x$ and $\lambda_y = 0$ to obtain:

$$\dot{\lambda}_x = -e^{-\rho s}\Psi_x + \lambda_x z$$

Solving the above differential equation yields:

$$\lambda_x(s) = \lambda_x(\widetilde{s})e^{-z(\widetilde{s}-s)} + \int_s^{\widetilde{s}} \Psi_x(v)e^{-z(v-s)-\rho v}dv \tag{26}$$

where $\Psi_x(s)$ is shorthand for $\Psi_x(t^{\Psi}(x^z(s), x^z(s))$. We must now check that $\lambda_x(s) \leq 0$, so that $\beta(s) \geq 0$. We know from (22) that $\lambda_x(\widetilde{s}) \leq 0$. Thus, it is sufficient to establish that $\Psi_x(s) \leq 0$

for all $s \in [0, \widetilde{s}[$. We need to do this for the case of positive contributions ($t^{\Psi}(x) > t^{W}(x)$) and the case of zero contributions ($t^{\Psi}(x) \leq t^{W}(x)$). If there are no contributions, using $W_x + W_x^* = (x^{ft} - x)/2$ we find

$$\Psi_x = -(a/2)(x - x^{ft}) + (p - 1 - x/2)$$

Since $x > x^{ft}$, the first term is negative. To show that the second term is also negative, note that there are two cases: (1) $t^{\Psi}(x) \leq t^{er}(x) \leq t^{W}(x)$ and (2) $t^{\Psi}(x) \leq t^{W}(x) \leq t^{er}(x)$. In case (1) $p(t^{er}(x), x) = 1$, which implies that $p(t^{\Psi}(x), x) < 1$. In case (2) we would have $x > x^{W}$, which implies that $p(t^{W}(x), x) < 1$ and hence $p(t^{\Psi}(x), x) < 1$. Thus, the second term above is negative.

If there are positive contributions, then (applying the Envelope Theorem) we obtain

$$\Psi_x = a(W_x(t^{W}(x), x) + W_x^*(t, x)) + (p - c - 1 + x(p_x - c_x))$$

Given that $t^{er}(x) > t^{\Psi}(x) > t^{W}(x)$, then $x > x^{W}$ and consequently $W_x(t^{W}(x), x) < 0$, as we showed in the proof of Proposition 4. W_x^* is always negative. $p - c - 1$ is zero at $t^{er}(x)$, hence it must be negative at $t^{\Psi}(x)$ given that $t^{\Psi}(x) < t^{er}(x)$. Hence, it suffices to show that $p_x - c_x < 0$ when evaluated at $t^{\Psi}(x)$. But $p_x - c_x = -1/2 - (1/x)\left[C_x - C/x\right]$. Since $C_x = (a/4)(t^{\Psi}(x) - t^{W}(x)) > 0$, then it is sufficient to establish that $1/2 - c(t^{\Psi}(x), x)/x > 0$. But in the proof of Proposition 1 we already established that $1/2 - c(t^{J}(x), x)/x > 0$. Given that $c(t^{J}(x), x) > c(t^{\Psi}(x), x)$, then this last inequality implies the previous one.

We have established that the conjectured paths for x and t together with the implied state variable y and costate variables λ_x and λ_y given by (26), (22), (20) and (25), (21), (18) in phases 1, 2, 3, respectively, and Kuhn-Tucker multipliers $\beta = -\lambda_x$ and ϕ given by (24), (23), and (19) in phases 1, 2, and 3, respectively, satisfy all the necessary conditions for an optimum. We now show that the conditions for sufficiency are also satisfied.

We need to show that the maximized Hamiltonian is concave in (x, y). The maximized Hamiltonian is:

$$H^0(u(x,y), t(x,y), x, y, \lambda_x, \lambda_y, s) = e^{-\rho s}\Psi(t(x,y), x) - \lambda_x z x + \lambda_y \left[(\rho + z)y - g(t(x,y), x)\right]$$

Clearly, it is sufficient to show that $d^2H^0/dx^2 < 0$.

Let us first analyze this in the first phase, where $t(x,y) = t^{\Psi}(x)$, $\Psi_t = 0$ and $\lambda_y = 0$. Differentiating and using $dt^{\Psi}/dx = -\Psi_{xt}/\Psi_{tt}$ yields $d^2H^0/dx^2 = (e^{-\rho s}/\Psi_{tt})\left(\Psi_{xx}\Psi_{tt} - \Psi_{xt}^2\right)$.

The SOC for $t^\Psi(x)$ requires that $\Psi_{tt} < 0$. Hence $d^2H^0/dx^2 < 0$ if and only if $\Psi_{xx}\Psi_{tt} - \Psi_{xt}^2 > 0$, which is a condition for Ψ to be concave in (x,t) at $(x^z(s), t^\Psi(x^z(s)))$. Differentiation yields

$$\Psi_{tt} = -a/2 + 3a/4 - 3a/4 = (a/2)(3/2-1) - 3a/4 = -a/2 < 0$$

and $\Psi_{xx}\Psi_{tt} - \Psi_{xt}^2 = (a/24)(5a+12+a) - 1/4$, which is positive given $a > 1$.

Now let us move to the second and third phases, where $t = t^{er}(x^z(s))$ and $g = 0$. After some simplifications, and using $L(x) \equiv e^{\rho s} d^2H^0/dx^2$, we obtain:

$$\begin{aligned} L(x) &= 2\Psi_{tx}dt^{er}/dx + \Psi_{xx} + \Psi_{tt}(dt^{er}/dx)^2 + \Psi_t d^2t^{er}/dx^2 \\ &= dt^{er}/dx - a/2 - 1 - a/2(dt^{er}/dx)^2 + \Psi_t d^2t^{er}/dx^2 \end{aligned}$$

We need to show that $L(x) < 0$. We need to consider the case in which there are no contributions and also the case when there are contributions. If there are no contributions, then $dt^{er}/dx = 1$, and hence $L(x) = -a < 0$.

If there are contributions, then define $\xi \equiv \sqrt{x(2a(3v-4) - x(4a-1))} > 0$. (Note that we showed above that $x(2a(3v-4) - x(4a-1)) > 0$.) After some manipulation we get,

$$L(x) = \frac{2(a+1)[x(4a-1) - a(3v-4)]\xi^2 + a^2(x+ax-a)(3v-4)^2 - (5a^2+2a-2)\xi^3}{9a\xi^3}$$

At $x = x^W = \frac{3v-4}{2}$, $L(x) = \frac{a}{3}\left(1 - \frac{4a(2-v)}{3v-4}\right)$. But $a \geq \frac{v-1}{2-v}$ implies that this is negative. Since $L(x^W) < 0$, we can prove that $L(x) < 0$ for $x \geq x^W$ by showing that $L'(x) \leq 0$ over this range. Differentiating and simplifying we obtain

$$L'(x) = \frac{a^2(3v-4)^2[a(3v-4) - x(a(8-3v) - 3(v-1))]}{3\xi^5}$$

Note that $a \geq \frac{v-1}{2-v}$ implies $a(8-3v) - 3(v-1) \geq \frac{2(v-1)}{2-v} > 0$, so using $x \geq x^W = \frac{3v-4}{2}$, we get

$$\begin{aligned} L'(x) &\leq \frac{a^2(3v-4)^2\left[a(3v-4) - \frac{3v-4}{2}(a(8-3v) - 3(v-1))\right]}{3\xi^5} \\ &= \frac{a^2(3v-4)^3[v-1-a(2-v)]}{2\xi^5} \leq 0. \end{aligned}$$

Q.E.D.

Proof of Proposition 8:

At time s_1, the analysis of the optimal agreement is exactly as in Proposition 7, except that x_1 may be lower than $\hat{x}$. Part (iii) of the proposition is thus a simple corollary of Proposition 7. For parts (i) and (ii), the first step is to calculate y immediately before s_1 as a function of x_1, which we denote $h(x_1)$. The value of y immediately after s_1 can be computed from the dynamics of Proposition 7. However, y jumps up at s_1 by the amount of ex-ante contributions, which are equal to the total losses for the two governments from applying tariff $\bar{t}(s)$ rather than free trade (which is what the two governments would choose in the absence of ex-ante lobbying). Note that these contributions are zero if $x_1 = x^{ft}$ and positive if $x_1 > x^{ft}$. Therefore, $h(x^{ft}) = 0$ and $h(x_1) < 0$ for $x_1 \in (x^{ft}, \hat{x}]$.

We now derive necessary conditions for an equilibrium and argue that they imply the features described in the proposition. First, let $g^J(x) \equiv g(t^J(x), x)$ and note that $g^J(x) = 0$ has only one solution, $x = \hat{x}$ (follows from Proposition 1). Second, note that any equilibrium has to satisfy the following two conditions before the agreement takes effect at time s_1:

1. $\dot{y} = (\rho + z)y - g^J(x)$ (almost everywhere, where $\dot{y}$ is defined);
2. If $y < 0$ then $\dot{x} = -zx$.

We now show that the path of $y(s)$ before $s = s_1$ must be $y = 0$ in a first phase (which may be empty) and $y < 0$ in a second phase (which is always non-empty). The key is to argue that y cannot touch zero from below before $s = s_1$. First note that in equilibrium it must be $g^J(x) \geq 0$ (if $g^J(x) < 0$ we would have $x > \hat{x}$; but x can only rise above $\hat{x}$ when $y = 0$, hence we would have $\dot{y} > 0$, leading instantly to $y > 0$, which cannot happen in equilibrium). But by condition 1 above this implies that, at a given point in time, if $y < 0$ then $\dot{y} < 0$. Thus, if y goes below zero it never touches zero again before s_1.

Together with condition 2, this implies that once y falls below zero, $\dot{x} = -zx$ until $s = s_1$. Also note that $y = 0$ for an interval of time implies $x = \hat{x}$. Therefore in the first (possibly empty) phase we have $y = 0$ and $x = \hat{x}$, and in the second phase we have $y < 0$ and $\dot{x} = -zx$ until $s = s_1$, at which point the agreement is signed.

To show that the second phase cannot be empty we proceed by contradiction. Imagine that this phase was empty. Since $s_1 > s_0$ this implies that the first phase would have to be non-empty, and $x_1 = \hat{x}$. But then $h(s_1) < 0$ so it is not possible that $y = 0$ just before s_1 (i.e., at $s = s_1 - \varepsilon$ for ε infinitely small).

Next we show that $x_1 > x^{ft}$, i.e. the reallocation of capital is not completed by time s_1. Suppose by contradiction that $x_1 \leq x^{ft}$. Noting that $h(x_1) = 0$ and $g^J(x) > 0$ for $x \leq x_1$, then $y > 0$ just before s_1, which is not possible in equilibrium. This also implies that if $s_1 - s_0$ is sufficiently large, then the first phase must be non-empty. It is also clear that if $s_1 - s_0$ is small then the first phase must be empty. To see this, just note that since $h(x_1) < 0$ then at $s = s_1 - \varepsilon$ the value of y must also be negative, ruling out a stage with $y = 0$ before s_1 if $s_1 - s_0 = \varepsilon$.

We have shown that any equilibrium must have the properties stated in the Proposition. Now we show that an equilibrium exists. Let $\pi(s, \hat{s}) \equiv g^J(\hat{x}e^{-z(s-\hat{s})})$ and note that the no-arbitrage condition for the second phase of adjustment (i.e., $s \in [\hat{s}, s_1)$) can be written as a differential equation $\dot{y} = (\rho + z)y - \pi(s, \hat{s})$ subject to the terminal condition $y(s_1) = h(x_1)$, where $x_1 = \hat{x}e^{-z(s_1-\hat{s})}$. The solution of this differential equation is

$$y(s; \hat{s}) = h(\hat{x}e^{-z(s_1-\hat{s})})e^{-(\rho+z)(s_1-s)} + \int_s^{s_1} \pi(v, \hat{s})e^{-(\rho+z)(v-s)}dv$$

In equilibrium the value of y at $s = \hat{s}$ is given by $y(\hat{s}; \hat{s})$. This implies that the equilibrium value of $\hat{s}$ must satisfy $y(\hat{s}; \hat{s}) \leq 0$. Alternatively, if $y(s_0; s_0) > 0$ then given that $y(s_1, s_1) \leq 0$ we see that - by continuity of the function $y(s; s)$ - there must exist a solution to $y(\hat{s}; \hat{s}) = 0$, and each solution $\hat{s}$ must be higher than s_0. This establishes that there is a first phase with no adjustment. **Q.E.D.**

References

Bagwell, Kyle and Robert W. Staiger (1999), "An Economic Theory of GATT", *American Economic Review*, 89(1), pp. 215-248.

Bagwell, K. and R. Staiger (2001), "Reciprocity, Non-discrimination and Preferential Agreements in the Multilateral Trading System", *European Journal of Political Economy*, 17(2), pp. 281-325.

Bagwell, K. and R. Staiger (2005), "Enforcement, Private Political Pressure and the GATT/WTO Escape Clause," *Journal of Legal Studies*, 34(2), pp. 471-513.

Bchir, Mohamed Hedi, Jean Sébastien and David Laborde (2005): "Binding Overhang and Tariff-Cutting Formulas: A Systematic, World-Wide Quantitative Assessment," mimeo.

Bond, Eric W. and Jee-Hyeong Park (2004): "Gradualism in Trade Agreements with Asymmetric Countries," *Review of Economic Studies*, 69, pp. 379-406.

Chisik, Richard, 2003: "Gradualism in Free Trade Agreements: a Theoretical Justification," *Journal of International Economics*, 59 (2) pp. 367-397.

Conconi, Paola and Carlo Perroni (2004), "Self-Enforcing International Agreements and Domestic Policy Credibility," mimeo.

Conconi, Paola and Carlo Perroni (2005), "The Economics of S&D Trade Regimes," mimeo.

Deveraux, Michael B. (1997), "Growth, Specialization and Trade Liberalization," *International Economic Review*, 38, pp. 565-586

Furusawa, Taiji and Edwin L.-C. Lai (1999): "Adjustment Costs and Gradual Trade Libealization," *Journal of International Economics*, 49, pp. 333-61.

Gawande, Kishore, Pravin Krishna and Michael. J. Robbins (2006), "Foreign Lobbies and U.S. Trade Policy," *Review of Economics and Statistics*, 88(3), pp. 563-571.

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

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

Grossman, Gene M. and Elhanan Helpman (1995b), "The Politics of Free-Trade Agreements", *American Economic Review*, 85(4), pp. 667-90.

Horn, Henrik, Giovanni Maggi and Robert W. Staiger (2006), "Trade Agreements as Endogenously Incomplete Contracts", NBER Working Paper No 12745.

Johnson, Harry G. (1954), "Optimum Tariffs and Retaliation, *Review of Economic Studies*, 21(2), pp. 142-153.

Krishna, Pravin (1998), "Regionalism and Multilateralism: A Political Economy Approach," *Quarterly Journal of Economics*, 113(1), pp. 227-52.

Lapan, Harvey E. (1988), "The optimal tariff, production lags, and time consistency," *American Economic Review*, 78(3), pp. 395-401.

Lockwood, Ben and Ben Zissimos (2005), "The GATT and Gradualism," mimeo.

Maggi, Giovanni (1999), "The role of multilateral institutions in international trade cooperation," *American Economic Review*, 89, pp. 190-214.

Maggi, Giovanni and Andrés Rodríguez-Clare (1998), "The Value of Trade Agreements in the Presence of Political Pressures," *Journal of Political Economy*, 106(3), pp. 574-601.

Maggi, Giovanni and Andrés Rodríguez-Clare (2006), "Import Tariffs, Export Subsidies, and the Theory of Trade Agreements," mimeo.

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

Mitra, Devashish, (2002): "Endogenous Political Organization and the Value of Trade Agreements," *Journal of International Economics*, 57, pp. 473-85.

Mussa, Michael L., (1986): "The Adjustment Process and the Timing of Trade Liberalization," in A. M. Choksi and D. Papageorgiou (eds.), *Economic Liberalization in Developing Countries*, Basil Blackwell, New York.

Ornelas, Emanuel (2005), "Rent Dissipation, Political Viability and the Strategic Adoption of Free Trade Agreements," *Quarterly Journal of Economics* 120(4), pp. 1475-1506.

Staiger, Robert W. (1995): "A Theory of Gradual Trade Libealization," in J. Levinsohn, A. Deardorff and R. Stern (eds.), *New Directions in Trade Theory*, University of Michigan Press, Ann Arbor, Michigan.

Staiger, Robert W. and Guido Tabellini (1987): "Discretionary Trade Policy and Excessive Protection," *American Economic Review*, 77, pp. 823-37.

Figure 1: The Long Run Non-Cooperative Equilibrium

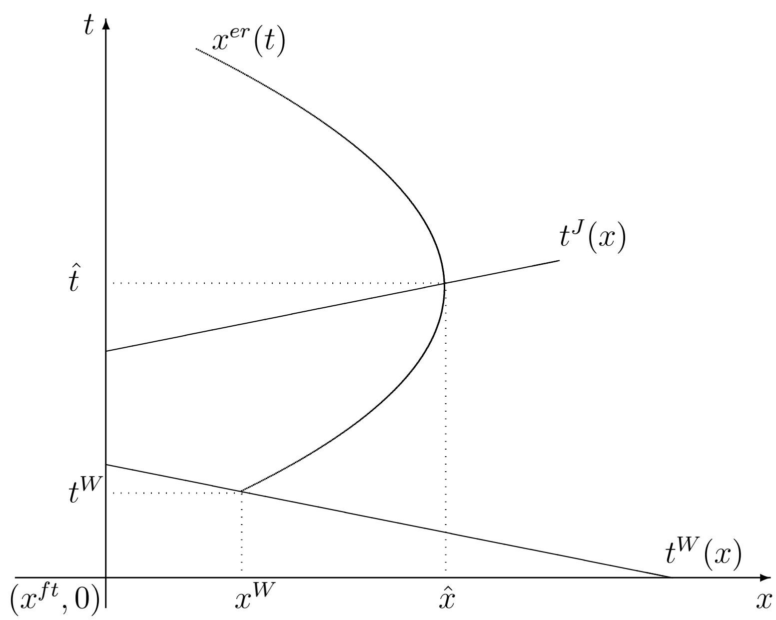

Figure 2: The Trade Agreement with Perfect Capital Mobility

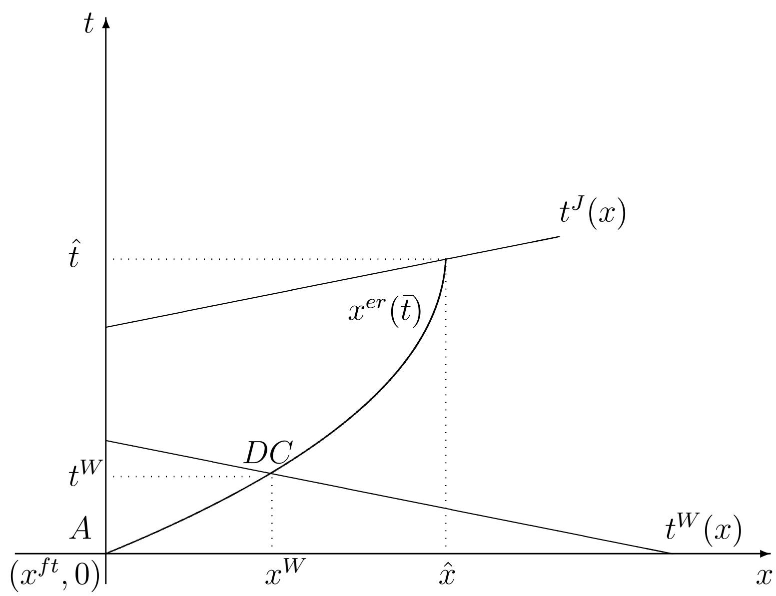

Figure 3: The Trade Agreement with Imperfect Capital Mobility

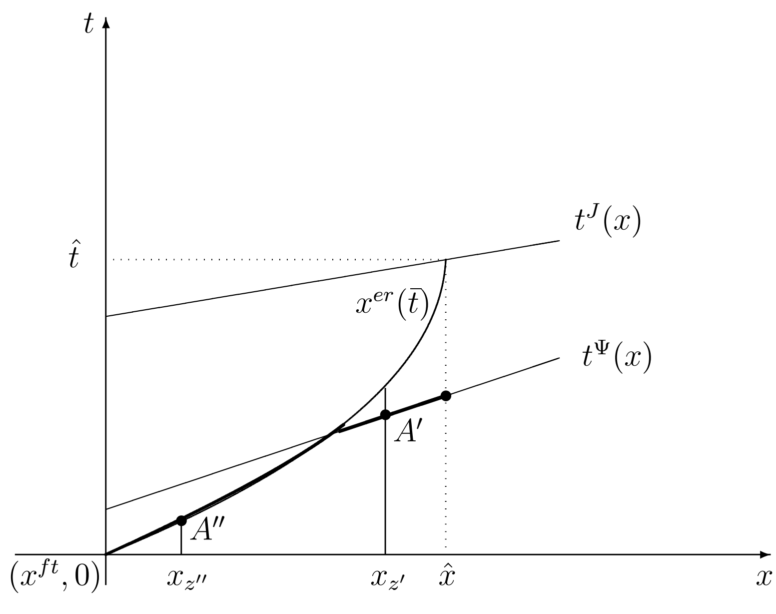

Figure 4: Ceilings vs Exact Tariff Commitments