Coercive Trade Policy*

Vincent Anesi†
Giovanni Facchini‡

September 29, 2014

Abstract

Empirical evidence suggests trade coercion exercised unilaterally is significantly less likely to induce concessions than coercion exercised through an international organization. In this paper we build a two-country model of coercion that can provide a rationale for this finding, and for how "weak" international institutions might be effective, even if their rulings cannot be directly enforced. In particular we show that if coercion is unilateral, the country requesting the policy change will demand a concession so substantial to make it unacceptable to its partner, and a trade war will ensue. If the parties can instead commit to an international organization (IO), compliance is more likely, because the potential IO ruling places a cap on the Foreign government's incentives to signal its resolve.

JEL classification numbers: F12, F16, L11.
Keywords: GATT, WTO, Dispute Settlement, Political Economy.

*We wish to thank Kyle Bagwell, Giovanni Maggi, Doug Nelson, Emanuel Ornelas and participants at the XIV SAET Conference in Tokyo.

†University of Nottingham and GEP. Address: School of Economics, Room B48, The Sir Clive Granger Building, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom. Email: vincent.anesi@nottingham.ac.uk.

‡University of Nottingham, University of Milan, CEPR, CES-Ifo, GEP, IZA and LdA. Address: School of Economics, Room B36, The Sir Clive Granger Building, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom. Email: giovanni.facchini@nottingham.ac.uk.

1 Introduction

In international trade disputes, coercion is often used against governments whose trade practices are deemed unfair. Trade coercion occurs when a "sender government" makes a demand backed by threats to use retaliatory sanctions against a "target government" if the latter does not acquiesce to this demand. There are typically two distinct methods of trade coercion: it can be exercised unilaterally or through multilateral institutions (e.g. GATT and WTO). In the case of unilateral coercion, the sender government makes a demand and (if necessary) retaliates one-sidedly, unconstrained by international obligations.[1] In the case of multilateral coercion, the sender uses instead an international institution's framework for trade dispute resolution.[2]

In this paper, we build on an empirical puzzle to develop a theory of trade coercion. The puzzle concerns the effectiveness of unilateral and multilateral coercions in getting target countries to concede to senders' demands: Empirical evidence (e.g. Busch 2000 and Pelc 2010) reveals in fact that a target of trade coercion from the US is significantly less likely to concede when coercion is unilateral than when it is multilateral. Given that neither GATT nor the WTO possess centralized enforcement power (Busch and Reinhardt 2000), the fact that these multilateral institutions can increase the chances of a sender government obtaining a concession presents an empirical puzzle. Why does unilateral coercion significantly reduce the likelihood of a target conceding? How can international trade institutions be effective if defendants can reject adverse rulings with impunity?

We address these questions by developing a theoretical model, which allow us to analyze the strategic incentives underlying trade coercion under three different institutional settings. The model depicts a dispute between two states, Home and Foreign, in which the Foreign government is dissatisfied with the trade policy implemented by the Home government. A key feature of trade coercion is the target government's lack of information on the sender government's domestic political constraints (e.g. Busch and Reinhardt 2000, Bagwell and Staiger 2005 and Beshkar and Bond 2012). To capture this idea, we assume that the political pressure exerted by the import-competing sector on the government in Foreign is private information, and is only known by the Foreign government. This political pressure plays a key role in shaping its level of resolve — i.e. the severity of its trade sanctions against the Home government — in a potential trade war.

Appraising the actual effectiveness of an international organization in dispute settlement

[1] A typical example was Section 301 of the 1974 US Trade Act, which allowed the United States to impose unilateral sanctions on countries whose trade practices were found to be unfair to US interests. This clause was invoked in several occasions – for instance in the much publicized dispute with Japan over automobiles of 1995, in which the US essentially bypassed the WTO and imposed sanctions unilaterally (see Puckett and Reynolds 1996 and Schoppa 1999).

[2] The WTO Dispute Settlement Mechanism is the leading institution of this kind, and since its inception, it has handled hundreds of cases. Several preferential trade agreements also include similar institutions. See for instance NAFTA's Dispute Settlement Process or MERCOSUR's Dispute Settlement Mechanism.

requires knowing what would happen if that institution did not exist — i.e. if there were no framework of rules governing trade coercion. For this reason, the first setting we examine is one in which unilateral coercion is the only option. The game begins with the Foreign government making a demand. The Home government can concede (ending the game with the implementation of the demanded tariff), or reject it (triggering a retaliatory trade war). In other words, it must decide which concessions are acceptable, that is, which tariff changes it would prefer to make rather than face Foreign's trade sanctions. Since the precise nature of these sanctions is uncertain and crucially depends on the privately observed level of resolve of the Foreign government, the latter has incentives to signal high levels of resolve by making excessive demands about the concessions required from Home to avoid retaliatory measures. Our characterization of equilibrium outcomes in this case reveals that such incentives lead the Foreign government to make requests that the Home government will not meet, thus causing a retaliatory trade war — even when there exist mutually advantageous policy concessions. This finding provides a possible explanation for the empirically observed lower effectiveness of unilateral coercion in obtaining concessions from target governments.

As we will show, a key factor in determining whether concessions can be obtained with multilateral coercion is the extent to which the sender government can commit not to bypass the dispute settlement process of the international organization through which coercion is channeled. To model the different strategic situations that may arise from differences in the sender's ability to commit to the international organization, we will examine two distinct variants of the previous model. In the first, the Foreign government is not allowed to bypass the international organization's dispute settlement process. As a result, multilateral coercion is its only option available. Dispute settlement is modeled by allowing the Foreign government to make a demand to the Home government prior to the international organization ruling. This assumption is intended to capture, e.g. the consultations stage of WTO disputes. If the Home government does not concede to the Foreign government's demand the international organization issues its ruling, whereas it remains inactive otherwise. As our aim is to investigate the effectiveness of weak international trade institutions — namely those that have no enforcement power and rely on the sender government itself to implement any retaliatory measures — the Home government is allowed not to comply with the ruling, thus triggering a trade war with the Foreign government. Our analysis shows that commitment to the international organization's ruling makes concessions more likely. Intuitively, the potential IO ruling places a cap on the Foreign government's incentives to signal its resolve with high demands. This results in the latter making more moderate requests, which can be accepted by Home.

In the second variant of the model, the Foreign government is only partially committed to the international organization's dispute settlement process, in the sense that it can choose between unilateral and multilateral coercion in an additional stage at the beginning of the game, committing itself to that choice.[3] This setting captures the environment created by Section 301 of the US Trade Act of 1974. In fact, this provision enabled the President to impose sanctions unilaterally against unfair trade practices, eliminating the need to observe existing international obligations (e.g. Puckett and Reynolds 1996). We show that the mere availability of the unilateral option prevents the foreign government from obtaining concessions in equilibrium. In fact, using multilateral coercion when unilateral coercion is available is perceived as a sign of the foreign government's weakness. Hence, incentives to signal higher levels of resolve to the Home government will lead the Foreign government to make unilateral demands which, as discussed above, cannot be accepted in equilibrium.

To the best of our knowledge, this is the first paper to provide a formal analysis of trade coercion. A large body of literature has studied international trade agreements as subgame perfect equilibria of infinitely-repeated prisoner-dilemma games, in which deviations from the (implicit) agreements are followed by indefinite play of high-tariff Nash equilibria. Papers in that literature study how international organizations' dispute settlement procedures can facilitate cooperation – e.g. Maggi (1999), Ludema (2001), Klimenko, Ramey, and Watson (2008), Limão and Saggi (2008), and Park (2011). However, in this tradition, trade dispute settlement is modeled as a set of conditions imposed on the off-the-equilibrium-path punishments that follow deviations, not explicitly as a coercion game.[4]

A recent literature has taken an incomplete-contracts approach to international trade agreements and dispute settlement — e.g. Bagwell, Mavroidis, and Staiger (2007), Beshkar (2010), Horn, Maggi, and Staiger (2010), Maggi and Staiger (2011) and Beshkar and Bond (2012). Its main focus is on the design of optimal institutions for international trade and dispute settlement in various informational/contractual environments. In contrast, the international organization's dispute settlement procedure is the main *exogenous* variable in our model. Our aim is not to study the normative aspects of trade institutions but, instead, to provide a positive theory of how commitment to such institutions may affect trade coercion outcomes.

The rest of the paper is organized as follows. Section 2 describes the model, while section 3 presents the main results of our analysis. In section 4 we discuss the substantive implications of our results and relate them to the existing empirical evidence. Section 5 concludes.

[3] As we will discuss in section 2.2, this assumption is consistent with empirical evidence on US trade coercion.

[4] Other notable recent examples are Bagwell and Staiger (2005), Martin and Vergote (2008), and Rosendorff (2005) who analyze repeated tariff games in which, as in our model, governments have private information about their relative valuations of import-competing sectors. Riezman (1991) and Hungerford (1991) also analyze repeated game settings with incomplete information but, in those papers, it is the countries' own levels of protection which are private information.

2 The Model

The goal of this section is twofold. We start by presenting the basic structure of the economy, and lie out next a simple model of trade coercion.

2.1 The Economic Environment

We consider a model with two large countries, Home and Foreign, trading between each other, which has been used in several previous analyses of trade negotiations.[5] Each economy is characterized by three sectors, $i = 0, 1, 2$. All goods are produced using a constant-returns-to-scale technology and are sold under conditions of perfect competition. The freely traded good 0 serves as the numeraire and is produced using labor alone. We choose units so that the international and domestic prices are both equal to one. We assume that aggregate labor supply, $L = \overline{L}$, is large enough to sustain production of a positive amount of good 0. This implies that in a competitive equilibrium the wage rate equals unity in each country. Goods 1 and 2 are manufactured using labor and a sector-specific input, which is available in fixed supply. Home is abundant in sector-specific input 2, while Foreign is abundant in sector-specific input 1. As a result, Home imports good 1, while Foreign imports good 2. We assume symmetry in factor endowments between the two countries. The domestic and international prices of a nonnumeraire good i are denoted by p_i and π_i, respectively, and the rent R_i, accruing to the specific factor in sector i, depends only on the producer price of the good, and can thus be expressed as $R_i(p_i)$. Industry supply is given by $Q_i(p_i) = \partial R_i/\partial p_i$.

Trade policies in the two countries consist of ad valorem import tariffs or subsidies, denoted by τ and τ^*, which drive a wedge between domestic and international prices.[6] In Home, the domestic price of good 1 is thus equal to $p_1 = (1+\tau)\pi_1$, with $\tau > 0$ ($\tau < 0$) representing an import tariff (subsidy); the domestic price of the export good is instead equal to $p_2 = \pi_2$. In Foreign, domestic prices are given by $p_1^* = \pi_1$ and $p_2^* = (1+\tau^*)\pi_2$.

The economy is populated by a continuum of agents, and the size of the population is normalized to one. Each agent shares the same quasi-linear and additively separable preferences, which can be written as

$$u(c_0, c_1, c_2) \equiv c_0 + \sum_{i=1}^{2} u_i(c_i), \tag{1}$$

where c_0 represents the consumption of the numeraire good, and c_1 and c_2 represent the consumption of the other goods. The sub–utility functions are assumed to be twice differentiable,

[5] See for instance Mayer (1981) and Conconi, Facchini, and Zanardi (2012).

[6] Following Johnson (1954) and Mayer (1981)), we restrict the set of policy tools available to import tariffs and subsidies. This allows us to describe the preferences of the two countries in the tariff space (τ, τ^*) and to easily characterize trade negotiations between them. Levy (1999), in his model of lobbying and international cooperation, has convincingly argued that export subsidies and taxes are rarely used, the only exception being probably agriculture.

increasing, and strictly concave.

Provided that income always exceeds the expenditure on the numeraire good, the domestic demand for good $i \in \{1,2\}$ can be expressed as a function of price alone, $D_i(p_i)$. Imports of good 1 by Home can then be written as $M_1(p_1) = D_1(p_1) - Q_1(p_1)$, while its exports are given by $X_2(p_2) = Q_2(p_2) - D_2(p_2)$.

World product markets of goods 1 and 2 clear when

$$M_1\Big((1+\tau)\pi_1\Big) - X_1^*(\pi_1) = 0, \tag{2}$$

$$M_2^*\Big((1+\tau^*)\pi_2\Big) - X_2(\pi_2) = 0. \tag{3}$$

From (2) and (3) we can derive an expression for world equilibrium prices as a function of the policies in the two countries, i.e., $\pi_1(\tau)$, $\pi_2(\tau^*)$. Tariff revenues in Home are

$$\mathbf{T}(\tau) = \tau\pi_1(\tau)M_1(\tau) \tag{4}$$

and are assumed to be redistributed uniformly among all domestic residents.

Individuals derive income from several sources: they all supply one unit of labor and earn wages; they also receive the same lump sum transfer (possibly negative) of trade policy revenues from the government and they own the same share of the specific inputs used in the production of goods 1 and 2. We assume that the Home government seeks to maximize aggregate welfare, which is defined as the sum of the income of all citizens (total labor income, industry rents and government revenues), plus consumer surplus, i.e.:

$$W(\tau,\tau^*) = 1 + R_1(\tau) + R_2(\tau^*) + \mathbf{T}(\tau) + \Omega(\tau) + \Omega(\tau^*), \tag{5}$$

where $\Omega(\tau) \equiv u\Big(D_1(\tau)\Big) - p_1 D_1(\tau)$ and $\Omega(\tau^*) \equiv u\Big(D_2(\tau^*)\Big) - p_2 D_2(\tau^*)$, i.e. the first term describes the surplus from the consumption of good 1 and the second from the consumption of good 2.

Foreign is identical to Home, with one important exception. While in the Home country the political influence of the import competing sector is normalized to one, in Foreign it is equal to $\gamma \in \big[\underline{\gamma}, \overline{\gamma}\big]$, and in the remainder of the paper γ will be referred to as the Foreign government's type.[7] As a result the objective function of the Foreign government is given by:

$$W^*(\tau,\tau^*) = 1 + \gamma R_1^*(\tau^*) + R_2^*(\tau) + \mathbf{T}^*(\tau^*) + \Omega(\tau^*) + \Omega(\tau), \tag{6}$$

As is standard in this class of models (e.g. Rosendorff 2005), we make the following natural assumptions about both governments' objective functions. First, for any given level of Home

[7] See Bagwell and Staiger (2005) for a similar setting.

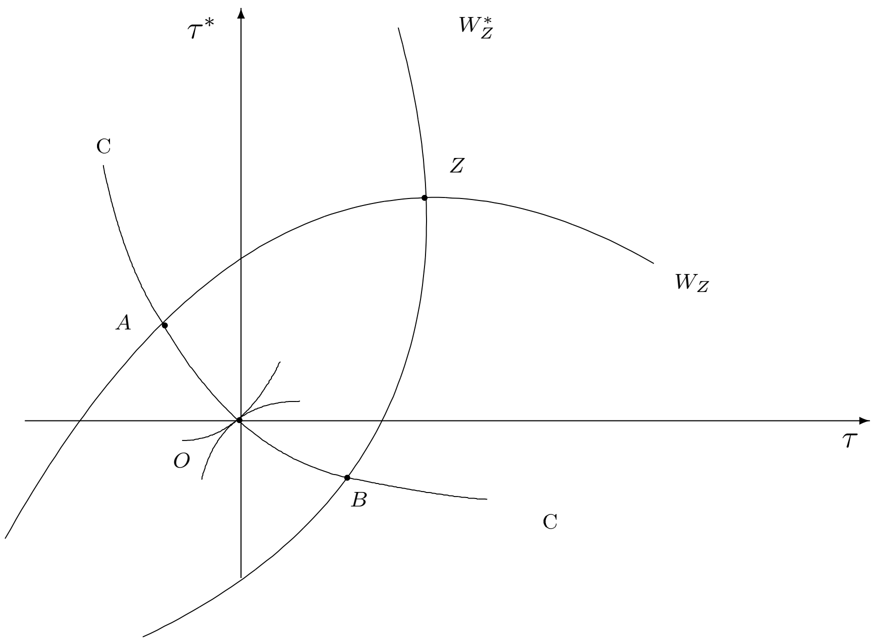

Figure 1: Policy preferences with $\gamma = 1$

tariff τ [resp. of Foreign tariff τ^* and type γ], $W(\tau, \cdot)$ [resp. $W^*(\cdot, \tau^*, \gamma)$] strictly decreases with τ^* [resp. with τ]. This simply ensures that, in each country, the losses incurred by domestic export firms when the other country raises its tariff always outweigh the benefits to domestic consumers. Second, for any given level of τ^* [resp. of τ and γ], $W(\cdot, \tau^*)$ [resp. $W^*(\tau, \cdot, \gamma)$] first increases and then decreases with τ [resp. with τ^*]. This ensures that $W(\cdot, \tau^*)$ and $W^*(\tau, \cdot, \gamma)$ have unique maximizers, which we denote by $\tilde{\tau}$ and $\tilde{\tau}^*(\gamma)$ respectively. (Additive separability in (5) and (6) implies that $\tilde{\tau}$ is independent of τ^*, and that $\tilde{\tau}^*(\gamma)$ is independent of τ.)

Tariffs $\tilde{\tau}$ and $\tilde{\tau}^*(\gamma)$ are clearly those which would be implemented if governments chose their policies non-cooperatively — or, using the language of the previous literature, if they engaged in a "trade war." Figure 1 provides an illustration for governments' preferences: W_H (resp. W_F^*) describes an indifference curve for the Home (resp. Foreign) government. A downward (resp. leftward) shift leads to higher values of the government's objective function. The policy pair $(\tilde{\tau}, \tilde{\tau}^*(\gamma))$ lies at the intersection between the two curves and describes the coordinates of point Z. Clearly the two governments could make themselves better off if they could agree on any tariff pair lying within the lense described by the two indifference curves.

2.2 A Simple Model of Trade Coercion

Let (τ_0, τ_0^*) be the status quo trade policy implemented by Home and Foreign. Consider a situation in which the Foreign government is dissatisfied with Home's existing policy, and has decided to use coercion to reduce Home's tariff, i.e. it threatens to increase τ_0^* if Home does not implement a new trade policy $\tau < \tau_0$.[8]

The goal of our analysis is to investigate whether and how different institutional arrangements affect the outcome of trade coercion. To this end, we develop a model with two active players, the Home and the Foreign governments, which possibly interact with an international organization (IO). At the preliminary stage, the Foreign government privately observes γ which is drawn from a cumulative distribution function F_0. We assume that F_0 has a continuous and strictly positive density over $\left[\underline{\gamma}, \overline{\gamma}\right]$. The sequence of events that follow the realization of γ depends on Foreign's institutional arrangements for trade coercion:

(i) Absence of IO membership. Suppose first that Foreign is not a member of the IO, so that coercion must be unilateral. In this case, the Foreign government threatens to increase its tariff unless the Home government acquiesces to a demand $\tau \leq \tau_0$. If the Home government concedes, reducing its tariff from τ_0 to τ, then Foreign does not impose any sanction. Then the policy vector (τ, τ_0^*) is implemented. If the Home government stands firm, then the Foreign government carries out its threat, thereby triggering a trade war.

(ii) Full Commitment to the IO. Suppose now that Foreign is a member of the IO, and is fully committed to its dispute settlement process — so that coercion must be conducted multilaterally. The process through which disputes are settled in international trade organizations is usually long and complex. It typically involves consultations between the sender and target (and potentially third parties and/or mediators) to reconcile their differences by themselves, IO panels' hearings and parties' rebuttals, several reports from the IO panel to the parties and, in the absence of an early settlement, rulings and appeals. Our aim here is to focus on the effects of incomplete information on multilateral negotiation outcomes and, therefore, to abstract away from any other complexity that such a situation might entail. To this end, and to ease comparison with the previous framework, we model proceedings as follows.

First, both parties observe the realization of the IO panel's "interpretation" of the trade agreement, τ^{io}. The Foreign government then makes a demand $\tau \leq \tau_0$. The Home government can concede to this demand (ending the game with the implementation of the policy pair (τ, τ_0^*)) or reject it. In the latter case the IO issues ruling τ^{io}. The Home government reacts to the ruling in one of two ways: compliance (ending the game with the implementation of policy pair (τ^{io}, τ_0^*)), or noncompliance. If it fails to comply with the ruling, then the IO authorizes Foreign to retaliate and a trade war ensues.

Although this is a highly abstract version of GATT-WTO proceedings, it contains all the

[8]Note that since we are interested in developing a model of coercion as opposed to bargaining, we do not allow the Foreign government to use its tariff as a bargaining instrument when it formulates its demand. In other words, in this model, as in Bagwell, Mavroidis, and Staiger (2007), the import-competing sector is only a "retaliation-good sector," in the sense that the Foreign government can only use its tariff τ^* as a retaliation instrument when coercion is unsuccessful.

elements needed to study the impact of incomplete information and IO membership on trade coercion, which is the main focus of the present paper.

(iii) Partial Commitment to the IO. As explained in the introduction, it is interesting to consider also an intermediate case in which the Foreign government initially decides whether to coerce unilaterally or multilaterally. The remainder of the game is as in (i) if it chooses to coerce unilaterally, and as in (ii) otherwise. This setting captures for instance the working of the Section 301 provision of the 1974 US Trade Act, under which action on a dispute could be unilateral or accompanied by a GATT/WTO complaint (Busch and Reinhardt 2000, Pelc 2010).

Before we proceed with the analysis, we need to discuss three of the assumptions of the model. First, we treat the IO ruling τ^{io} as exogenous. Note that our main goal is to study how countries' commitment to international dispute settlement mechanisms affect trade coercion outcomes. Consequently, in our model the IO dispute settlement process is taken as given. The value of τ^{io} can simply be interpreted as the governments' (common) expectations about the ideal ruling of the decisive IO-panel member. More specifically, one can think of the IO as an organization with its own social welfare function (maximized by τ^{io}), which is unaffected by the political pressure from domestic actors. Beyond intrinsic policy preferences concerning the current situation, this objective function may also be influenced by other external factors — e.g. consistency with previous rulings and setting precedents in anticipation of potential future disputes. A second assumption of the model is that, once the Foreign government has filed a complaint with the IO, it always complies with the IO ruling and empirical evidence supports this view. In fact, as observed by Pelc (2010), "... once the United States began GATT proceedings, it did not turn back to unilateralism." In particular, the United States never retaliated unilaterally nor threatened to do so after a panel decision was reached. Finally, we assume that, even in the case of full commitment to the IO, noncompliance to a ruling leads to a trade war. This evidently does not mean that the IO falls apart whenever a defendant spurns its ruling. In reality, the WTO only authorizes the complainant to retaliate on a noncomplying defendant within certain limits. However, even such constrained retaliatory trade sanctions might cause the target to retaliate in turn, leading to escalation into further sanctions. We thus assume – for simplicity – that a trade war follows noncompliance. It is important to note though that all of our qualitative results carry over to alternative settings with constrained retaliation.

Each variant of the model describes a sequential game of incomplete information. We solve it by looking for (pure strategy) perfect Bayesian equilibria, which are defined as follows: (a) the Home government's beliefs are generated by Bayesian updating whenever possible and (b) in each stage governments' actions are optimal, given their beliefs and their opponents' strategies.

In order to eliminate equilibria which rely on implausible beliefs off the equilibrium path, we use criterion D1 from Cho and Kreps (1987). Intuitively, this refinement requires that if the set of Home government's actions that make some foreign government's type γ willing to deviate is strictly smaller than the set of actions that make some other type γ' willing to deviate, then the Home government should believe that type γ' is infinitely more likely to deviate than γ is.[9] In the remainder of the paper, any reference to an "equilibrium" is to a perfect Bayesian equilibrium consistent with criterion D1.[10]

3 International Trade Institutions and Coercion Outcomes

In this section we characterize in turn the equilibria that will emerge from the three institutional settings described in the previous section.

3.1 Benchmark: Coercive Trade Policy in the Absence of the IO

Both because it is empirically relevant and because it provides a benchmark to compare outcomes with those possible when the Foreign government can coerce multilaterally, we start by analyzing the case in which the Foreign country is not a member of the IO.

Trade wars and reservation demands. To solve the game, we begin with the last stage in which the two governments engage in a trade war. Although this continuation game may involve the presence of asymmetric information, it always has a unique equilibrium outcome: the Home government adopts its ideal tariff $\tilde{\tau}$, irrespective of the Foreign government's policy choice; likewise, the type-γ Foreign government adopts its ideal tariff $\tilde{\tau}^*(\gamma)$, irrespective of the Home government's policy choice.

Given the outcome of a trade war, consider now the Home government's decision of whether to concede to the Foreign government's demand τ. Suppose that its beliefs about γ are given by some c.d.f. F. It will concede to demand τ if and only if its payoff from conceding exceeds its expected payoff from triggering a trade war; that is

$$W(\tau, \tau_0^*) \geq \int_{\underline{\gamma}}^{\overline{\gamma}} W(\tilde{\tau}, \tilde{\tau}^*(\gamma))\, dF(\gamma) \ .$$

Let the smallest value of τ that satisfies the above inequality be denoted by $T(F)$. This is the Home government's "reservation demand," or the minimum demand it will accept rather than engage in a trade war. In what follows, we will sometimes indulge in a slight abuse of

[9] This is a strengthening of the Intuitive Criterion, which has no bite in this game. See the Appendix for the formal definition.

[10] In order to limit the number of possible cases (without affecting the paper's conclusions), we also assume that in case of a tie, a player will prefer to agree than to disagree with the other player or the IO.

notation and denote by $T(\gamma)$ the Home government's reservation tariff when its beliefs assign probability 1 to type γ. Similarly, the type-γ Foreign government's reservation demand $T^*(\gamma)$ — that is, the Home tariff at which the Foreign government is indifferent between settling and engaging in a trade war — is defined as the largest value of tariff τ that satisfies

$$W^*\left(\tau, \tau_0^*, \gamma\right) \geq W^*\left(\tilde{\tau}, \tilde{\tau}^*(\gamma), \gamma\right)$$

(recall that $W^*\left(\tau, \tau_0^*, \gamma\right)$ decreases as τ increases).

It can be easily shown that $T(\gamma)$ and $T^*(\gamma)$ are both strictly decreasing in γ. An increase in γ causes the trade-war tariff of the Foreign government, $\tilde{\tau}^*(\gamma)$, to rise. As $W\left(\tilde{\tau}, \tilde{\tau}^*(\gamma)\right)$ decreases with $\tilde{\tau}^*(\gamma)$ (and therefore with γ), the Home government is willing to implement a lower tariff to avoid a trade war. In contrast, applying the Envelope Theorem reveals that $W^*\left(\tilde{\tau}, \tilde{\tau}^*(\gamma), \gamma\right)$ increases with γ; so that greater political pressure from its import-competing sector makes the Foreign government less willing to tolerate high tariffs applied by Home. To insure that the model always has an equilibrium, we assume throughout our analysis that $T^*\left(\underline{\gamma}\right) < T\left(\underline{\gamma}\right)$. This assumption implies that there is no room for compromise when the Foreign government is not prepared to resort to high retaliatory tariffs.

The ineffectiveness of unilateral coercion. Can the Foreign government obtain a concession from the Home government in equilibrium? This question is answered in the following

Proposition 1. *Suppose that there is no IO — so that coercion must be unilateral. There exists an equilibrium and, in any equilibrium, the Foreign government always fails to obtain a concession from the Home government.*

Proof. See Appendix A.1. □

To understand the intuition for this result, note that upon observing the demand τ by the Foreign government, the Home government — uninformed about the level of political pressure γ that has emerged in the Foreign country — updates its beliefs. Given these new beliefs, say F, it concedes to τ if and only if $\tau \geq T(F)$. As its reservation demand $T(\gamma)$ is decreasing in γ, the best strategy for the Foreign government is to signal high values of γ by requiring a low level of τ. Indeed, trade wars are less costly to Foreign governments that are very sensitive to the well–being of the import sector (characterized by a high-γ)— the Foreign reservation demand $T^*(\gamma)$ decreases with γ — and the Foreign government is therefore more likely to risk a trade war when γ is large. Understanding this, the Home government rationally infers higher values of γ from a demand for a lower tariff. Such beliefs lead the foreign policy-maker to go too far, however, and to make requests which the Home government is not prepared to meet. This signaling spiral leads all types of Foreign government to make unsuccessful demands, and a trade war will ensue in every equilibrium.

3.2 Coercive Trade Policy with Full Commitment to the IO

We now turn to the analysis of the consequences of full commitment to the IO on trade coercion outcomes. One of the questions this paper seeks to answer is how international trade institutions, despite their lack of enforcement power, can be effective in settling disputes. We have just shown how the logic of unilateral trade coercion locks the Foreign government into signaling spirals leading to trade wars. Despite being unable to enforce its rulings, can the IO's dispute settlement process do a better job of obtaining concessions from the Home government?

The answer is positive, and the intuition is that full commitment to the IO's dispute settlement process may offer the Foreign government an opportunity to break the spiral of unilateral coercion. To see how this can occur in equilibrium, suppose that $\tau^{io} \geq T(F_0)$. Consider first the stage in which the Home government must decide whether or not to comply with the IO ruling τ^{io}. Failure to comply would trigger a trade war. Therefore, it follows from the analysis of the trade-war stage we have developed in the previous section that it chooses to comply if and only if $\tau^{io} \geq T(F)$, where the c.d.f. F stands for the updated beliefs about the Foreign government's type γ at this stage. This implies that, when confronted with some demand τ from the Foreign government, the Home government's optimal strategy is to concede if and only if $\tau \geq \max\{\tau^{io}, T(F)\}$. As long as $T(F) > \tau^{io}$, the same signaling incentives as under unilateral coercion drive the Foreign government to be tougher, thus signaling high values of γ and reducing the Home government's reservation tariff $T(F)$. When $T(F) \leq \tau^{io}$, however, the Home government's reservation demand becomes constant and equal to τ^{io}: pessimistic beliefs cannot reduce it any further. Hence, the Foreign government is faced with two alternatives: make a successful demand $\tau \geq \tau^{io}$; or make an unsuccessful one, following which the Home government will comply with the IO ruling τ^{io}. As $W^*(\tau, \tau_0^*, \gamma)$ strictly decreases with τ for all γ, making demand $\tau = \tau^{io}$ is a best response for all types of Foreign government. Since the Home government's beliefs must be correctly derived from the Foreign government's equilibrium strategy and Bayes' rule, its beliefs when it receives the same demand τ^{io} from all types of Foreign government must be given by F_0. As $\tau^{io} \geq T(F_0)$, this confirms that there is an equilibrium in which all types of Foreign government make demand τ^{io}, to which the Home government concedes.

Our next result shows that the condition $\tau^{io} \geq T(F_0)$ is also necessary for a trade war to be avoided. If τ^{io} is too low then, as in the case of unilateral coercion, the Foreign government's demands spiral down to unacceptable levels leading to a trade war.

Proposition 2. *Suppose the Foreign government is fully committed to the IO — so that coercion must be multilateral. There always exists an equilibrium, and the following is true in any equilibrium:*

(i) If $\tau^{io} \geq T(F_0)$, then: either all types of Foreign government obtain the concession τ^{io}; or they all make unsuccessful demands following which the Home government complies with the

IO ruling.

(ii) if $\tau^{io} < T(F_0)$, then all types of Foreign government make unsuccessful demands following which the Home government fails to comply with the IO ruling.

Proof. See Appendix A.2. □

Combined with Proposition 1, Proposition 2 shows that an international organization can affect the outcome of trade coercion and prevent trade wars, even though it has no enforcement power. It also suggests a possible explanation for why trade coercion appears to be more effective in obtaining concessions from target governments when conducted multilaterally. We will elaborate on the empirical and normative implications of the equilibrium analysis in Section 4.

3.3 Coercive Trade Policy with Partial Commitment to the IO

Under partial commitment to the IO, the Foreign government is allowed to choose whether to coerce the Home government unilaterally or multilaterally. Suppose that $\tau^{io} \geq T(F_0)$, and that the Foreign government's type γ satisfies $T(\gamma) < \tau^{io} < T^*(\gamma)$, so that both countries are better off implementing τ^{io} than engaging in a trade war. Our analysis so far reveals that the signaling incentives inherent in unilateral coercion would lead the Foreign government to make inefficient demands to the Home government. To avoid this outcome, the type-γ Foreign government would therefore be expected to adopt multilateral coercion. Some authors argue, however, that taking a trade dispute to an international organization signals a lack of resolve — i.e., a low γ — by the sender government (e.g. Reinhardt 2000; Pelc 2010). The next proposition provides a formalization of their argument.

Proposition 3. *Suppose the Foreign government is only partially committed to the IO — so that it can choose between unilateral and multilateral coercion. There exists an equilibrium in which all types of Foreign government coerce unilaterally and fail to obtain a concession. In addition, a trade war arises with probability one in any equilibrium.*

Proof. See Appendix A.3. □

In other words, partial commitment to the IO yields the same outcome as absence of membership: In both cases, the Foreign government fails to obtain a concession from the Home government, and a trade war ensues.

Note though that the Foreign government's coercive policy has now two components: the demand τ and the method of coercion (unilateral vs. multilateral) through which this demand is made. A deviation from multilateral to unilateral coercion in this case conveys the same signal as a deviation to a smaller demand in the absence of an IO: the Home government therefore anticipates tougher retaliatory measures in case of a trade war. As in the unilateral-coercion game, such beliefs induce the Home government to concede to lower unilateral demands. This in turn drives the Home government to (unilaterally) ask for even lower tariffs until its demands become unacceptable.

These incentives to coerce unilaterally to signal high resolve can only disappear when in equilibrium all types of Foreign government make unsuccessful demands (either unilaterally or multilaterally), thus leading to a trade war. In this case, the Home government interprets any deviation by its foreign counterpart as an attempt to escape this outcome and, consequently, infers that the Foreign government's type γ must be low. It is therefore optimal for the Home government to only accept demands so high that the Foreign government prefers to engage in a trade war.

4 Implications

Our theoretical model provides novel insights on the influence of international trade institutions on coercion outcomes. Importantly, our results are consistent with the stylized facts that have been uncovered in the existing empirical literature. In this section we briefly review these empirical findings and explain how they relate to our analysis.

Unilateral vs. multilateral coercion: the influence of international trade institutions. Busch and Reinhardt (2000) observe that, during the GATT period, only two-fifths of the rulings in favor of the complainant resulted in full compliance by the defendant — whereas in nearly a third of the cases, defendants failed to comply at all. Even though the establishment of the WTO dispute settlement mechanism improved the situation, as Rossmiller (1994) pointed out, the WTO remains a "court with no bailiff." These observations prompt the following question: Can a multilateral institution influence coercion outcomes despite its lack of enforcement power? Empirical evidence uncovered by Pelc (2010), suggests that this is indeed the case. Focusing on the US experience between 1975 and 2000, he finds that disputes that went through the GATT, rather than relying only on Section 301, are 34 percentage points more likely to result in a concession.

Pelc (2010) suggests that it is the perceived illegitimacy of unilateral coercion and the importance of reputation which decrease the likelihood of a target conceding. While resistance to institutionally constrained demands entails the reputational cost of being branded a violator, resistance to unilateral threats — regarded as illegitimate by the rest of the world — yields a reputational benefit: It decreases the likelihood of being unilaterally targeted again in the future. Our formal analysis provides an alternative rationale, which focuses on the role played by the sender government's incentives. On the one hand, unilateral coercion creates signaling spirals leading the sender government to make unacceptable demands. On the other, commitment to a multilateral organization can break these spirals and allow the sender government to obtain concessions.

Early dispute settlements. Analyzing evidence on more than 600 GATT/WTO disputes from 1948 through 1999, Busch and Reinhardt (2000) observe that in a majority of cases (about 55%), no panel was ever established, and a further 8% of them ended prior to the issuance of a panel report. Paraphrasing them, a key question is why should target governments settle early given that they can spurn adverse rulings with impunity. They argue that the source of early concessions lies in the normative power of GATT/WTO rulings and in the pressure to abide by the norm: An adverse ruling may weaken the target government's political position in its own country, as well as its position in ongoing multilateral trade talks. As a result, if the target government is uncertain about the IO ruling, then it may prefer to concede beforehand.

Consistently with the evidence, Proposition 2(i) shows that pre-ruling settlements may occur in equilibrium.[11] Importantly though, in our setting, the mechanism at work is different: IO rulings do not convey any normative or reputational costs. When the Foreign government anticipates an unfavorable IO ruling (i.e. when $\tau^{io} \geq T(F_0)$), it expects the Home government to comply with this ruling. This leads the Foreign government to abandon aggressive strategies, and to make more accommodating demands to which the Home government is willing to concede.[12] Thus, it is mainly the sender government's (rather than target's) incentives which are affected by the prospect of the IO decision.

An alternative rationale for international trade agreements. Our model suggests a possible explanation for another empirical puzzle: Given that membership in an international trade organization may limit the (coercive) policy discretion of a national government, why would the latter choose to join a supranational body? Most of the existing literature on this topic suggests that states become members of such institutions to solve the coordination problem created by the terms of trade externality from tariffs (e.g. Bagwell and Staiger 1999). Our analysis reveals, however, that another driving force may emanate from informational asymmetries in trade coercion. By helping to explain why demands channeled through the multilateral system may be more successful than unilateral demands, our model provides a new rationale for states' commitment to multilateral institutions.

To see this, suppose that we add an initial stage to the game in which the Foreign government decides whether or not to fully commit to the IO. If $\tau^{io} < T(F_0)$, then it is indifferent between all institutional arrangements: a trade war is inevitable. Suppose instead that

[11] The proof of Proposition 2 (Section A.2 in the Appendix) shows that something even stronger is true if $\tau^{io} > T^*(\overline{\gamma})$: in *all* equilibria the Foreign government obtains an early concession from the Home government.

[12] In fact, Busch and Reinhardt (2000) point out that among those disputes ending prior to a ruling, 67% exhibit full or partial concession by the target government.

$T(F_0) \leq \tau^{io} \leq T^*(\gamma)$. An immediate corollary of Propositions 1-3 is that, in this case, the Foreign government is better-off fully committing to the IO.

The role of commitment to international organizations. Proposition 3 shows that institutions allowing sender governments to choose between unilateralism and multilateralism can reduce the effectiveness of coercion. A leading historical example of the coexistence of these two coercion methods is represented by Section 301 of the US Trade Act of 1974. This provision allowed the United States to take a number of unilateral retaliatory actions against any foreign measures deemed to violate existing agreements or otherwise impeding its interests. At the same time, the US retained access to the dispute settlement system provided by the GATT-WTO (Pelc 2010).

As argued by Pelc (2010) the availability of unilateral coercion did not deliver the expected results, and in fact the US "ultimately found it in its interest to ... push for greater formal constraints in the Uruguay Round that ultimately raised the costs of unilateralism further." In our model, if we allowed the Foreign government to choose between full and partial commitment to the IO, then it would strictly prefer the former whenever $T(F_0) \leq \tau^{io} < T^*(\gamma)$. The Foreign government would indeed be better off making a successful demand τ^{io} under full commitment (Proposition 2) than making an unsuccessful demand under partial commitment (Proposition 3). Unlike Pelc's explanation based on the illegitimacy of unilateral coercion, our result though stems from the Foreign government's strategic incentives created by the presence of a unilateral option. Even though the Foreign government would be better off if this option were not available, incentives to signal higher levels of resolve to the Home government by deviating from multilateral to unilateral coercion eventually lead the Foreign government to make unacceptable demands (Subsection 3.3). These incentives are reminiscent of Reinhardt (2000) observation that taking a dispute to the GATT was a signal of the complainant's lack of resolve.

5 Concluding Remarks

This paper is a first attempt at analyzing the strategic interactions that underlie coercive trade policy. We have studied trade coercion in settings where sender governments may show their resolve by demanding more concessions from target governments. We have seen how the temptation to exaggerate can reduce the likelihood of targets conceding. This problem is especially severe when the sender government is not (fully) committed to a multilateral dispute settlement mechanism. Then, unbound by international commitments, the sender may make excessive demands which are unacceptable to its target. Institutions through which demands are channeled thus matter to coercion outcomes. In accordance with empirical evidence, our results indicate that full commitment to (even weak) multilateral trade institutions makes trade coercion more effective in obtaining concessions from target governments.

There are a number of research avenues opened up by our results, two of which we will briefly discuss. First, our positive theory of the impact of multilateral institutions on trade coercion outcomes naturally prompts a normative question: What would an optimal dispute settlement mechanism be in the presence of informational asymmetries?[13] Answering this question would require a richer framework, i.e. one that would further our understanding of the effects of settlement mechanisms both on membership in international trade institutions and on target governments' policy choices that are likely to trigger coercive responses.

As we noted at the outset (c.f. footnote 8), our analysis focused on coercion itself and not on its ultimate origin. It would be interesting to investigate why do dissatisfied governments use coercion instead of potentially more efficient bargaining approaches. Trade coercion typically involves two policy instruments: the target's trade policy which is the source of the sender's discontent, and the sender's policy which is only used as a retaliation instrument. By focusing its demand on the former instrument, the sender government leaves out mutually advantageous agreements which would be available if its demand would involve instead a combination of both instruments. But we leave this as a topic for future research.

Appendix

A Proofs of the Propositions

A.1 Proof of Proposition 1

To prove Proposition 1, we must show that: (i) in any equilibrium of the game without the IO, the Home government never concedes to the Foreign government's demands; and (ii) there exists an equilibrium in which the Home government never concedes to the Foreign government's demands.

Claim 1: Suppose that the Foreign government can only coerce unilaterally. In any equilibrium, the Home government never concedes to its demands.

Proof. First of all, observe that only one demand can successfully be made in equilibrium. To see this, suppose that two different demands τ_1 and τ_2 are made in equilibrium by types γ_1 and γ_2, respectively. Assume without loss of generality that $\tau_1 < \tau_2$. By definition of an equilibrium,

[13] Maggi and Staiger (2011) answer a similar question, but in a complete-information setting where states of the world are "vague" and subject to interpretation by contracting governments.

type γ_2 must find it profitable to make successful demand τ_2; hence, $T^*(\gamma_2) \geq \tau_2 > \tau_1$. But this implies that type γ_2 could profitably deviate by making claim τ_1: $W^*(\tau_1, 0, \gamma_2) > W^*(\tau_2, 0, \gamma_2)$.

Now we establish the claim in two steps: (1) we first show that if a demand is successful in equilibrium, then it must emanate from a single type; and (2) we then show that this is impossible in equilibrium.

(1) We proceed by contradiction. Suppose that multiple types make a successful demand, say τ, in some equilibrium. From our initial observation above, all the other equilibrium demands are unsuccessful. Let $\Gamma_\tau \subseteq [\underline{\gamma}, \overline{\gamma}]$ be the set of types that demand τ, and let $\gamma^{\text{sup}} \equiv \sup \Gamma_\tau$ (observe that, by assumption, $\gamma^{\text{sup}} \in (\underline{\gamma}, \overline{\gamma}])$. By definition of a PBE, we must have $\tau \leq T^*(\gamma)$ for all $\gamma \in \Gamma_\tau$ — otherwise, some type in Γ_τ could profitably deviate by making an unacceptable demand — and, therefore, $\tau \leq T^*(\gamma^{\text{sup}})$. As $T^*(\gamma)$ is a strictly decreasing function, this implies that $\tau < T^*(\gamma)$ for all $\gamma < \gamma^{\text{sup}}$; so that all types $\gamma < \gamma^{\text{sup}}$ strictly prefer τ to a trade war. Hence, in equilibrium, all types $\gamma < \gamma^{\text{sup}}$ must make the unique successful demand τ. Furthermore, by definition of a PBE, all types $\gamma > \gamma^{\text{sup}}$ must prefer a trade war to τ: $\tau > T^*(\gamma)$ for every $\gamma > \gamma^{\text{sup}}$ (recall that indifferent types choose to avoid a trade war). By continuity of $T^*(\cdot)$, therefore, we must have $T^*(\gamma^{\text{sup}}) = \tau$. Being indifferent between τ and a trade war, the type-γ^{sup} Foreign government chooses τ. We have thus established that $\Gamma_\tau = [\underline{\gamma}, \gamma^{\text{sup}}]$.

Confronted with demand τ, the Home government — whose updated beliefs F_τ assign a probability of 1 to the event "$\gamma \in [\underline{\gamma}, \gamma^{\text{sup}}]$" — optimally chooses to concede in the equilibrium under consideration. As the distribution of types has full support on $[\underline{\gamma}, \gamma^{\max}]$, this implies that $\tau \geq T(F_\tau) > T(\gamma^{\text{sup}})$.

Now take any tariff $\tau' \in (T(\gamma^{\text{sup}}), \tau)$, and observe that no type of Foreign government demands τ' in equilibrium. Indeed, by definition, all types in Γ_τ demand $\tau \neq \tau'$. As for types γ outside Γ_τ, they must be greater than γ^{sup}. Therefore, if type $\gamma > \gamma^{\text{sup}}$ demanded $\tau' > T(\gamma^{\text{sup}}) > T(\gamma)$, then the Home government would concede, thus contradicting our previous result that only one demand can be successful in equilibrium. All the premises of Lemma 1 are thus satisfied: When confronted with demand τ', the Home government believes that the Foreign government's type is lower than γ^{sup} with probability 0. As $\tau' > T(\gamma^{\text{sup}})$, the Home government concedes to demand τ' (off the equilibrium path). As $T^*(\gamma_\varepsilon) > \tau'$, this implies that demanding τ' is a profitable deviation for the type-γ^{sup} Foreign government, giving the desired contradiction. As a consequence, Γ_τ is either a singleton or an empty set.

(2) Suppose γ_τ is the unique type that makes a successful demand τ in some equilibrium. Bayesian updating implies that demand τ fully reveals the type of the Foreign government. Therefore, $T(\gamma_\tau) \leq \tau \leq T^*(\gamma_\tau)$ — otherwise either the Home government or the type-γ_τ Foreign government could profitably deviate from their equilibrium strategies. From our assumption that $T^*(\underline{\gamma}) < T(\underline{\gamma})$, this in turn implies that $\gamma_\tau \neq \underline{\gamma}$. Now take any type $\gamma < \gamma_\tau$.

By assumption, a trade war occurs when the Foreign government is of type γ (γ_τ is the only type that makes a successful demand). As $T^*(\gamma) > T^*(\gamma_\tau) \geq \tau$, however, the type-$\gamma$ Foreign government strictly prefers τ to a trade war. It could therefore profitably deviate by making the successful demand τ. Combined with (1), this proves that in any equilibrium all types of Foreign government make unsuccessful demands. $\square$

Claim 2: There exists an equilibrium of the game without the IO, in which the Home government never concedes to the Foreign government's demands.

Proof. Let k be a strictly positive number and consider the following strategy profile and beliefs: The type-γ Foreign government demands a tariff $\tau_k(\gamma) \equiv T(\gamma) - k$; the Home government's strategy when confronted with a demand τ is to concede if and only if $\tau \geq T\left(\underline{\gamma}\right)$; it believes that the Foreign government is of type γ with probability 1 when confronted with demand $\tau_k(\gamma)$, for all $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$, and that it is of type $\underline{\gamma}$ when confronted with any other demand.

It is readily checked that the Home government's beliefs satisfy Bayes' rule whenever possible. By Lemma 2, they also satisfy Criterion D1. It also readily checked that, given these beliefs, the Home government's strategy is a best response to the Foreign government's: given its beliefs, accepting any offer below [resp. above] $T\left(\underline{\gamma}\right)$ would make the Home government strictly worse-off [resp. better-off] than triggering a retaliatory trade war. Finally, as the Home government rejects any demand below $T\left(\underline{\gamma}\right)$, the only possible deviation for the Foreign government would be to make a demand $\tau \geq T\left(\underline{\gamma}\right)$. But, as $T\left(\underline{\gamma}\right) > T^*\left(\underline{\gamma}\right) \geq T^*(\gamma)$ for all types $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$, such a deviation would not be profitable. $\square$

A.2 Proof of Proposition 2

We prove Proposition 2 in four steps. Steps 1 and 2 show that, in any equilibrium, either all types of Foreign government successfully demand τ^{io} or they all make unsuccessful demands. Step 3 shows that all types successfully demand τ^{io} in any equilibrium if and only if $\tau^{io} \geq T(F_0)$. Finally, Step 4 shows that a trade war never arises in equilibrium when $\tau^{io} \geq T(F_0)$, and that all types make obtain concession τ^{io} when $\tau^{io} > T^*(\overline{\gamma})$. Finally, Step 5 proves existence of and characterizes equilibria when $\tau^{io} < T(F_0)$, showing that: all types of Foreign government fail to obtain a concession from the Home government; and the latter never complies with the IO ruling — thus completing the proof of the proposition.

Step 1: If the Foreign government makes a successful demand in equilibrium, then this demand must be τ^{io}.

Consider an equilibrium in which a demand τ is successfully made by a nonempty set of Foreign-government types Γ_τ. Let F be the Home government's updated beliefs after receiving this demand. Obviously, τ is the only successful proposal made in equilibrium — otherwise all types making the highest demands could profitably deviate by making the lowest demand. As it is optimal for the Home government to concede to τ, we must have $\tau \geq \tau^{io}$.

Now suppose by contradiction that $\tau > \tau^{io}$. As T^* is a decreasing function and indifferent players prefer agreements over disagreements, the set of types demanding τ must be of the form $\left[\underline{\gamma}, \gamma_0\right]$, with $\gamma_0 \in \left[\underline{\gamma}, \overline{\gamma}\right]$. We distinguish between two different cases:

• *Case 1:* $\gamma_0 > \underline{\gamma}$. In this case, we have $\tau \geq \max\{T(F), \tau^{io}\} \geq T(F) > T(\gamma_0)$. Consider a deviation from τ to $\tau' \in (\max\{T(\gamma_0), \tau^{io}\}, \tau)$ — the same argument as in that proof can be used to show that τ' can only be made off the equilibrium path. By Lemma 3, reasonable beliefs F' must assign zero probability to the event $\{\gamma < \gamma_0\}$ following demand τ'. This implies that $T(F') \leq T(\gamma_0) < \tau'$, which in turn implies that demand $\tau' < \tau$ would be successful. By definition of a PBE, this is impossible: all types in $\left[\underline{\gamma}, \overline{\gamma}\right]$ can profitably deviate.

• *Case 2:* $\gamma_0 = \underline{\gamma}$. In this case, demand τ reveals that the Foreign government's type is $\underline{\gamma}$. As it is optimal for the Home government to concede to τ, we must have $\tau \geq \max\left\{T\left(\underline{\gamma}\right), \tau^{io}\right\} \geq T\left(\underline{\gamma}\right) > T^*\left(\underline{\gamma}\right)$. But this implies that the type-$\underline{\gamma}$ Foreign government could profitably deviate by making an unacceptable demand $\tau' < \tau^{io}$ (whether this leads to compliance with τ^{io} or with a trade war, it ends up strictly better off).

Step 2: In any equilibrium, either all types of Foreign government successfully demand τ^{io} or they all make unsuccessful demands.

To prove this statement, we will show that, in any equilibrium, if some type successfully demands τ^{io} then all types do (step 1). We proceed by contradiction: Suppose that a nonempty subset of types $\Gamma^{io} \neq \left[\underline{\gamma}, \overline{\gamma}\right]$ make the only successful demand τ^{io} in some equilibrium. As $T^*(\gamma)$ is a strictly decreasing function (and indifferent types are assumed to prefer a successful over an unsuccessful demand), Γ^{io} must be of the form $\left[\underline{\gamma}, \gamma_0\right]$ with $\gamma_0 \geq \underline{\gamma}$.

Let F represent the Home government's beliefs when it receives demand τ^{io}. As it concedes to τ^{io} in equilibrium, $\tau^{io} \geq \max\{T(F), \tau^{io}\} \geq T(F)$. From our initial assumption, there must be a type γ' outside $\left[\underline{\gamma}, \gamma_0\right]$ which makes an unsuccessful demand, say τ', in equilibrium. Bayesian updating implies that the Home government's beliefs assign zero probability to the event $\{\gamma \leq \gamma_0\}$ following demand τ'. As $T(\cdot)$ is strictly decreasing in γ, this in turn implies that $\tau^{io} \geq T(F) \geq T(\gamma_0) \geq T(F')$ where F' represents the Home government's beliefs following demand τ'. Hence, the Home government complies with the IO ruling after rejecting demand τ' in this equilibrium, leaving the type-γ' Foreign government indifferent between its unsuccessful equilibrium demand τ' and the successful demand τ^{io}. According to our indifference condition, it should then demand τ^{io} instead of τ'.

Step 3: There is an equilibrium in which all types make a successful demand if and only if $T(F_0) \leq \tau^{io}$.

Necessity. If all types of Foreign government demand τ^{io} in equilibrium, then the Home government's beliefs when receiving this demand are given by F_0. As a consequence, we must have $\tau^{io} \geq \max\{T(F_0), \tau^{io}\} \geq T(F_0)$.

Sufficiency. Suppose that $T(F) \leq \tau^{io}$, and consider the following strategy profile and beliefs: All types of Foreign government demand τ^{io}; the Home government concedes to (multilateral) demand τ if and only if $\tau \geq \tau^{io}$, and always accepts the IO's ruling; it maintains its initial beliefs F_0 if it receives demand τ^{io}, and believes that the Foreign government is of type $\overline{\gamma}$ otherwise.

As $\tau^{io} \geq T(F_0) > T(\overline{\gamma})$, the Home government's beliefs ensure that it is always optimal for it to comply with the IO ruling and to concede to demand $\tau \geq \tau^{io}$ from the Foreign government. Anticipating that it will get payoff $W(\tau^{io}, 0)$ if it does not concede to the Foreign government's demand, it is also optimal for the Home government not to concede to any $\tau < \tau^{io}$.

Given the Home government's strategy, the Foreign government has two options: (i) to make a successful demand $\tau \geq \tau^{io}$ and thus get a payoff of $W^*(\tau, 0, \gamma)$; or (ii) to make an unsuccessful demand and thus get a payoff of $W^*(\tau^{io}, 0, \gamma)$. As $W^*(\cdot, 0, \gamma)$ is a strictly decreasing function for all $\gamma \in \Gamma$, demanding τ^{io} is the best strategy for any type of Foreign government.

Finally, it is readily checked that the Home government's beliefs satisfy Bayes' rule whenever possible. Moreover, by Lemma 4, they are reasonable.

Step 4: If $\tau^{io} \geq T(F_0)$, then a trade war never arises in equilibrium. In addition, if $\tau^{io} > T^*(\overline{\gamma})$ then all types of Foreign government make unsuccessful demands in any equilibrium.

Suppose that $\tau^{io} \geq T(F_0)$. To prove the statements above, we must first show that the Home government complies with the IO ruling whenever it rejects a demand from the Foreign government on the equilibrium path. To this end, consider an equilibrium — say σ — in which some type of Foreign government makes an unsuccessful demand. From Step 2, this implies that all types make unsuccessful demands. Let T^σ be the set of demands made by all types of Foreign government in σ, and let $\{\Gamma_\tau\}_{\tau \in T^\sigma}$ be a partition of $[\underline{\gamma}, \overline{\gamma}]$ such that all types in Γ_τ demand τ in equilibrium. Suppose first that the Home government rejects the IO ruling after rejecting any demand $\tau \in T^\sigma$. Letting F_τ denote the Home government's beliefs following demand τ, this would imply that $\tau^{io} < T(F_\tau)$ for all $\tau \in T^\sigma$; contradicting our assumption that $\tau^{io} \geq T(F_0)$.

Suppose now that T^σ can be partitioned into two nonempty, disjoint subsets T_1 and T_2 such that the Home government always concedes [resp. does not always concede] to the IO ruling after rejecting any $\tau \in T_1$ [resp. any $\tau \in T_2$]. In particular, observe that if a type γ prefers the IO ruling τ^{io} to a trade war, then so do all types $\gamma' < \gamma$ (recall that T^* is a strictly decreasing function). As σ is an equilibrium, no type that makes a demand in T_1 can profitably deviate by mimicking a type that makes a demand in T_2, and vice versa. This implies that there exists a threshold type γ_0 such that all types smaller [resp. larger] than γ_0 belong to T_1 [resp. to T_2]. Thus in turn implies that the Home government learns that the

Foreign government's type is lower [resp. greater] than γ_0 when it receives a demand $\tau_1 \in T_1$ [resp. a demand $\tau_2 \in T_2$]. As $\widetilde{W}(\gamma)$ is a strictly decreasing function, Bayesian updating then implies that $\mathbb{E}_{F_{\tau_1}}\left[\widetilde{W}(\gamma)\right] \geq \widetilde{W}(\gamma_0) \geq \mathbb{E}_{F_{\tau_2}}\left[\widetilde{W}(\gamma)\right]$ for all $\tau_1 \in T_1$ and all $\tau_2 \in T_2$. However, in equilibrium, the Home government prefers τ^{io} to a trade war after rejecting τ_1 and (strictly) prefers a trade war to τ^{io} after rejecting τ_2; that is

$$W\left(\tau^{io}, 0\right) \geq \mathbb{E}_{F_{\tau_1}}\left[\widetilde{W}(\gamma)\right] \geq \mathbb{E}_{F_{\tau_2}}\left[\widetilde{W}(\gamma)\right] > W\left(\tau^{io}, 0\right) \ ,$$

which is of course impossible. We have thus established that, in equilibrium σ, the Home government always complies with the IO ruling after rejecting any $\tau \in T^{\sigma}$.

Now suppose that $\tau^{io} \geq T(F_0)$ and $\tau^{io} > T^*(\overline{\gamma})$, and that there is an equilibrium in which some (and therefore all) types make unsuccessful demands. From Step 3 and the argument in the previous paragraph, we know that all equilibrium demands lead to the implementation of τ^{io}. This implies that demand τ^{io} must be unsuccessful in equilibrium; otherwise all Foreign government's types would be indifferent between their equilibrium unsuccessful demands and τ^{io} (and would therefore choose to demand τ^{io}). This in turn implies that demand τ^{io} is followed by a trade war; otherwise the Home government would be indifferent between conceding and not conceding to τ^{io} and, therefore, would choose to concede. By definition of an equilibrium, no type of Foreign government can profitably deviate by demanding τ^{io} (thus triggering a trade war); that is: $W^*(\tau^{io}, 0, \gamma) \geq \widetilde{W}^*(\gamma)$ or, equivalently, $\tau^{io} \leq T^*(\gamma)$ for all $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$. As $T^*(\cdot)$ is a strictly decreasing function, this is equivalent to $\tau^{io} \leq T^*(\overline{\gamma})$, thus contradicting the assumption that $\tau^{io} > T^*(\overline{\gamma})$.

Step 5: If $\tau^{io} < T(F_0)$, then: (i) all types of Foreign government fail to obtain a concession from the Home government; and (ii) the latter never complies with the IO ruling. Such an equilibrium exists.

Suppose that $\tau^{io} < T(F_0)$. Part (i) is an immediate consequence of Steps 2 and 3. To prove part (ii), suppose that there is an equilibrium in which a nonempty set of types of Foreign government, say Γ^{io}, make unsuccessful demands followed by compliance with τ^{io}. Observe that $\Gamma^{io} \neq \left[\underline{\gamma}, \overline{\gamma}\right]$: a nonempty subset of types must make unsuccessful demands followed by trade wars. To see this, suppose instead that all types' demands lead the Home government to comply with τ^{io}. Letting F_{τ} denote the Home government's beliefs following demand τ, this would imply that $\tau^{io} \geq T(F_{\tau})$ for all on-the-equilibrium-path demands τ and, therefore, that $\tau^{io} \geq T(F_0)$; thus contradicting $\tau^{io} < T(F_0)$.

By definition of an equilibrium, $\gamma \in \Gamma^{io}$ if and only if $T^*(\gamma) \geq \tau_0$ (otherwise γ could profitably deviate by mimicking a type outside Γ^{io}). As T^* is a strictly decreasing function, there exists a threshold type $\gamma_0 < \overline{\gamma}$ such that $\Gamma^{io} = \left[\underline{\gamma}, \gamma_0\right]$. This implies that, when the Home government receives a demand τ' from a type outside Γ^{io}, its updated beliefs F' assign a zero probability to the event $\{\gamma \leq \gamma_0\}$. Hence, $T(F') \leq T(F_\tau) \leq \tau^{io}$ for any demand τ made by a type in Γ^{io}. But this implies that the Home government should comply with τ^{io} after rejecting demand τ'. We have thus established that the Home government never concedes to the Foreign government's demands, and never complies with the IO ruling in an equilibrium.

We now have to prove that such an equilibrium exists. We argue that the following strategy profile and system of beliefs constitute an equilibrium: All types of foreign government demand τ^{io}; the Home government concedes to demand τ if and only if $\tau \geq T(\underline{\gamma})$; it never complies with the IO ruling; and it believes that the Foreign government's type is $\underline{\gamma}$ if the latter demands $\tau \neq \tau^{io}$, and maintains its initial beliefs F_0 otherwise.

To see that the Foreign government does not have a profitable deviation, observe that it could only change the equilibrium outcome (i.e. a trade war) by making a demand $\tau \geq T(\underline{\gamma})$. As $T^*(\gamma) \leq T^*(\underline{\gamma}) < T(\underline{\gamma})$, this would be unprofitable to all Foreign government's types $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

As the Home government's beliefs are F_0 when it receives demand τ^{io} and $\tau^{io} < T(F_0)$, it is optimal for it not comply with ruling τ^{io} after rejecting demand τ^{io}. This in turn implies that it is also optimal to reject demand τ^{io}. When it receives a demand $\tau \neq \tau^{io}$, the Home government believes that the Foreign government is of type $\underline{\gamma}$. As $T(\underline{\gamma}) > T(F_0) > \tau^{io}$, it is optimal for the Home government to trigger a trade war by rejecting the IO ruling. This in turn implies that it is a best response to concede to demand τ if and only if $\tau \geq T(\underline{\gamma})$.

Finally, it is readily checked that the Home government's beliefs satisfy Bayes' rule whenever possible. Moreover, Lemma 5 shows that they also satisfy criterion D1.

A.3 Proof of Proposition 3

We prove Proposition 3 in two steps:

Step 1: There exists an equilibrium in which all types of Foreign government coerce unilaterally and fail to obtain a concession.

Let κ be a strictly positive number and consider the following strategy profile and beliefs: The type-γ Foreign government makes unilateral demand $\tau_\kappa(\gamma) \equiv T(\gamma) - \kappa$; the Home government concedes to a unilateral demand τ if and only if $\tau \geq T(\underline{\gamma})$; concedes to a multilateral demand τ if and only if $\tau \geq \max\{T(\underline{\gamma}), \tau^{io}\}$; it complies with the IO ruling if and only if $\tau^{io} \geq T(\underline{\gamma})$; it believes that the Foreign government is of type γ when it is confronted with unilateral demand $\tau_\kappa(\gamma)$, for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$, and that it is of type $\underline{\gamma}$ when confronted with any other demand.

To see that these strategy profile and system of beliefs constitute an equilibrium, note first that the Home government's beliefs are consistent with Bayes' rule whenever possible. Moreover, Lemma 6, shows that they are reasonable. The Foreign government can only change the outcome by making either a unilateral demand $\tau \geq T\left(\underline{\gamma}\right)$ or a multilateral demand $\tau \geq \max\left\{T\left(\underline{\gamma}\right), \tau^{io}\right\}$. As $T^{*}(\gamma) \leq T^{*}\left(\underline{\gamma}\right) < T\left(\underline{\gamma}\right) \leq \max\left\{T\left(\underline{\gamma}\right), \tau^{io}\right\}$, however, such deviations can only make it worse off. Finally, it is readily checked that, given its beliefs, the Home government's strategy is a best response to the Foreign government's.

Step 2: In any equilibrium, a trade war arises with a probability of one.

To prove this statement, we will establish in turn that in equilibrium: (i) if all types of Foreign government make unilateral demands, then the Home government never concedes; (ii) if all types make multilateral demands, then the Home government never concedes to those demands and never complies with the IO ruling; and (iii) if some types coerce unilaterally and others multilaterally, then all their demands are unsuccessful and lead to a trade war.

(i) If all types coerce unilaterally in equilibrium, then by the same argument as in Proposition 1 they all fail to obtain a concession (all deviations available in the game without IO are still available). Hence, a tarde war ensues for all possible realizations of the Foreign government's type.

(ii) Consider an equilibrium in which all types of Foreign government coerce multilaterally, and suppose (by contradiction) that some type's demand does not lead to a trade war. By the same argument as in Proposition 2, this implies that $\tau^{io} \geq T\left(F_0\right)$ and that all types' demands lead to the implementation of τ^{io}. This in turn implies that $\tau^{io} \leq T^{*}\left(\overline{\gamma}\right)$ — otherwise the type-$\overline{\gamma}$ Foreign government could profitably deviate by making an unacceptable unilateral demand $\tau' < T\left(\overline{\gamma}\right)$. Lemma 7 shows that, in such a case, reasonable beliefs must assign a probability of 1 to type $\overline{\gamma}$ following any (off-the-equilibrium-path) unilateral demand $\tau' < \tau^{io}$.

Now consider a deviation to unilateral demand $\tau' \in \left(T\left(\overline{\gamma}\right), \tau^{io}\right)$ (observe that $T\left(\overline{\gamma}\right) < T\left(F_0\right) \leq \tau^{io}$). As the Home government believes that this demand emanates from the type-$\overline{\gamma}$ government, it should concede to it. This makes the deviation profitable for all types of Foreign government.

(iii) Consider an equilibrium in which $\left[\underline{\gamma}, \overline{\gamma}\right]$ can be partitioned into two nonempty subsets Γ_1 and Γ_2 such that all types in Γ_1 [resp. Γ_2] coerce multilaterally [resp. unilaterally]. Proceeding by contradiction, assume that in this equilibrium, a trade war is avoided for some realization of the Foreign government's type. By the same argument as in Proposition 1, all types in Γ_2 fail to obtain a concession; so that a trade war occurs if $\gamma \in \Gamma_2$. Therefore, the types avoiding a trade war must be in Γ_1. By the same argument as in Proposition 2, tariff τ^{io} must then be implemented whenever the Foreign government's type is in Γ_1.

By definition of an equilibrium, types in Γ_1 cannot profitably deviate by mimicking types in Γ_2, and vice versa. As $T^{*}(\gamma)$ is strictly decreasing in γ, this implies that there must be a type $\gamma_0 \in \left(\underline{\gamma}, \overline{\gamma}\right)$ such that $\gamma_0 = \left(T^{*}\right)^{-1}\left(\tau^{io}\right)$ and $\Gamma_1 = \left[\underline{\gamma}, \gamma_0\right]$. ($\gamma_0 > \underline{\gamma}$ because $T^{*}\left(\underline{\gamma}\right) < T\left(\underline{\gamma}\right)$; and $\gamma_0 < \overline{\gamma}$ because by assumption $\Gamma_2 \neq \emptyset$.) We distinguish between two different cases:

(a) If $T^{*}\left(\gamma_0\right) \leq T\left(\gamma_0\right)$, then we have $\tau^{io} = T^{*}\left(\gamma_0\right) \leq T\left(\gamma_0\right) < T(\gamma)$ for all $\gamma < \gamma_0$. This implies that there must a demand τ emanating from some type, or some subset of types, in Γ_1 such that the Home government's updated beliefs F_τ satisfy $T(F_\tau) > \tau^{io}$. This in turn implies that it is optimal for the Home government to reject both demand τ and ruling τ^{io} in order to trigger a trade war — this is a contradiction.

(b) If $T^*(\gamma_0) > T(\gamma_0)$, then $\tau^{io} > T(\gamma_0)$. Consider a unilateral demand $\tau' \in (T(\gamma_0), \tau^{io})$. Observe that this demand is only made off the equilibrium path: types $\gamma \leq \gamma_0$ make multilateral demands, and types $\gamma > \gamma_0$ make unsuccessful demands (as $T(\gamma) < T(\gamma_0) < \tau'$ for all $\gamma > \gamma_0$, the Home government would concede to τ' if it emanated from types $\gamma > \gamma_0$ in equilibrium). Furthermore, Lemma 8 shows that the Home government's beliefs F' when it receives unilateral demand τ' must assign zero probability to the event $\{\gamma < \gamma_0\}$; so that $T(F') \leq T(\gamma_0) < \tau'$ (recall that $T(\gamma)$ is strictly decreasing in γ). This implies that if some type of Foreign government deviated to unilateral demand τ', then the Home government would concede. As $\tau' < \tau^{io} = T^*(\gamma_0)$, this deviation is profitable to all types in Γ_1 — this is again a contradiction.

B Reasonable Beliefs

As explained in the main text, there is a unique equilibrium in the trade-war continuation game, in which the type-γ Foreign government always chooses $\tilde{\tau}^*(\gamma)$ and the Home government always chooses $\tilde{\tau}$. Similarly, the Home government's decision of whether or not to concede when confronted with the international organization's ruling τ^{io} is uniquely determined by sequential rationality.

However, multiplicity arises in the earlier stages of the model where, anticipating equilibrium moves in subsequent subgames, governments play a signaling game. In order to rule out PBEs supported by "unreasonable" beliefs off the equilibrium path, we concentrate on pure strategy equilibria that satisfy Cho' and Kreps' (1987) criterion D1 (see Fudenberg and Tirole 1991, and Ramey 1996).

Fix an equilibrium, and let $\widehat{W}^*(\gamma)$ be the payoff of the type-γ Foreign government in this equilibrium. According to criterion D1, what types of Foreign government can reasonably be thought to choose an off-the-equilibrium-path demand τ'? Let $MBR(F, \tau')$ be the Home government's set of mixed best responses to τ' when it has beliefs F about the Foreign government's type. Next, define $D_F(\gamma, \tau')$ be the set of mixed best responses $\alpha \in MBR(F, \tau')$ that make type γ strictly prefer τ' to its equilibrium strategy — that is, the type-γ Foreign government's expected payoff when the Home government adopts any strategy in $D_F(\gamma, \tau')$ is strictly greater than $\widehat{W}^*(\gamma)$. Thus, $D(\gamma, \tau') \equiv \bigcup_F D_F(\gamma, \tau')$ can be interpreted as the set of Home government's responses that make the type-γ Foreign government willing to deviate to τ'. The set $D^0(\gamma, \tau')$ of mixed best responses that make the type-γ Foreign government exactly indifferent is defined analogously. Accordingly, a type γ is deleted following demand τ' under criterion D1 if there is another type γ' such that $[D(\gamma,\tau') \cup D^0(\gamma,\tau')] \subset D(\gamma',\tau')$. In words, if the set of Home government's responses that make type γ willing to deviate to τ' is strictly smaller than the set of best responses that make type γ' willing to deviate, than the Home government should believe that type γ' is infinitely more likely to deviate to τ' than type γ is.

Now we establish a series of lemmata that are used in the proofs of our main results.

Coercion without the IO

Lemma 1. *Consider an equilibrium in which some some subset of types of the form $[\underline{\gamma},\gamma^{\sup}]$, with $\gamma^{\sup} > \underline{\gamma}$, obtain a concession τ. Reasonable beliefs assign zero probability to all types $\gamma < \gamma^{\sup}$ following any (off-the-equilibrium-path) demand $\tau' \in (T(\gamma^{\sup}),\tau)$.*

Proof. Consider a deviation to demand $\tau' \in (T(\gamma^{\sup}),\tau)$. By definition of an equilibrium, all types $\gamma \in [\underline{\gamma},\gamma^{\sup}]$ prefer successful demand τ to a trade war; that is: $W^*(\tau,0,\gamma) \geq \widetilde{W}^*(\gamma)$ for all $\gamma \in [\underline{\gamma},\gamma^{\sup}]$. In addition, $\tau' < \tau$ implies that:

$$\widetilde{W}^*(\gamma) \leq W^*(\tau,0,\gamma) < W^*(\tau',0,\gamma) \ \text{, for all } \gamma \in \Gamma_\tau \ . \tag{7}$$

Take an arbitrary type $\gamma \in [\underline{\gamma},\gamma^{\sup})$. The Home government's mixed best response α makes the type-γ foreign government prefer τ' to its equilibrium demand τ if and only if:

$$\alpha W^*(\tau',0,\gamma) + (1-\alpha)\widetilde{W}^*(\gamma) \geq W^*(\tau,0,\gamma) \ .$$

(Our restrictions on τ' ensure that any $\alpha \in [0,1]$ is a best response for some beliefs.) This inequality can be rewritten as

$$\begin{aligned}\alpha \geq \bar{\alpha}(\gamma) &\equiv \frac{W^*(\tau,0,\gamma) - \widetilde{W}^*(\gamma)}{W^*(\tau',0,\gamma) - \widetilde{W}^*(\gamma)} \\ &= \frac{w^*(0) - w^*(\tilde{\tau}^*(\gamma)) + \gamma[\Pi_1^*(0) - \Pi_1^*(\tilde{\tau}^*(\gamma))] + \Pi_2^*(\tau) - \Pi_2^*(\tilde{\tau})}{w^*(0) - w^*(\tilde{\tau}^*(\gamma)) + \gamma[\Pi_1^*(0) - \Pi_1^*(\tilde{\tau}^*(\gamma))] + \Pi_2^*(\tau') - \Pi_2^*(\tilde{\tau})} \ .\end{aligned}$$

The inequalities in (7) guarantee that $\bar{\alpha}(\gamma) \in [0,1)$ for all $\gamma \in [\underline{\gamma},\gamma^{\sup}]$. Furthermore, as $\Pi_2^*(\tau') > \Pi_2^*(\tau)$, the sign of the derivative of $\bar{\alpha}$ is the same as the sign of

$$\frac{d}{d\gamma}[w^*(0) - w^*(\tilde{\tau}^*(\gamma)) + \gamma[\Pi_1^*(0) - \Pi_1^*(\tilde{\tau}^*(\gamma))]] = \Pi_1^*(0) - \Pi_1^*(\tilde{\tau}^*(\gamma)) < 0 \ .$$

(The equality follows from the Envelope Theorem: $\tilde{\tau}^*(\gamma)$ is the maximizer of $w^*(\cdot) + \gamma\Pi_1^*(\cdot)$.) Hence, $\bar{\alpha}$ is strictly decreasing.

This implies that, for any $\gamma' \in [\underline{\gamma},\gamma^{\sup}]$ such that $\gamma' > \gamma$, $D(\gamma,\tau') \cup D^0(\gamma,\tau') = [\bar{\alpha}(\gamma),1] \subset (\bar{\alpha}(\gamma'),1] = D(\gamma',\tau')$. Criterion D1 then requires that, when confronted with demand τ', the Home government believes that the Foreign government is of type γ with probability 0. As γ

was taken arbitrarily in $\left[\underline{\gamma},\gamma^{\sup}\right)$, this establishes the lemma. $\square$

Lemma 2. *Consider an equilibrium in which a trade war ensues after every type's demand. Beliefs which assign a probability of 1 to type $\underline{\gamma}$ following any (off-the-equilibrium-path) unilateral demand τ' are reasonable.*

Proof. If $\tau' < T(\overline{\gamma})$, then the lemma is trivial: the only best response for the Home government is to reject demand τ'. This implies that $D(\gamma,\tau') = \emptyset$ for all types $\gamma \in \left[\underline{\gamma},\overline{\gamma}\right]$ and, therefore, that it is impossible to eliminate type $\underline{\gamma}$.

If $\tau' \in \left[T(\overline{\gamma}), T\left(\underline{\gamma}\right)\right]$, then any $\alpha \in [0,1]$ may be a best response. As all types of foreign government make unsuccessful demands in equilibrium, we have

$$D(\gamma,\tau') = \begin{cases} (0,1] & \text{if } \tau' < T^*(\gamma) \ , \\ \emptyset & \text{otherwise,} \end{cases}$$

for all $\gamma \in \left[\underline{\gamma},\overline{\gamma}\right]$. As T^* is strictly decreasing, this implies that $D\left(\underline{\gamma},\tau'\right) \supseteq D(\gamma,\tau')$ for all $\gamma \in \left[\underline{\gamma},\overline{\gamma}\right]$. Therefore, beliefs which assign a probability of 1 to type $\underline{\gamma}$ following any (off-the-equilibrium-path) demand $\tau' \in \left[T(\overline{\gamma}), T\left(\underline{\gamma}\right)\right]$ are reasonable.

Finally, if $\tau' > T\left(\underline{\gamma}\right)$ then $D(\gamma;\tau',u) = \{1\}$ if $\tau' < T^*(\gamma)$, and $D(\gamma;\tau') = \emptyset$. By the same argument as above, it is impossible to eliminate type $\underline{\gamma}$ using criterion D1. $\square$

Coercion with Full Commitment to the IO

Lemma 3. *Consider an equilibrium in which a set of types of the form $\left[\underline{\gamma},\gamma_0\right]$, with $\gamma_0 \in \left(\underline{\gamma},\overline{\gamma}\right]$, make a successful demand $\tau > \tau^{io}$. Reasonable beliefs must assign zero probability to all types $\gamma < \gamma_0$ following any (off-the-equilibrium-path) demand $\tau' \in (\tau^{io},\tau)$.*

Proof. Consider a deviation $\tau' \in (\tau^{io},\tau)$, and let F' be the Home government's beliefs following this demand. If F' makes the Home government indifferent between conceding and not conceding to τ', then $W(\tau^{io},0) < W(\tau',0) = W(T(\gamma),0)$ — so that a trade war ensues when the Home government rejects τ'. In addition, all types in $\left[\underline{\gamma},\gamma_0\right]$ must prefer τ to a trade war; otherwise we would have $W^*(\tau,0,\gamma) < \min\left\{W^*(\tau',0,\gamma), \widetilde{W}^*(\gamma)\right\}$ for some type $\gamma \in \left[\underline{\gamma},\gamma_0\right]$ (which could the profitably deviate by making an unacceptable offer $\tau'' < \tau^{io}$). These observations imply that we can use the same argument as in the proof of Lemma 1 to obtain the result. $\square$

Lemma 4. *Suppose that, in equilibrium, all types of Foreign government successfully demand τ^{io}. Beliefs which assign a probability of 1 to type $\overline{\gamma}$ following any (off-the-equilibrium-path) multilateral demand $\tau' \neq \tau^{io}$ are reasonable.*

Proof. Take an arbitrary (off-the-equilibrium-path) multilateral demand $\tau' \neq \tau^{io}$. Throughout this proof, the Home government's updated beliefs following demand τ' are denoted by F'.

Suppose first that $\tau' < \tau^{io}$. In this case, it is never a best response for the Home government to concede to τ^{io}; so that $MBR\,(F',\tau') = \{0\}$. If its beliefs F' are such that $\tau^{io} \geq T\,(F')$ (i.e., it concedes to the IO ruling τ^{io}), then any type of Foreign government is indifferent between its successful equilibrium demand τ^{io} and the unsuccessful demand τ'; so that $D_{F'}\,(\gamma,\tau') = \emptyset$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$. If its beliefs F' are such that $\tau^{io} < T\,(F')$ (i.e., it does not concedes to the international organization's ruling, thus triggering a trade war), then the type-γ Foreign government strictly prefers demanding τ' over demanding τ^{io} if and only if: $\widetilde{W}^{*}(\gamma) > W^{*}\,(\tau^{io},0,\gamma)$ or, equivalently, $T^{*}(\gamma) < \tau^{io}$. This implies that

$$D_{F'}\,(\gamma,\tau') = \begin{cases} \{0\} & \text{if } T^{*}(\gamma) < \tau^{io} \ , \\ \emptyset & \text{otherwise,} \end{cases}$$

As $T^{*}(\gamma)$ is strictly decreasing in γ, this implies that $D_{F'}\,(\gamma,\tau') \subseteq D_{F'}\,(\overline{\gamma},\tau')$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$. This in turn implies that $D\,(\gamma,\tau') \subseteq D\,(\overline{\gamma},\tau')$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$, thus proving that beliefs which assign a probability of 1 to type $\overline{\gamma}$ following any (off-the-equilibrium-path) demand $\tau' < \tau^{io}$ are reasonable.

Suppose now that $\tau' > \tau^{io}$. If the Home government's beliefs F' are such that $\tau^{io} \geq T(F')$ (i.e., it concedes to the international organization's ruling τ^{io}), then any type of Foreign government is always worse off making the unsuccessful demand τ'; so that $D_{F'}\,(\gamma,\tau') = \emptyset$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$. If its beliefs F' are such that $\tau^{io} < T(F')$ (i.e., it does not concede to the international organization's ruling), then we must distinguish between three different cases:

(i) If F' is such that $T(F') < \tau'$, then the unique best response for the Home government is to accept τ' with a probability of 1: $MBR\,(F',\tau') = \{1\}$. As $\tau' > \tau$, any type of Foreign government is worse off. Hence, $D_{F'}\,(\gamma,\tau') = \emptyset$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$.

(ii) If F' is such that $T(F') = \tau'$, then the Home government is indifferent between conceding and not conceding to τ': $MBR\,(F',\tau') = [0,1]$. An $\alpha \in [0,1]$ makes the type-γ Foreign government (strictly) prefer τ' to its equilibrium demand τ^{io} if and only if:

$$\alpha W^{*}\,(\tau',0,\gamma) + (1-\alpha)\widetilde{W}^{*}(\gamma) > W^{*}\left(\tau^{io},0,\gamma\right) \ .$$

This implies that

$$D_{F'}\,(\gamma,\tau') = \begin{cases} [0,\bar{\alpha}(\gamma)) & \text{if } \widetilde{W}^{*}(\gamma) > W^{*}\,(\tau^{io},0,\gamma) \ , \\ \emptyset & \text{otherwise,} \end{cases}$$

where

$$\bar{\alpha}(\gamma) \equiv \frac{w^{*}\,(0) - w^{*}\,(\tilde{\tau}^{*}(\gamma)) + \gamma\,[\Pi_1^{*}\,(0) - \Pi_1^{*}\,(\tilde{\tau}^{*}(\gamma))] + \Pi_2^{*}\,(\tau^{io}) - \Pi_2^{*}\,(\tilde{\tau})}{w^{*}\,(0) - w^{*}\,(\tilde{\tau}^{*}(\gamma)) + \gamma\,[\Pi_1^{*}\,(0) - \Pi_1^{*}\,(\tilde{\tau}^{*}(\gamma))] + \Pi_2^{*}\,(\tau') - \Pi_2^{*}\,(\tilde{\tau})} \ .$$

As $\Pi_2^*(\tau') < \Pi_2^*(\tau^{io})$, the sign of the derivative of $\bar{\alpha}$ is the same as the sign of

$$-\frac{d}{d\gamma}\left[w^*(0) - w^*(\tilde{\tau}^*(\gamma)) + \gamma\left[\Pi_1^*(0) - \Pi_1^*(\tilde{\tau}^*(\gamma))\right]\right] > 0$$

(The argument is the same as in the proof of Lemma 1).) Hence, $\bar{\alpha}$ is strictly increasing. This implies that $D_{F'}(\gamma, \tau') \subseteq D_{F'}(\overline{\gamma}, \tau')$ for all $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$.

(iii) If F' is such that $T(F') > \tau'$, then the unique best response for the Home government is to reject τ': $MBR(F', \tau') = \{0\}$. As a consequence, the type-γ Foreign government is better-off demanding τ' rather than τ^{io} if and only of it prefers a trade war over agreement on τ^{io} or, equivalently, $T^*(\gamma) < \tau^{io}$. Hence,

$$D_{F'}(\gamma, \tau') = \begin{cases} \{0\} & \text{if } T^*(\gamma) < \tau^{io}, \\ \emptyset & \text{otherwise.} \end{cases}$$

As $T^*(\gamma)$ is strictly decreasing in γ, this implies that, for any beliefs F', $D_{F'}(\gamma, \tau') \subseteq D_{F'}(\overline{\gamma}, \tau')$ for all $\gamma \in \Gamma$. We have thus proved that the latter relation is true for all possible beliefs and, therefore, that $D(\gamma, \tau') \subseteq D(\overline{\gamma}, \tau')$ for all $\gamma \in \Gamma$. As a result, beliefs which assign a probability of 1 to type $\overline{\gamma}$ following any (off-the-equilibrium-path) demand $\tau' > \tau^{io}$ are reasonable. This completes the proof of the lemma. □

Lemma 5. *Consider an equilibrium in which all types of Foreign government unsuccessfully make demand τ^{io}, following which the domestic government does not comply the IO ruling. Beliefs that assign a probability of 1 to type $\underline{\gamma}$ following any (off-the-equilibrium-path) demand $\tau' \neq \tau^{io}$ are reasonable.*

Proof. To prove the lemma, it suffices to show that $D(\gamma, \tau') \subseteq D\left(\underline{\gamma}, \tau'\right)$ for all $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$ (so that $\underline{\gamma}$ cannot be eliminated).

Suppose first that $\tau' < \tau^{io}$. In this case, it is never a best response for the domestic government for any beliefs F it may have (it receives $\max\left\{W(\tau^{io}, 0), \mathbb{E}_F\left[\widetilde{W}(\gamma)\right]\right\} \geq W(\tau^{io}, 0) > W(\tau', 0)$ by rejecting τ'); so that $MBR(F, \tau') = \{0\}$. If its beliefs F are such that $\tau^{io} \geq T(F)$ (i.e., it concedes to the IO ruling τ^{io}), then the type-γ foreign government strictly prefers unsuccessful demand τ' to the equilibrium trade war if and only if $\tau^{io} < T^*(\gamma)$; so that

$$D_F(\gamma, \tau') = \begin{cases} \{0\} & \text{if } T^*(\gamma) > \tau^{io}, \\ \emptyset & \text{otherwise.} \end{cases}$$

As $T^*(\gamma)$ is strictly decreasing in γ, this implies that $D_F(\gamma, \tau') \subseteq D_F\left(\underline{\gamma}, \tau'\right)$ for all $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$.

If its beliefs F are such that $\tau^{io} < T(F)$ (i.e., it does not concedes to the IO ruling, thus triggering a trade war), then any type of foreign government is indifferent between its unsuccessful equilibrium demand τ^{io} and the unsuccessful demand τ'; so that $D_F(\gamma, \tau') = \emptyset$

for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

Suppose now that $\tau' > \tau^{io}$. If the domestic government's beliefs F are such that $\tau^{io} \geq T(F)$ (i.e., it concedes to the IO ruling τ^{io}), then any type of foreign government is always worse off making the successful demand τ'; so that $D_F(\gamma, \tau') = \emptyset$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$. If its beliefs F are such that $\tau^{io} < T(F)$ (i.e., it does not concede to the IO ruling), then we must distinguish between three different cases:

(i) If F is such that $T(F) < \tau'$ — so that $\widetilde{W}(\gamma) < W(\tau', 0)$ — then the unique best response for the domestic government is to accept τ' with a probability of 1: $MBR(F, \tau') = \{1\}$. Therefore, the type-γ foreign government is strictly better off demanding τ' if and only if $\tau' < T^*(\gamma)$. As $T^*(\gamma)$ is strictly decreasing in γ, this implies that $D_F(\gamma, \tau') \subseteq D_F(\underline{\gamma}, \tau')$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

(ii) If F is such that $T(F) = \tau'$ — so that $\widetilde{W}(\gamma) = W(\tau', 0)$ — then the domestic government is indifferent between conceding and not conceding to τ': $MBR(F, \tau') = [0, 1]$. As, the type-γ foreign government strictly prefers successful demand τ' to the equilibrium trade war if and only if $\tau^{io} < T^*(\gamma)$, we have

$$D_F(\gamma, \tau') = \begin{cases} (0, 1] & \text{if } T^*(\gamma) > \tau' \ , \\ \emptyset & \text{otherwise.} \end{cases}$$

As $T^*(\gamma)$ is strictly decreasing in γ, this implies that $D_F(\gamma, \tau') \subseteq D_F(\underline{\gamma}, \tau')$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

(iii) If F is such that $T(F) > \tau'$ — so that $\widetilde{W}(\gamma) > W(\tau', 0)$ — then the unique best response for the domestic government is to accept τ' with zero probability: $MBR(F, \tau') = \{0\}$. Therefore, all types of foreign government are indifferent between their equilibrium demand and τ'; so that $D_F(\gamma, \tau') = \emptyset$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

We have thus showed that the following is true for all domestic government's beliefs F: $D_F(\gamma, \tau') \subseteq D_F(\underline{\gamma}, \tau')$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$. This in turn implies that $D_F(\gamma, \tau') \subseteq D_F(\underline{\gamma}, \tau')$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$. It is therefore impossible to eliminate type $\underline{\gamma}$. $\square$

Coercion with Partial Commitment to the IO

Observe that, in this version of the model, the Foreign government makes two choices: a coercion mode and a demand to the Home government. Therefore, a deviation is now of the form (τ', c) where $c \in \{u, m\}$ is the coercion mode adopted by the Foreign government when it deviates — u meaning "unilateral," and m "multilateral."

Lemma 6. *Consider an equilibrium in which all types of Foreign government make unsuccessful unilateral demands. Beliefs that assign a probability of 1 to type $\underline{\gamma}$ following any off-the-equilibrium-path demand are reasonable.*

Proof. We can apply the same argument as in Lemma 2 to show that beliefs assigning a probability of 1 to type $\underline{\gamma}$ following any deviation to a unilateral demand are reasonable. Now consider a deviation to a multilateral demand τ'. Suppose first that $\tau' > \tau^{io}$. Consider first a system of beliefs F such that the Home government complies with the IO ruling; that is, $\tau^{io} \geq T(F)$. In this case, the only best response for the Home government is to accept τ' with a probability of 1: $MBR(F, \tau', m) = \{1\}$. This implies that the type-γ Foreign government strictly prefers demanding τ' over its equilibrium demand if and only if $\tau' < T^*(\gamma)$. Hence,

$$D_F(\gamma, \tau', m) = \begin{cases} \{1\} & \text{if } T^*(\gamma) > \tau' , \\ \emptyset & \text{otherwise.} \end{cases}$$

As $T^*(\gamma)$ is strictly decreasing in γ, this implies that $D_F(\gamma, \tau', m) \subseteq D_F(\underline{\gamma}, \tau', m)$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

Consider now a system of beliefs F such that the Home government does not comply with the IO ruling; that is, $\tau^{io} < T(F)$. We must distinguish between three different situations:

(i) If $\tau' > T(F)$, then $MBR(F, \tau', m) = \{1\}$. We can use then use the same argument as above to obtain that $D_F(\gamma, \tau', m) \subseteq D_F(\underline{\gamma}, \tau', m)$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

(ii) If $\tau' < T(F)$, then $MBR(F, \tau', m) = \{0\}$. Therefore all types of Foreign government are perfectly indifferent between demanding τ' multilaterally and their equilibrium unilateral demand. Hence, $D_F(\gamma, \tau', m) = \emptyset$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

(iii) If $\tau' = T(F)$, then $MBR(F, \tau') = [0, 1]$. In this case,

$$D_F(\gamma, \tau', m) = \begin{cases} (0, 1] & \text{if } T^*(\gamma) > \tau' , \\ \emptyset & \text{otherwise.} \end{cases}$$

As $T^*(\gamma)$ is strictly decreasing in γ, this implies that $D_F(\gamma, \tau', m) \subseteq D_F(\underline{\gamma}, \tau', m)$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

Suppose now that $\tau' = \tau^{io}$. If the Home government's beliefs F are such that $\tau^{io} \geq T(F)$ — i.e. it complies with the IO ruling — then it is indifferent between accepting and rejecting demand τ': $MBR(F, \tau', m) = [0, 1]$. This implies that the type-γ Foreign government strictly prefers demanding τ' over its equilibrium demand if and only if $W^*(\tau', 0, \gamma) = W^*(\tau^{io}, 0, \gamma) > \widetilde{W}^*(\gamma)$ (or $\tau' < T^*(\gamma)$). Hence,

$$D_F(\gamma, \tau', m) = \begin{cases} (0, 1] & \text{if } T^*(\gamma) > \tau' , \\ \emptyset & \text{otherwise.} \end{cases}$$

As $T^*(\gamma)$ is strictly decreasing in γ, this implies that $D_F(\gamma, \tau', m) \subseteq D_F(\underline{\gamma}, \tau', m)$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.

If the Home government's beliefs F are such that $\tau^{io} < T(F)$ — i.e. it does not comply with the IO ruling — then its only best response is to accept $\tau' = \tau^{io}$ with a zero probability:

$MBR(F,\tau') = \{0\}$. Therefore all types of Foreign government are perfectly indifferent between demanding τ' multilaterally and their equilibrium unilateral demand. Hence, $D_F(\gamma,\tau',m) = \emptyset$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$.

Finally, suppose that $\tau' < \tau^{io}$. In this case, it is never a best response for the Home government for any beliefs F it may have (it receives $\max\left\{W(\tau^{io},0), \mathbb{E}_F\left[\widetilde{W}(\gamma)\right]\right\} \geq W(\tau^{io},0) > W(\tau',0)$ by rejecting τ'); so that $MBR(F,\tau',m) = \{0\}$. If its beliefs F are such that $\tau^{io} \geq T(F)$ (i.e., it concedes to the IO ruling τ^{io}), then the type-γ Foreign government strictly prefers unsuccessful demand τ' to the equilibrium trade war if and only if $\tau^{io} < T^*(\gamma)$; so that

$$D_F(\gamma,\tau',m) = \begin{cases} \{0\} & \text{if } T^*(\gamma) > \tau^{io}, \\ \emptyset & \text{otherwise.} \end{cases}$$

As $T^*(\gamma)$ is strictly decreasing in γ, this implies that $D_F(\gamma,\tau',m) \subseteq D_F(\underline{\gamma},\tau',m)$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$.

If its beliefs F are such that $\tau^{io} < T(F)$ (i.e., it does not concedes to the IO ruling, thus triggering a trade war), then any type of Foreign government is indifferent between its unsuccessful equilibrium unilateral demand and the unsuccessful demand τ'; so that $D_F(\gamma,\tau',m) = \emptyset$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$.

We thus established that, for any system of beliefs F, $D_F(\gamma,\tau',m) \subseteq D_F(\underline{\gamma},\tau',m)$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$. Taking the union over all possible beliefs, we obtain $D(\gamma,\tau',m) \subseteq D(\underline{\gamma},\tau',m)$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$. This proves that beliefs that assigns probability 1 to type $\underline{\gamma}$ are reasonable. $\square$

Lemma 7. *Consider an equilibrium in which all types of Foreign government coerce multilaterally, and all their demands are followed by the implementation of $\tau^{io} \leq T^*(\overline{\gamma})$. Reasonable beliefs must assign a probability of 1 to type $\overline{\gamma}$ following any (off-the-equilibrium-path) unilateral demand $\tau' < \tau^{io}$.*

Proof. Consider a deviation to a unilateral demand $\tau' < \tau^{io}$. If the Home government's beliefs, F, are such that its unique best response is to concede to τ' with a probability of 1, then all types of Foreign government are strictly better-off demanding τ' unilaterally: $D_F(\gamma,\tau',u) = \{1\}$ and $D_F^0(\gamma,\tau',u) = \emptyset$ for all $\gamma \in [\underline{\gamma},\overline{\gamma}]$. If the Home government's beliefs are such that its unique best response is to concede to τ' with a zero probability, then all types $\gamma < \overline{\gamma}$ are strictly worse-off ($\tau' < \tau^{io} \leq T^*(\overline{\gamma}) < T^*(\gamma)$ for all $\gamma < \overline{\gamma}$). This implies that $D_F(\gamma,\tau',u) = D_F^0(\gamma,\tau',u) = \emptyset$ for all $\gamma < \overline{\gamma}$.

Finally, if the Home government's beliefs are such that it is indifferent between conceding and not conceding to τ'. In this case, a best response $\alpha \in MBR(F,\tau',u) = [0,1]$ makes the type-γ Foreign government prefer to demand τ' unilaterally if and only if

$$\alpha W^*(\tau',0,\gamma) + (1-\alpha)\widetilde{W}^*(\gamma) \geq W^*(\tau^{io},0,\gamma)$$

or, equivalently,

$$\alpha \geq \bar{\alpha}(\gamma) \equiv \frac{W^*(\tau^{io}, 0, \gamma) - \widetilde{W}^*(\gamma)}{W^*(\tau', 0, \gamma) - \widetilde{W}^*(\gamma)} \in [0, 1)$$

(with $\bar{\alpha}(\gamma) > 0$ for all $\gamma < \overline{\gamma}$). Therefore, $D_F(\gamma, \tau', u) = (\bar{\alpha}(\gamma), 1]$ and $D_F^0(\gamma, \tau', u) = \{\bar{\alpha}(\gamma)\}$ for all $\gamma \in [\underline{\gamma}, \overline{\gamma}]$. It is readily checked that $\bar{\alpha}$ is a strictly decreasing function (see the proof of Lemma 1). Hence, taking the union over all possible beliefs F, we obtain

$$\left[D(\gamma, \tau', u) \cup D^0(\gamma, \tau', u)\right] = [\bar{\alpha}(\gamma), 1] \subset (\bar{\alpha}(\overline{\gamma}), 1] = D(\overline{\gamma}, \tau', u) \ .$$

This proves completes the proof of Lemma 7. □

Lemma 8. *Consider an equilibrium in which: all types in* $[\underline{\gamma}, \gamma_0]$, *with* $\gamma_0 = (T^*)^{-1}(\tau^{io})$, *make multilateral demands followed by the implementation of tariff* τ^{io}; *and all types in* $(\gamma_0, \overline{\gamma}]$ *make unsuccessful unilateral demands. Reasonable beliefs must assign zero probability to all types* $\gamma \leq \gamma_0$ *following any (off-the-equilibrium-path) unilateral demand* $\tau' \in (T(\gamma_0), \tau^{io})$.

Proof. Observe that, in terms of equilibrium payoffs, this is similar to the case without IO in which all types in $[\underline{\gamma}, \gamma_0]$ successfully demand τ^{io} and all types in $(\gamma_0, \overline{\gamma}]$ fail to obtain a concession. We can therefore replicate the argument of Lemma 1 (replacing $\gamma^{\sup}$ by γ_0) to prove that all types $\gamma \leq \gamma_0$ must be eliminated according to the D1 criterion. □

References

Bagwell, K., P. Mavroidis, and R. Staiger (2007). Auctioning countermeasures in the WTO. *Journal of International Economics 73*, 309–332.

Bagwell, K. and R. W. Staiger (1999). An economic theory of GATT. *American Economic Review 89*, 215–248.

Bagwell, K. and R. W. Staiger (2005). Enforcement, private political pressure, and the general agreement on tariffs and trade/world trade organization escape clause. *Journal of Legal Studies 34*, 471–525.

Beshkar, M. (2010). Trade skirmishes and safeguards: A theory of the WTO dispute settlement process. *Journal of International Economics 82*, 35–48.

Beshkar, M. and E. W. Bond (2012). Cap and escape in trade agreements. mimeo, Vanderbilt University.

Busch, M. L. (2000). Accommodating unilateralism? U.S. section 301 and GATT/WTO dispute settlement. mimeo Queen's University.

Busch, M. L. and E. Reinhardt (2000). Bargaining in the shadow of the law: Early settlement in GATT/WTO disputes. *Fordham International Law Journal 24*, 158–172.

Cho, I. K. and D. Kreps (1987). Signaling games and stable equilibria. *Quarterly Journal of Economics 102*, 179–221.

Conconi, P., G. Facchini, and M. Zanardi (2012). Fast track authority and international trade negotiations. *American Economic Journal: Economic Policy 4*, 146–189.

Fudenberg, D. and J. Tirole (1991). *Game Theory.* Cambridge, MA: Massachusetts Institute of Technology Press.

Horn, H., G. Maggi, and R. W. Staiger (2010). Trade agreements as endogenously incomplete contracts. *American Economic Review 100*, 394–419.

Hungerford, T. L. (1991). GATT: a cooperative equilibrium in a noncooperative trading regime? *Journal of International Economics 31*, 357–369.

Johnson, H. G. (1954). Optimum tariffs and retaliation. *Review of Economic Studies 21*, 142–153.

Klimenko, M., G. Ramey, and J. Watson (2008). Recurrent trade agreements and the value of external enforcement. *Journal of International Economics 74*, 475–499.

Levy, P. (1999). Lobbying and international cooperation in tariff setting. *Journal of International Economics 47*, 345–370.

Limão, N. and K. Saggi (2008). Tariff retaliation versus financial compensation in the enforcement of international trade agreements. *Journal of International Economics 76*, 48–60.

Ludema, R. (2001). Optimal international trade agreements and dispute settlement procedures. *European Journal of Political Economy 72*, 355–376.

Maggi, G. (1999). The role of multilateral institutions in international trade cooperation. *American Economic Review 89*, 190–214.

Maggi, G. and R. Staiger (2011). The role of dispute settlement procedures in international trade agreements. *Quarterly Journal of Economics 126*, 475–515.

Martin, A. and W. Vergote (2008). On the role of retaliation in trade agreements. *Journal of International Economics 76*, 61–77.

Mayer, W. (1981). Theoretical considerations on negotiated tariff adjustments. *Oxford Economic Papers 33*, 135–153.

Park, J. H. (2011). Enforcing international trade agreements with imperfect private monitoring. *Review of Economic Studies 78*, 1102–1134.

Pelc, K. J. (2010). Constraining coercion? Legitimacy and its role in U.S. trade policy 1975-2000. *International Organization 64*, 65–96.

Puckett, A. L. and W. L. Reynolds (1996). Rules, sanctions and enforcement under section 301: At odds with the WTO? *The American Journal of International Law 90*, 675–689.

Ramey, G. (1996). D1 signaling equilibria with multiple signals and a continuum of types. *Journal of Economic Theory 69*, 508–531.

Reinhardt, E. (2000). Aggressive multilateralism: The determinants of GATT/WTO dispute initiation. mimeo Emory University.

Riezman, R. (1991). Dynamic tariffs with asymmetric information. *Journal of International Economics 30*, 267–283.

Rosendorff, P. (2005). Stability and rigidity: Politics and design of the WTO's dispute settlement procedure. *American Political Science Review 99*, 389–400.

Rossmiller, G. E. (1994). Discussion. In G. Anania, C. A. Carter, and A. F. McCalla (Eds.), *Agricultural Trade Conflicts and GATT: New Dimensions in U.S.-European Agricultural Trade Relations*, pp. 262–263. Boulder, Colorado: Westview Press.

Schoppa, L. J. (1999). The social context in coercive international bargaining. *International Organization 53*, 307–342.