The GATT/WTO as an Incomplete Contract

Henrik Horn (IIES, Stockholm University)
Giovanni Maggi (Princeton University and NBER)
Robert W. Staiger (University of Wisconsin and NBER)

September 2005
(preliminary draft)

1. Introduction

The World Trade Organization (WTO) regulation of trade in goods – the General Agreement on Tariffs and Trade (GATT) – is a highly incomplete contract.[1] It directly binds only trade policies, leaving significant discretion over domestic policy instruments with trade impact to national governments. The policies that are bound, are bound rigidly, and are thus not adaptable to stochastic economic or political shocks. And most of the provisions are vaguely worded, and leave for adjudicating bodies to determine the actual ambit of the agreement.

A sizeable economic literature seeks to shed light on various aspects of this incompleteness. The typical approach is to impose exogenous restrictions on the set of policy instruments that can be included in a trade agreement, and examine what the agreement can accomplish given these limitations.[2]

This literature has helped shed light on several important aspects of the agreement. But it has in our view an important limitation in that the incompleteness of the agreement is assumed rather than endogenously derived. It is easy to accept as a general statement that a trade agreement has to be incomplete because of the immense costs that would be involved in reaching a fully efficient agreement (if this is a practical possibility at all). However, there are many different ways in which an agreement can save on contracting costs, so there are many different forms that the incompleteness could take. The general purpose of this paper is to take the analysis of the GATT as an incomplete contract one step further, by endogenously determining the choice of contract form. In contrast to the existing literature, we will thus not take the structure of the agreement for given, and analyze possible consequences thereof, but characterize circumstances under which various forms of trade agreements may arise.

[1] For simplicity, we will use the term "GATT" in a somewhat imprecise sense. More formally, the WTO regulates trade in goods through "GATT 1994", consisting of the original 1947 GATT agreement plus a number of agreed modifications and interpretations, plus a number of special agreements, such as those governing safeguards, anti-dumping, technical barriers to trade and health measures.

[2] An incomplete list of papers that fall into this category is Copeland (1990), Bagwell and Staiger (2001), Battigalli and Maggi (2003), and Horn (2003).

The GATT consists of a large number of provisions. But we believe that one can identify certain structural features of the agreement that seem important from an incomplete contracting point of view:

1. The agreement binds the level of trade instruments, such as import quotas and tariffs, requiring that the bindings fulfil the Most-Favored Nation (MFN) provision.

2. Domestic instruments are left at the discretion of national governments. But internal policies have to respect the NT clause (and similar provisions in other agreements regulating goods trade), and there is a regulation of subsidies.

3. The bindings are rigid. But there are "escape mechanisms" allowing countries e.g. to unilaterally impose temporary protection (such as Art. XIX GATT safeguards), or to renegotiate bindings; anti-dumping has also increasingly come to be used this way.

4. Tariff bindings are typically "weak" in that they stipulate a maximum permitted level for a policy variables, but allow for lower levels.

5. There are no bindings of trade volumes or prices, but only policy measures (except perhaps for safeguards). The legal instrument that in terms of interpretation is closest to such a binding – the "Non-Violation" instrument (NV) provided for through GATT Article XXIII.1(b) – has especially since the advent of the WTO played a very modest role, at least in dispute settlement processes.[3]

The ultimate aim of this line of research is to understand why the regulation of goods trade has taken this particular form. We believe that the analysis to follow sheds light on at least some of these core features.

The analytical starting point of the paper is the notion that legislators face two fundamental problems when forming a trade agreement. The first is that there is very significant *uncertainty* concerning the circumstances that will prevail during the life-time of the agreement. This uncertainty suggests that the agreement should be highly adaptable to the contingencies that unfold. However, the second salient feature of the contracting situation is that there are *contracting costs*. There are different types of costs associated with forming a trade agreement. There are costs in terms of effort and time for working out bargaining proposals and for evaluating proposals made by others, and there are costs involved in verifying that trading partners abide by the rules of the agreement. While contracting costs can take many different forms, it is probably safe to say that contracting costs tend to be higher when the agreement is more detailed, in terms of the contingencies that it specifies and the number of policies that it seeks to constrain. This basic idea will be reflected in our formalization of contracting costs.

It is hard to dispute that contracting costs constitute a severe constraining factor for the design of a trade agreement. The choice of the structure of the agreement will have to reflect a

[3] The exact meaning of this provision is hard to determine from the text itself. But it is often seen as protecting market access expectations of governments against changes in policies by their trading partners – even when these policies are not contracted over – which would have the effect of upsetting the market access that a government could have reasonably expected based on a prior GATT/WTO negotiation.

trade off between the performance of the contract as such and the associated contracting costs. The more specific purpose of this paper is hence to highlight the role that the combination of uncertainty and contracting costs may play for explaining core features of the GATT.

We will work with a linear, partial equilibrium, two-country setting, in order to be able to easily compare welfare levels achieved from a number of different forms of agreements. In the industry under study, the Home country is a natural importer. But in the background there are many such industries, with Home importing in some and exporting to Foreign in some. In this otherwise non-distorted economy, Home may experience two types of externalities: a consumption externality, and/or a production externality. Home is for simplicity the only policy-active country in this industry, and has access to an import tariff, separate internal taxes on Home and Foreign produced products, as well as a production subsidy. Uncertainty will play a central role in the analysis, and we will employ various formulations of the uncertainty. Generally speaking, all the parameters of the model may be uncertain, that is, both the parameters describing demand and supply conditions, as well as those capturing the externalities.

In the absence of an agreement, Home would manipulate the terms of trade in standard fashion. This is of course a suboptimal situation from a global point of view, and there is therefore scope for an agreement that improves world welfare. Were it not for the externalities, the optimal trade agreement (at least in the absence of contracting costs) would be very simple: it would just stipulate free trade and no other policy intervention. But due to the externalities, the contracting problem is substantially more complex, for several reasons. As will be shown, the first best agreement will now require instruments to be state contingent, offsetting the externalities in a Pigouvian fashion, and this first best outcome will require a substantially more complex agreement, to be implemented.

In order to formalize the notion of contracting costs we employ a very simple representation, but one that we believe captures certain important aspects of contracting costs. Inspired by Battigalli and Maggi (2002), we will assume that these costs depend both on the degree to which a contract is state contingent, and on the scope of its coverage of policy instruments. As a result of these costs, the parties may find it worthwhile to use a simpler contract form than the one required to implement first best policies. As pointed out by Battigalli and Maggi (2002), there are two essential ways in which the parties to a contract can save on contracting costs. One is that the agreement is (partially or fully) *rigid* – i.e. it is insensitive to changes in the underlying economy. The other is that it leaves *discretion* in the governments' choices of policies. In our framework, one would naturally expect that the costs of making the contract more contingent pushes toward rigidity, and the costs of contracting over policies pushes toward discretion, and this is indeed the case, but we will show that the combination of contracting costs and uncertainty may have also much more subtle implications.

The paper proceeds as follows. Section 2 describes our model of the economy and our formalization of contracting costs. The section also presents two benchmark scenario. One is the no-agreement outcome – that is the Nash equilibrium – and the other is the first-best outcome. The approach of the paper is to view an optimal agreement as one that maximizes global welfare minus contracting costs. The no-agreement and first-best outcomes can thus be seen as the outcomes of two extreme forms of contracting costs, the first-best outcome resulting when contracting costs are zero, and the no-agreement outcome resulting when they are sufficiently high to dominate any gains that could be had from an agreement.

Section 3 characterizes the optimal trade agreement within a simple class of contracts, and how it depends on contracting costs, on the degree and type of uncertainty, and on the expected demand and supply conditions. When uncertainty is small – in a sense to be made precise – we are able to characterize the optimal contract with analytical tools. When uncertainty is large, on the other hand, the optimization problem is too complex for analytical solution, and for this reason we consider a simple parametrization of the model and we use numerical techniques to study the optimal contract.

Here we offer a quick preview of our main results, referring the reader to the next sections for an intuitive explanation of these results.

At a broad level, we find that, as contracting costs increase, the optimal agreement is initially fully state-contingent, then it becomes increasingly rigid, and then starts allowing for discretion, eventually reaching the situation where it is optimal to have no contract at all.

While the above insight applies to the contract as a whole, our model also offers interesting predictions on the way that contractual incompleteness varies across policy instruments. We find that the optimal contract tends to leave more discretion on domestic policy instruments than on border measures. More specifically, while for a range of contracting costs it is optimal to bind import taxes while leaving domestic instruments discretionary, it is never optimal to leave import taxes to discretion and contract only over domestic instruments.

The role of uncertainty depends in subtle ways on its source. Broadly speaking, when uncertainty concerns mostly the extent of externalities, the optimal contract tends to feature state-contingency and/or discretion. On the other hand, when the uncertainty concerns mostly demand and supply conditions, the preferred contract tends to feature greater rigidity.

Leaving discretion over domestic instruments is more likely to be optimal, *ceteris paribus*, (i) when demand is more elastic; (ii) when supply is more elastic; and (iii) when the expected level of demand is higher. We also find that the impact of the expected demand level on the optimal degree of discretion tends to be stronger when there is more uncertainty.

The remaining part of the paper extends the analysis to shed light on several other core aspects of the GATT that we believe are best understood from an incomplete contracts perspective.

Section 4 investigates the role of the NT clause, by extending the set of possible agreements to include a simple formalization of NT as it applies to taxation. We find that the NT rule is likely to save on contracting costs, since it does not require binding domestic consumption taxes. It has the virtue of allowing for some ex post flexibility, since the importing country can set the general consumption tax level in response to stochastic disturbances. But it is not a perfect substitute to bindings, since it allows the common tax level to be set opportunistically. As a result, the optimal contract tends to be NT-based when contracting costs are high. The optimal contract also tends to be NT-based when the flexibility it implies is valuable. Consequently, an NT-based contract is more likely to be optimal when there is large uncertainty concerning the consumption externality.

Section 5 examines the usefulness of an NV provision as a means to economize on contracting costs. The NV clause allows countries to avoid contracting directly over domestic policy instruments, thereby saving on the costs of specifying and verifying the values of these instruments. On the other hand, the clause requires verification of the state of demand and supply, which is costly. Hence, an NV-based contract tends to be optimal when the cost of verifying policies is large relative to the cost of state verification.

In section 6 we argue that the presence of contracting costs may explain why GATT stipulates weak bindings – e.g. maximum tariff levels – rather than strict bindings. More specifically, we show that the optimal contract may include *rigid* weak bindings. This type of binding combines rigidity and discretion, since the ceiling does not depend on the state of the world, and the government has discretion to set the policy below the ceiling. Thus we find that rigidity and discretion may be complementary ways to economize on contracting costs.

The general conclusion that we see emerging from this analysis is that salient features of the GATT, such as the fact that the agreement binds trade policies but not internal instruments, that it focuses mostly on binding policies rather than economic outcomes, the presence of the NT provision, and perhaps the more limited role of NVs, can be understood as resulting from the interaction between contracting costs, uncertainty and the features of the underlying economy.

2. The Model

We adopt a partial equilibrium perspective, according to which there are potentially many goods produced, consumed and traded between a home country and a foreign country. For simplicity, we concentrate on a single good, for which the home country is the natural importer. We then characterize the impact of internationally negotiated contracts on the production, consumption and trade of this good.

We are interested in exploring contracting possibilities over a rich set of instruments. To this end, we assume that the home government can use an import tariff (τ), an internal tax on consumption of the domestically produced good (t_h), an internal tax on consumption of the imported product (t_f), and a production subsidy to domestic firms (s). All instruments are expressed in specific terms. For simplicity, we assume that the foreign (exporting) government has no policies available in the sector under consideration, though our results generalize naturally to a setting in which the foreign government also makes policy choices.

The goods markets in the two countries are integrated, and prices differ only to the extent of government intervention. Throughout we focus on non-prohibitive levels of government intervention that do not choke off all trade. Let p and p^*denote the prices paid by consumers in the home and foreign country, respectively, with asterisks denoting variables in the foreign country here and throughout. Due to the possibility of consumer arbitrage, and to the absence of taxation in the foreign country, we have the following relationship between home and foreign consumer prices:

$$p = p^* + \tau + t_f.$$

For a foreign firm to sell in both countries, it must receive the same price for sales in the foreign-country market as it receives after taxes for sales in the home-country market:

$$q^* = p - \tau - t_f,$$

where q^* is the price received by a foreign firm for sales in the foreign-country market. Due to the absence of taxation or other trade costs in the foreign country, we also have that producer and consumer prices in the foreign country are equalized, or

$$q^* = p^*.$$

Finally, let q denote the home-country producer price, i.e., the price received by a home firm for sales in the home-country market. The relationship between the home-country producer price and the home-country consumer price is given by

$$q = p - t_h + s.$$

We can express the above pricing relationships in more compact form as

$$\begin{aligned} p &= p^* + T, \text{ and} \\ q &= p^* + T + S, \end{aligned}$$

where $T \equiv \tau + t_f$ and $S \equiv s - t_h$. These are the two key price wedges that will be used in the analysis to follow; the first one is the wedge between the home-country consumer price and the foreign-country price (equal to T), and the second one is the wedge between the home-country producer price and the foreign-country price (equal to $T + S$). Note that τ and t_f are perfectly substitutable policy instruments, and the same is true for s and t_h. Thus, while it is appropriate to refer to τ as a "border measure" and to t_f, t_h and s as "internal measures," we will also sometimes refer to T as the total tax on imports, or simply as the "import tax," and to S as the "effective production subsidy."

We turn next to specifying the demand, supply and market-clearing conditions. Home and foreign demand functions take a simple linear form:

$$\begin{aligned} D(p) &= \alpha - \beta p, \text{ and} \\ D^*(p^*) &= \alpha^* - \beta^* p^*, \end{aligned}$$

where $\alpha > 0$, $\beta > 0$, $\alpha^* > 0$, and $\beta^* > 0$. The home and foreign supply functions are also linear:

$$\begin{aligned} X(q) &= \lambda q, \text{ and} \\ X^*(q^*) &= \lambda^* q^*, \end{aligned}$$

where $\lambda > 0$ and $\lambda^* > 0$. Market clearing requires that world demand equal world supply, or

$$\alpha - \beta p + \alpha^* - \beta^* p^* = \lambda q + \lambda^* q^*.$$

The above system, together with the arbitrage relationships, yields expressions for the three market clearing prices as functions of T and S:

$$p(T, S) = [\alpha + \alpha^* + (\beta^* + \lambda^*)T - \lambda S]/A,$$

$$q(T, S) = [\alpha + \alpha^* + (\beta^* + \lambda^*)T + (\beta + \beta^* + \lambda^*)S]/A, \text{ and}$$
$$p^*(T, S) = q^*(T, S) = [\alpha + \alpha^* - (\beta + \lambda)T - \lambda S]/A,$$

where $A \equiv \lambda + \lambda^* + \beta + \beta^*$. At the market clearing prices, home import volume, M, is equal to foreign export volume, E^*, and is given by

$$M(T,S) = E^*(T,S) = [\alpha(\beta^* + \lambda^*) - \alpha^*(\beta + \lambda) - (\beta + \lambda)(\beta^* + \lambda^*)T - \lambda(\beta^* + \lambda^*)S]/A.$$

Note that $M(T = 0, S = 0) > 0$, and hence the home country is a natural importer of the good under consideration, provided that

$$\frac{\alpha}{\lambda + \beta} > \frac{\alpha^*}{\lambda^* + \beta^*}.$$

We will henceforth assume that this condition is met. For future use we may also define implicitly the locus of policies that prohibit trade according to $M(T^a(S), S) \equiv 0$. Explicit calculations yield

$$T^a(S) = [\frac{\alpha}{\lambda + \beta} - \frac{\alpha^*}{\lambda^* + \beta^*}] - \frac{\lambda}{\lambda + \beta} \cdot S.$$

We assume that each government's objective corresponds to the welfare of its representative citizen. For the foreign-country government, who we recall has no policy instruments of its own in the sector under consideration, this objective is simply the sum of foreign consumer and producer surplus, which we denote by CS^* and PS^*, respectively. Hence, the objective of the foreign-country government, $W^*(T,S)$, is given by

$$W^*(T,S) = CS^*(T,S) + PS^*(T,S),$$

where

$$CS^*(T,S) \equiv \int_{p^*(T,S)}^{\alpha^*/\beta^*} D^*(p^*)dp^*; \text{ and } PS^*(T,S) \equiv \int_0^{q^*(T,S)} X^*(q^*)dq^*.$$

In the home country, in addition to home consumer and producer surplus (CS and PS, respectively), a further surplus consideration is the net revenue generated by the home-government's policy intervention. The home government's net revenue is composed of its revenue from the import tax ($T \cdot M$) minus its expenditure on the effective production subsidy ($S \cdot X$). Moreover, we allow the possibility that there may exist several kinds of externalities in the home country that introduce a divergence between national income and national welfare. Specifically, we assume that there is a positive production externality equal to σX with $\sigma > 0$, and a negative consumption externality equal to $-\gamma D$ with $\gamma > 0$ (nothing substantial in the analysis would change if the signs of the externalities were different). These externalities enter directly and separably into the representative home-country citizen's utility and do not cross borders. Hence, the home-country government's objective, $W(T,S)$, is given by the sum of consumer surplus, producer surplus, tax revenue and the valuation of the externalities associated with home-country production and consumption, or

$$W(T,S) = CS(T,S) + PS(T,S) + T \cdot M(T,S) - S \cdot X(T,S) + \sigma X(T,S) - \gamma D(T,S),$$

where

$$\begin{aligned} D(T,S) &\equiv D(p(T,S));\; X(T,S) \equiv X(q(T,S)); \\ CS(T,S) &\equiv \int_{p(T,S)}^{\alpha/\beta} D(p)dp; \text{ and } PS(T,S) \equiv \int_0^{q(T,S)} X(q)dq. \end{aligned}$$

2.1. The Nash equilibrium and efficient policies

We first derive the Nash equilibrium policies, which we take to represent the policy choices made in the absence of any agreement between the home and foreign governments. With the foreign government passive, the Nash equilibrium policies are defined by the two first-order conditions characterizing the home government's best-response policy choices. These first-order conditions simplify to

$$\begin{aligned} \frac{dW(T,S)}{dT} &= 0 \Longrightarrow \frac{E^*(S,T)}{\beta^*+\lambda^*} - T + \frac{\lambda}{\beta+\lambda}(\sigma - S) + \frac{\beta}{\beta+\lambda}\gamma = 0, \text{ and} \\ \frac{dW(T,S)}{dS} &= 0 \Longrightarrow \frac{E^*(S,T)}{\beta^*+\lambda^*} - T + \frac{\beta+\beta^*+\lambda^*}{\beta^*+\lambda^*}(\sigma - S) - \frac{\beta}{\beta^*+\lambda^*}\gamma = 0. \end{aligned}$$

These two first-order conditions define, respectively, the best-response level of T given S, which we denote $T^R(S)$, and the best-response level of S given T, which we label $S^R(T)$. These two best-response functions will play an important role in what follows.

Solving the system, we may derive the following expressions for the Nash equilibrium import tax and effective production subsidy choices of the home government, which we denote by T^{NE} and S^{NE}, respectively:[4]

$$\begin{aligned} T^{NE} &= \gamma + \frac{E^*(S^{NE},T^{NE})}{\beta^*+\lambda^*} = \gamma + \frac{p^*}{\eta^*}, \text{ and} \\ S^{NE} &= \sigma - \gamma, \end{aligned}$$

where η^* is the elasticity of the foreign export supply (itself evaluated at S^{NE} and T^{NE}).

Recalling that $T \equiv \tau + t_f$ and $S \equiv s - t_h$, we note that there are many equivalent policy combinations that correspond to the Nash policy choices T^{NE} and S^{NE}. One of these combinations is $\{\tau = \frac{p^*}{\eta^*}, t_h = t_f = \gamma, s = \sigma\}$. This particular policy combination makes it transparent that in the Nash equilibrium the home-country government sets its traditional (Johnson, 1953-54) "optimal tariff" – the inverse of the Nash equilibrium foreign export supply elasticity – to exploit its power over the terms of trade (p^*), applies a Pigouvian production subsidy at the level of the production externality, and applies a uniform Pigouvian consumption tax at the level of the consumption externality.

Finally, solving for the explicit expression for T^{NE} in terms of the underlying parameters yields

$$T^{NE} = \frac{(\beta^*+\lambda^*)[\alpha + A\gamma - \lambda(\sigma-\gamma)] - (\beta+\lambda)\alpha^*}{(\beta^*+\lambda^*)[A+(\beta+\lambda)]}.$$

[4] It is not hard to verify that W is jointly concave in (T,S), which ensures that the first-order conditions are sufficient.

We focus throughout on parameter combinations for which trade is not prohibited in the Nash equilibrium. These parameter combinations are defined by the restriction that $T^{NE} < T^a(\sigma-\gamma)$. In light of our assumption that the home country is a natural importer of the good under consideration, it is direct to verify that this restriction is met in the absence of externalities (i.e., when $\sigma = 0$ and $\gamma = 0$), and that the restriction in effect places upper limits on the magnitude of the externality parameters σ and γ.

Having characterized the Nash equilibrium policy choices, we turn next to the globally efficient policies. The globally efficient policies are those policies that maximize "global welfare," that is, the sum of home and foreign welfare:[5]

$$\Omega(T, S) \equiv W(T, S) + W^*(T, S).$$

It is direct to verify that the efficient import tax and effective production subsidy choices of the home government, which we denote by T^{eff} and S^{eff}, respectively, are given by

$$\begin{aligned} T^{eff} &= \gamma, \text{ and} \\ S^{eff} &= \sigma - \gamma. \end{aligned}$$

Hence, efficient policy combinations ensure that the relevant price wedges only reflect externalities, not terms-of-trade considerations. In particular, the wedge between the domestic consumer price and the foreign price (T) should be equal to the consumption externality γ (Pigouvian consumption tax), and the wedge between the domestic producer price and the foreign price ($S + T$) should be equal to the production externality σ (Pigouvian production subsidy).

At this point it is convenient to emphasize a feature of the Nash policy choices and their relation to the efficient policy choices that will turn out to be important for interpreting our results in the following sections. In particular, notice that the Nash choice of effective production subsidy is efficient ($S^{NE} = S^{eff}$), and the nature of the policy inefficiency associated with the Nash equilibrium is then entirely reflected in a Nash level of import taxes that is too high – and Nash trade volumes that are therefore too low – relative to their efficient levels ($T^{NE} > T^{eff}$). The inefficiently high level of T reflects in turn the unilateral incentive to manipulate the terms of trade with the choice of import taxes. Therefore, it is accurate to say that the potential gains from contracting in this setting arise entirely from the ability to control the incentive to utilize import taxes to manipulate the terms of trade. As a consequence of this feature, while the contracts we consider below may impose constraints beyond the choice of import taxes, we will nevertheless at times refer to these contracts as "trade agreements," because they represent attempts to solve what is evidently at its core a trade – and trade policy – problem.

The feature we emphasize above is quite general (see Bagwell and Staiger, 2001), and for the costly contracting environment we describe in the next section it gives rise to an important implication concerning the potential value of contracting over S while leaving T unrestricted. In particular, as we next establish, a contract that restricts the level of S while leaving T to the government's discretion can achieve *no* improvement over the Nash equilibrium.

[5] By defining globally efficient policies in this way, we are implicitly assuming that the two governments can transfer surplus between them in a lump sum fashion.

To see that this is the case, observe first that a contract that constrains only S cannot achieve greater surplus than the maximal surplus from contracting over S alone when contracting costs are zero: when contracting costs are zero, this defines the level of the feasible contracting surplus from contracting over S alone, and the introduction of strictly positive contracting costs can only reduce the feasible contracting surplus from this level. Recalling that $T^R(S)$ is the best-response level of T given S, this maximal surplus is given by the value of $\Omega(T^R(S), S)$ evaluated at the optimal level of S, with the optimal level of S defined in turn by the associated first-order condition $d\Omega/dS = 0 \Longrightarrow T^{R'}(S) = -\ \Omega_S/\Omega_T$. This first-order condition requires that the slope of $T^R(S)$ be equated with the slope of an iso-Ω curve in (T, S) space. It is direct to verify that the slope of $T^R(S)$ is

$$T^{R'}(S) = -\frac{\lambda}{\beta + \lambda}.$$

The slope of an iso-Ω curve in general is given by $-\frac{\Omega_S}{\Omega_T} = \frac{W_S + W_S^*}{W_T + W_T^*}$. However, since at the Nash equilibrium $W_S = W_T = 0$, the slope of the iso-Ω curve at the Nash equilibrium point is

$$-\frac{\Omega_S}{\Omega_T} = -\frac{W_S^*}{W_T^*} = -\frac{\frac{dW^*}{dp^*} \cdot \frac{dp^*}{dS}}{\frac{dW^*}{dp^*} \cdot \frac{dp^*}{dT}} = -\frac{\frac{dp^*}{dS}}{\frac{dp^*}{dT}} = -\frac{\lambda}{\beta + \lambda} \tag{2.1}$$

Therefore, the slope of the iso-Ω curve at the Nash point is equal to the slope of the $T^R(S)$ curve, and as a consequence, the level of S that maximizes $\Omega(T^R(S), S)$ is the Nash equilibrium level S^{NE}. We may conclude, then, that a contract that constrains only S cannot achieve greater surplus than $\Omega(T^R(S^{NE}), S^{NE})$, which is just the Nash equilibrium surplus $\Omega(T^{NE}, S^{NE})$. The next lemma records this result.[6]

Lemma 1. *Constraining the effective subsidy S while leaving the import tax T to discretion cannot improve over the Nash equilibrium.*

At a broad level, the intuition for this result is very simple. Contracting over S alone is useless because, as we have emphasized above, the inefficiency that arises in the noncooperative equilibrium concerns T, not S. We can also be a little more precise about the logic behind Lemma 1. The key steps of the argument are two. First, as equation 2.1 makes transparent, the slope of the iso-Ω curve at the Nash point is equal to the slope of the iso-p^* line. This is an immediate consequence of the fact that Home policies affect Foreign welfare only through terms of trade, and that at the Nash point small changes in home policies have no first-order effect on Home welfare. Second, the best-response import tax line $T^R(S)$ coincides with the iso-p^* line through the Nash point. The reason is that the Nash tariff implements the terms of

[6] Notice that this result is distinct from and not contradictory to the result emphasized by Copeland (1990), that negotiating over tariffs can always generate surplus even if other instruments are non-negotiable. Copeland's result implies that the inclusion of tariffs in the set of instruments over which negotiations occur is *sufficient* for the possibility of gains from negotiations. The result we report in Lemma 1 implies that the inclusion of tariffs in the set of instruments over which negotiations occur is also *necessary* for the possibility of gains from negotiations.

trade that is optimal from the Home point of view, and constraining S away from its reaction curve triggers a change in T that brings the terms of trade back to its optimal level.

We emphasize that, in a world of costless contracting, the result we have highlighted in Lemma 1 would be irrelevant, because if contracts are costless they would always be written in a way that placed constraints on all of the policy instruments that enter into the determination of the terms of trade. But in a world of costly contracting, one has to consider contracts that place constraints on only a subset of these instruments, and this result then gains relevance. In particular, in such a world Lemma 1 implies that it can never be optimal to constrain internal measures but not import taxes. We will confirm the importance of this insight when we analyze optimal contracts in the presence of contracting costs.

2.2. The costs of contracting

We think of the vector of exogenous parameters $\boldsymbol{\theta} = (\alpha, \alpha^*, \beta, \beta^*, \lambda, \lambda^*, \sigma, \gamma)$ as a random vector whose value is not known at the time of writing the contract. We will sometimes refer to these as the state-of-the-world variables, or simply the "state" variables. We let $\boldsymbol{\Theta}$ denote the support of $\boldsymbol{\theta}$.

We formalize contracting costs in a very stylized way. There are two kinds of elementary contracting costs: the costs of including *state* variables in the contract $(\alpha, \alpha^*, \beta, \beta^*, \lambda, \lambda^*, \sigma, \gamma)$, and the costs of including *policy* variables in the contract (τ, t_f, s, t_h). We think of the cost of including a given variable in the contract as capturing both the cost of describing this variable (i.e. defining the variable, how it should be measured etc., along the lines of the "writing costs" emphasized by Battigalli and Maggi, 2002) as well as the cost of verifying its value ex-post. We assume that, if a variable is included in the contract, the court automatically verifies its value ex-post, incurring the associated verification cost.[7] A broader interpretation of these contracting costs might also include negotiation costs: it is reasonable to think that negotiation costs are higher when there are more policy instruments on the table, and when there are more relevant contingencies to be discussed for a given policy instrument.

The cost of contracting over a state variable is c_s and the cost of contracting over a policy variable is c_p. We assume that, if a variable is included in the contract, the associated cost is incurred only once, regardless of how many times that variable is mentioned in the contract; in other words, there is no cost in "recalling" a given variable after the first time it appears in the contract.

Summarizing, the cost of writing a contract is given by

$$C = c_s \cdot n_s + c_p \cdot n_p,$$

where n_s and n_p denote, respectively, the number of state variables and policy variables included in the contract.

A few examples can be useful to illustrate our assumptions on contracting costs:

[7] Of course this is a strong assumption. In reality, the WTO verifies compliance with the contract only if there is a complaint by one of the contracting parties. We expect that similar qualitative insights would emerge in a richer model with verification "on demand" to the extent that verification occurs in equilibrium at least with some probability.

Example 1: The contract $\{\tau = 2\}$ specifies a rigid binding on the level of the home tariff, and costs c_p.

Example 2: The contract $\{t_h = t_f = 3\}$ specifies a rigid commitment for the level of the home consumption tax on home-produced and on foreign-produced goods, and costs $2c_p$.

Example 3: The contract $\{s = \sigma\}$ specifies a state-contingent commitment for the level of the home production subsidy, and costs $c_p + c_s$.

Example 4: The contract $\{\tau = 0; s = \sigma; t_h = t_f = \gamma\}$ specifies a rigid binding on the level of the home tariff, a state-contingent commitment for the level of the home production subsidy, and a rigid commitment for the level of the home consumption tax on home-produced and on foreign-produced goods, and costs $4c_p + 2c_s$.

It might be reasonable to assume that it is more costly to contract over internal measures (t_f, s, t_h) than over tariffs (τ), because in reality it is easier to verify border measures than internal measures. But as will become clear below, in this case our qualitative results would only be strengthened. So in the interests of parsimony, we do not introduce this distinction, and rather maintain the assumption of a common contracting cost for border and internal measures.

For future reference, we introduce two definitions. First, we say that two contracts are *equivalent* if they implement the same outcome and have the same cost. Introducing a notion of contract equivalence in this context is necessary because, as will become clear shortly, for any given contract there exist many other contracts that implement the same outcome and have the same cost.

Second, we refer to the *efficiently-written first-best* contract as the least costly among the contracts that implement the first best outcome. We will often label this the *EWFB* contract.

3. Optimal Contracts

We begin our characterization of optimal contracts by describing the efficiently-written first-best (EWFB) contract. In the first part of the paper we focus on *instrument-based* contracts, i.e. contracts that impose (possibly contingent) constraints on policy instruments. We defer until a later section the discussion of *outcome-based* contracts, i.e. contracts that impose constraints on equilibrium outcomes such as prices or trade volumes.

Observe first that since the two policy instruments τ and t_f are perfect substitutes and matter only through their sum T, when we view them as contractual variables they are perfect *complements*: constraining one of the two instruments but not the other would have no effect. The same is true for the domestic instruments s and t_h, which matter only through their sum S. Hence, as a starting point we can think of T and S as the relevant policy variables, and the associated contracting costs are $2c_p$ for each of these variables. Moreover, for now we restrict our search to contracts that impose separate equality constraints on T and S. To be concrete, we allow for clauses of the type $(T = \gamma)$ or $(S = 10)$, but not for clauses of the type $(T + S = \sigma)$.[8]

[8] We note that, given our assumptions on the costs of contracting, clauses of the kind $S + T = \sigma$ cannot be strictly optimal, because one can implement the same oucome as this contract with another contract that separately pins down the values of T and S, and this would not cost more than the original contract. In a later section we will consider contracts that include a national-treatment clause $(t_h = t_f)$ and will argue that this kind of clause may be desirable under a slightly different set of assumptions on contracting costs. But under

We label this class of contracts $\mathcal{K}_0$. In the following sections we will consider broader classes of contracts.

We assume that there is *some* uncertainty in σ and γ, in the sense that $E(|\sigma - \bar{\sigma}|) > 0$ and $E(|\gamma - \bar{\gamma}|) > 0$ where $\bar{\sigma}$ is the expected value of σ and $\bar{\gamma}$ is the expected value of γ; as will become apparent, this ensures that there is a positive benefit from writing a contingent contract. It is clear that the contract $\{T = \gamma; S = \sigma - \gamma\}$ implements the first best outcome. This contract costs $4c_p + 2c_s$. It is also easy to verify that one cannot implement the first best outcome with a contract in $\mathcal{K}_0$ that costs less than $4c_p + 2c_s$. The following proposition states the result:

Proposition 1. *Any efficiently-written first-best contract in* $\mathcal{K}_0$ *is equivalent to* $\{T = \gamma; S = \sigma - \gamma\}$.

Next we look for the optimal contract in $\mathcal{K}_0$, that is, the contract that maximizes expected global welfare net of contracting costs. Of course the qualitative structure of the optimal contract depends on the values of the exogenous parameters. A natural and convenient way to characterize the optimal contract is to track how the optimal contract changes as the level of elementary contracting costs increase. We will consider a proportional increase in the two elementary contracting costs (c_p, c_s). To express our results in a simple comparative-statics fashion, we parameterize contracting costs as follows: $c_p = c$, $c_s = k \cdot c$. The parameter k captures the cost of contracting over a state variable relative to that of contracting over a policy variable, while c captures the general level of elementary contracting costs, which we will henceforth refer to simply as "contracting costs." We will keep k fixed and consider changes in c. Our qualitative results would be the same if we allowed c_p and c_s to vary in a non-proportional way, as long as they co-vary.

The EWFB contract yields expected global welfare net of contracting costs (henceforth simply "expected net global welfare") equal to $\Omega(T = \gamma, S = \sigma - \gamma) - (4 + 2k) \cdot c$.[9] When contracting costs are zero, the EWFB contract is of course optimal. But for sufficiently high contracting costs the EWFB contract cannot be optimal, because as c rises eventually even the empty contract (which costs nothing and yields expected global welfare of $\Omega(T = T^{NE}, S = \sigma - \gamma))$ will yield higher expected net global welfare. The basic question is then whether other (incomplete) contracts will become optimal as contracting costs rise from zero, reflecting an attractive trade-off between the surrender of expected gross-of-contracting-costs global welfare (henceforth, simply "expected gross global welfare") for a reduction in the overall costs of the contract, and if so how the optimal contract is shaped by this trade-off.

In principle, with eight state variables and four policy variables, the optimal restructuring of a contract to economize on overall contracting costs could be enormously complex. This complexity is exacerbated by the inherent non-separability of the contracting problem that we study, a feature that seems an unavoidable part of the problem faced by trade negotiators but which precludes characterizing the optimal contract "task by task" as in Battigalli and Maggi

the current set of assumptions, there can be no strict gains from this kind of contract relative to contracts that are separable in T and S.

[9]Here and throughout the rest of the paper, we use $\Omega(\cdot)$ to denote the expected global welfare gross of contracting costs, suppressing the expectations operator for notational simplicity.

(2002). But knowledge of the EWFB contract significantly reduces this complexity. In particular, notice that the EWFB contract contains only two of the possible eight state variables, σ and γ. This property reflects two important points: first, the government objectives we adopt assume that the distributional impacts of local price movements are unimportant to governments, implying that globally optimal policy intervention exists only to correct externalities (σ and γ);[10] and second, we have chosen to focus here on contracts that impose constraints on policies (which optimally depend only on σ and γ) as opposed to economic magnitudes such as prices and quantities directly (which optimally depend on all eight state variables).[11] Together with the perfect policy substitutability between τ and t_f and between s and t_h, the implication is that opportunities for economizing on overall contracting costs starting from the EWFB contract are relatively easy to describe. As we demonstrate below, this in turn permits some general principles to emerge in a transparent fashion.

Specifically, between the EWFB contract which costs $(4+2k)\cdot c$ and the empty contract which costs nothing, there are five cost classes of contracts that warrant consideration: contracts costing $(2+2k)\cdot c$; contracts costing $(2+k)\cdot c$; contracts costing $(4+k)\cdot c$; contracts costing $4\cdot c$; and contracts costing $2\cdot c$. Of course, within each cost class there are many possible contracts, but we may now use Lemma 1 to reduce the complexity of the problem further. Recalling from Lemma 1 that it can never be optimal to constrain S but not T, we then have five kinds of contracts to consider, corresponding to the five cost classes listed just above: contracts that constrain T as a function of two state variables; contracts that constrain T as a function of one state variable; contracts that constrain T and S as functions of one state variable; contracts that constrain T and S in a non-state-contingent fashion; and contracts that constrain T in a non-state-contingent fashion.

We are able to derive strong analytical results regarding the form of the optimal contract for the case in which the degree of uncertainty is relatively small. For this reason we will first focus on the case of small uncertainty, and then we will turn to the case of large uncertainty, where we will utilize both analytical and numerical methods to gain further insights.

Small uncertainty

In this section we focus on the case in which uncertainty is small. A simple measure of the degree of uncertainty in the state $\boldsymbol{\theta}$ is given by $u \equiv E(||\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}||)$, where $\bar{\boldsymbol{\theta}}$ is the expected value of $\boldsymbol{\theta}$ and $||\cdot||$ is the 1-norm. We will consider the case in which u is small. We keep the assumption that there is positive uncertainty in both σ and γ, so that the EWFB contract is the one identified in the previous proposition.

In the next proposition we let $\Omega^{\max}_{\{K\}}$ denote the maximum expected gross global welfare that

[10] In particular, these government objectives rule out the possibility of political economy motives. In the presence of such motives, the EWFB contract would in general depend on all (eight) state variables in the model, and the characterization of optimal incomplete contracts would as a consequence be considerably more complex. We leave the introduction of political economy motives into our incomplete contracting setting for future work.

[11] The approach to contracting taken by the GATT/WTO may be viewed as somewhere in between the two extremes of contracting over policies and contracting over economic magnitudes such as prices and quantities directly. In a later section, we will consider more complex contracting possibilities in an effort to capture elements of the GATT/WTO approach.

can be obtained with a contract of type K.

Proposition 2. *Consider the contract class $\mathcal{K}_0$. There exists $\hat{u} > 0$ such that, if $u < \hat{u}$, then:*
For $c \in (0, c_0)$, the optimal contract is the EWFB contract $\{T = \gamma; S = \sigma - \gamma\}$.
For $c \in (c_0, c_1)$, the optimal contract is of the form $\{T = T(\cdot); S = S(\cdot)\}$, where T and S are contingent on a single state variable, either γ or σ.
For $c \in (c_1, c_2)$, the optimal contract is of the form $\{T = \bar{T}; S = \bar{S}\}$.
For $c \in (c_2, c_3)$, the optimal contract is of the form $\{T = \bar{T}\}$.
For $c > c_3$, the optimal contract is the empty contract.
The critical levels of c satisfy $0 < c_0 \leq c_1 < c_2 \leq c_3$, with $c_2 < c_3$ iff $\Omega^{\max}_{\{T=\bar{T}\}} > \frac{1}{2}\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} + \frac{1}{2}\Omega_{\{\emptyset\}}$.

The proof of this proposition and all others not proved in the text are contained in the Appendix. In order to describe in a general manner the way in which contractual incompleteness in trade agreements grows as contracting costs rise, we follow Battigalli and Maggi (2002) and identify two forms of incompleteness: excessive *rigidity*, which occurs when contractual obligations do not include state contingencies that are included in the EWFB; and excessive *discretion*, which occurs when contractual obligations that are included in the EWFB are simply missing (in what follows for simplicity we will omit the qualifier "excessive"). According to Proposition 2, in an environment of small uncertainty the contractual incompleteness of trade agreements evolves in a particular way as contracting costs rise: once we leave the EWFB contract, we first see the possibility of partial rigidity in T and/or S (when the contract is contingent on one state variable only), followed by full rigidity in both T and S, followed by the possibility of rigidity in T and discretion in S, and finally followed by complete discretion (the empty contract).

A general insight when uncertainty is small which is highlighted by Proposition 2 is that, as contracting costs rise, the optimal contract first becomes increasingly rigid, and then it becomes increasingly discretionary. Intuitively, as contracting costs rise, economizing on them becomes more attractive, raising two possibilities: remove states from the contract (hence increasing rigidity) or remove policy variables from the contract (hence increasing discretion). In the case of small uncertainty, the cost of rigidity (only being right "on average") is very small compared to the cost of discretion (permitting terms-of-trade manipulation), and so it is always better to economize on contracting costs by first removing state variables (as in $\{T = T(\cdot); S = S(\cdot)\}$ and finally $\{T = \bar{T}; S = \bar{S}\}$) and only then to begin removing policy variables (as in $\{T = \bar{T}\}$).

While we have described just above an insight that applies to the contract as a whole, Proposition 2 also has something interesting to say about the way contractual incompleteness varies across policy instruments. According to Proposition 2, the effective subsidy S tends to be more discretionary than T; more specifically, for a range of sufficiently high contracting costs it is optimal to contract over import taxes T while leaving the effective production subsidy S to discretion, but as Lemma 1 indicates and Proposition 2 confirms it is *never* optimal to leave T to discretion and contract only over S. In this way, Proposition 2 provides a general prediction that incomplete contracts negotiated in settings such as the GATT/WTO should always include commitments over import taxes, and should only introduce commitments over other internal measures as the contract becomes more complete. This prediction resonates to some degree with the broad approach taken by the GATT/WTO, which has been to first establish a base of commitments over import tax levels, and only later to broaden the contract to explicitly take on various internal measures. Notice, too, that our prediction does not rely on an assumption that embodies the commonly-held view that border measures are more transparent than internal measures and are therefore less costly to contract over, an assumption that would only reinforce this prediction.

The last point of the proposition states a condition under which the interval (c_2, c_3) is nonempty, that is, under which the contract $\{T = \bar{T}\}$ is optimal for some level of c.[12] The condition is that $\Omega^{\max}_{\{T=\bar{T}\}}$ must be closer to $\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}$ than to $\Omega_{\{\emptyset\}}$. This condition requires that there are diminishing gains from contracting over additional policy instruments. More specifically, this condition may be rewritten equivalently as $[\Omega^{\max}_{\{T=\bar{T}\}} - \Omega_{\{\emptyset\}}] > [\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} - \Omega^{\max}_{\{T=\bar{T}\}}]$, indicating that the contract $\{T = \bar{T}\}$ is optimal for some level of c if and only if switching from the empty contract to a contract that ties down T yields a higher (gross) gain than switching from a contract that ties down T to one that ties down both T and S.

We now go a step further and ask how changes in the underlying parameters of the model favor or disfavor the contract $\{T = \bar{T}\}$. We focus on the home-country parameters α, β and λ. To simplify, we consider the case of no uncertainty ($u = 0$), and point out where the introduction of a small amount of uncertainty would change the comparative-statics results we report below. With no uncertainty, the only contracts that we need to consider are $\{T = \bar{T}; S = \bar{S}\}$ (which in this case is the EWFB), $\{T = \bar{T}\}$, and the empty contract. We will examine how α, β and λ affect the cost interval over which $\{T = \bar{T}\}$ is optimal, that is (c_2, c_3), and the cost interval over which $\{T = \bar{T}; S = \bar{S}\}$ is optimal, that is $(0, c_2)$. A simple way to capture whether changes in underlying model parameters favor the contract $\{T = \bar{T}\}$ or rather the contract $\{T = \bar{T}; S = \bar{S}\}$ is to report how c_2/c_3 varies with the parameter of interest, with a rising c_2/c_3 indicating that the contract $\{T = \bar{T}; S = \bar{S}\}$ is favored by the parameter change relative to the contract $\{T = \bar{T}\}$. We first report our results in the following Remarks (which we prove with Maple calculations available on request), and then develop an interpretation of our findings.

Remark 1. *Assume* $u = 0$. *Then:*
c_2/c_3 *is weakly decreasing in* β, *with* $c_2/c_3 = 1$ *for* β *sufficiently small.*
c_2/c_3 *is weakly increasing in* λ, *with* $c_2/c_3 < 1/2$ *if* λ *is sufficiently small.*

Remark 2. *Assume* $u = 0$. *Then:*
c_2/c_3 *is independent of* α.

According to Remark 1, the contract $\{T = \bar{T}\}$ is disfavored relative to the contract $\{T = \bar{T}; S = \bar{S}\}$ as β falls, and disappears completely as an optimal contract (for any contracting costs) if β is sufficiently small. On the other hand, $\{T = \bar{T}\}$ is favored relative to $\{T = \bar{T}; S = \bar{S}\}$ as λ falls, and is sure to be an optimal contract for a greater range of contracting costs than

[12] The proposition does not state conditions under which the interval (c_0, c_1) is nonempty, that is under which the contract $\{T = T(\cdot); S = S(\cdot)\}$ is optimal for some level of c. Hence, as c increases from zero towards c_1, it *may* be optimal to switch from the EWFB contract to a contract that is contingent on a single state variable, either γ or σ. Whether this is the case and which of the two state variables is chosen depends on parameters, and in particular on the degree of uncertainty in σ relative to that in γ.

is $\{T = \bar{T}; S = \bar{S}\}$ if λ is sufficiently small. Finally, according to Remark 2, α has no bearing on the relative attractiveness of the contracts $\{T = \bar{T}\}$ and $\{T = \bar{T}; S = \bar{S}\}$.

To interpret the findings of Remarks 1 and 2, it is useful to introduce and characterize a particular feature of the underlying contracting environment, namely, *the degree of substitutability between policy instruments*. We consider how the results of Remark 1 can be interpreted through the effect of changes in the underlying model parameters on this feature. As we now explain, as a general matter high substitutability across instruments tends to raise the cost of discretion over S, and therefore tends to favor the contract $\{T = \bar{T}; S = \bar{S}\}$ over the contract $\{T = \bar{T}\}$. In fact, at a very basic level we have already seen that the degree of instrument substitutability plays an important role in contract design. To confirm this, recall that a comparison of the Nash policy combination $\{\tau = \frac{p^*}{\eta^*}, t_h = t_f = \gamma, s = \sigma\}$ and the efficient policy combination $\{\tau = 0, t_h = t_f = \gamma, s = \sigma\}$ would suggest that the fundamental contracting problem for a trade agreement to solve is to place constraints on the border measure τ so as to eliminate terms-of-trade manipulation. However, as we have already observed, τ and t_f are perfectly substitutable policy instruments, comprising a single composite instrument $T \equiv \tau + t_f$, and so controlling terms-of-trade manipulation requires at a minimum placing constraints on T. To move beyond this basic implication, we now observe that constraints on T alone can be effective in controlling terms-of-trade manipulation only if the degree of substitutability is low between T and the remaining internal measures s and t_h which, as we have observed previously, are themselves perfectly substitutable and comprise a single composite instrument $S \equiv s - t_h$. Consequently, whether or not there exists a level of contracting costs for which it is optimal to leave the choice of S to the government's discretion depends on the degree of substitutability between the import tax T and the effective production subsidy S: if S is effective at manipulating terms of trade when T is constrained, i.e. if discretion in S can be used to largely "undo" the constraint on T, then $\Omega^{\max}_{\{T=\bar{T}\}}$ will be closer to $\Omega_{\{\emptyset\}}$ than to $\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}$, and the contract $\{T = \bar{T}\}$ will be dominated for all contracting cost levels.

The degree of substitutability between the two instruments T and S can be captured formally in the following way. First recall that, if the contract dictates an import tax level T but leaves discretion over S, the home government will choose S according to the best-response function $S^R(T)$, and therefore the associated level of gross global welfare will be $\Omega(T, S^R(T))$. In this case, then, a change in the agreed-upon level of T has two effects on Ω: a direct effect, captured by the partial derivative $\partial\Omega/\partial T$, and an indirect effect through the induced change in S, captured by $(\partial\Omega/\partial S)(dS^R/dT)$. The total effect of a change in T is given by the total derivative $d\Omega/dT = \partial\Omega/\partial T + (\partial\Omega/\partial S)(dS^R/dT)$. When T and S are close substitutes, the impact on Ω of a change in T will be largely offset by the induced change in S, and hence $d\Omega/dT$ will be close to zero. On the other hand, if T and S are poor substitutes, the indirect effect through S will be small and therefore $d\Omega/dT$ will be close to $\partial\Omega/\partial S$.

The above discussion suggests that the ratio $\frac{d\Omega/dT}{\partial\Omega/\partial T}$ can serve as a useful index of (inverse) instrument substitutability between T and S: when T and S are close substitutes this ratio is close to zero, and when T and S are poor substitutes it is close to one. It is natural to evaluate this ratio at the Nash equilibrium policies, (T^{NE}, S^{NE}). This leads to the following index of (inverse) instrument substitutability between T and S:

$$\left(\frac{d\Omega(T,S^R(T))/dT}{\partial\Omega(T,S^R(T))/\partial T}\right)_{T=T^{NE}} = \frac{\beta A^2}{(\beta+\lambda)(A^2-\lambda)}$$

It can be shown that this index takes on a value of zero when $\beta = 0$, a value of one when $\lambda = 0$, and is monotonically increasing in β and decreasing in λ. We note as well that this index is independent of α.

Accordingly, we may say that the degree of substitutability between T and S is decreasing in β, increasing in λ, and independent of α.[13] Recalling now that high substitutability between T and S raises the cost of discretion over S and hence disfavors the contract $\{T = \bar{T}\}$, we may conclude that reductions in β work against the optimality of the contract $\{T = \bar{T}\}$ through the instrument-substitutability effect (i.e., reductions in β tend to raise c_2/c_3), reductions in λ favor the optimality of the contract $\{T = \bar{T}\}$ through the instrument-substitutability effect (i.e., reductions in λ tend to reduce c_2/c_3), and changes in α have no impact on the optimality of the contract $\{T = \bar{T}\}$ through the instrument-substitutability effect (i.e., changes in α leave c_2/c_3 unchanged). These observations suggest that the instrument-substitutability effect provides a useful interpretation of the findings in Remarks 1and 2.

At this point it might be tempting to conclude that the degree of substitutability between policy instruments is the only feature of the contracting environment that needs to be understood for interpreting the impacts of changes in α, β and λ on the nature of the optimal contracts. And for the case of $u = 0$ this conclusion seems warranted. However, as we report in the next section, for $u > 0$, the independence of c_2/c_3 in α no longer holds, indicating that the finding of Remark 2 is special to the $u = 0$ environment, and suggesting that other features in addition to the instrument-substitutability effect are at work in determining the nature of the optimal contracts. For this reason, we now consider how a change in α can alter the nature of the optimal contracts through its impact on a second feature of the contracting environment, namely, *the magnitude of the foreign export supply elasticity* (evaluated at the Nash equilibrium). As we next explain, as a general matter the cost of discretion over S tends to rise as α rises and the magnitude of the foreign export supply elasticity falls. This by itself would tend to disfavor the contract $\{T = \bar{T}\}$ relative to the contract $\{T = \bar{T}; S = \bar{S}\}$. But in the case of $u = 0$ this cost rises *proportionally* to the total gains from contracting, and so in this case there is no implied effect on c_2/c_3, as Remark 2 indicates.

To understand, recall that if the contract dictates an import tax level T but leaves discretion over S, the home government will choose S according to the best-response function $S^R(T)$. As we have observed previously, this best response is equal to S^{eff} when the home government is

[13] Intuitively, it may be seen that the substitutability between T and S is very high when β is very low by recalling that the potential gains from contracting in this setting arise entirely from the ability to control the incentive to manipulate the terms of trade, and that the terms of trade is given by $p^*(T,S) = [\alpha + \alpha^* - (\beta + \lambda)T - \lambda S]/A$. Hence, when β is very low, T affects the terms of trade mostly through the supply channel, as S does, and so S is a good substitute for T for purposes of terms-of-trade manipulation. Similarly, it may be seen intuitively that S is not a good substitute for T for purposes of terms-of-trade manipulation when λ is very low, because in this case S has very little ability to influence the terms of trade.

on its import–tax reaction curve, but more generally it is direct to verify that

$$S^R(T) = S^{eff} + \frac{A^2 - (\beta+\lambda)^2}{A^2 - \lambda(\beta+\lambda)}(T^{NE} - T). \tag{3.1}$$

From equation 3.1 it follows that the difference between $S^R(T)$ and S^{eff} is directly proportional to the distance between the home government's best-response import tax T^{NE} and T, with the factor of proportionality independent of α. Therefore, changes in α which increase the difference between T^{NE} and T^{eff}, which is to say changes in α which decrease the magnitude of the Nash equilibrium foreign export supply elasticity, increase the difference between $S^R(T^{eff})$ and S^{eff} and hence increase the cost of discretion over S. [14] Intuitively, if the magnitude of the Nash equilibrium foreign export supply elasticity is low, then the incentive to alter policies to manipulate the terms of trade is high, and so controlling this incentive with a contract over T but leaving S to discretion will result in costly distortions in S for purposes of manipulating the terms of trade.

Observing now that increases in α reduce the Nash equilibrium foreign export supply elasticity, we may conclude that increases in α raise the cost of discretion over S through the export-supply-elasticity effect. Indeed, when $u = 0$ we find that $\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} - \Omega^{\max}_{\{T=\bar{T}\}}$ is increasing in α, confirming this prediction. However, as indicated above, when $u = 0$ we find as well that $\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} - \Omega_{\{\emptyset\}}$ and therefore $\Omega^{\max}_{\{T=\bar{T}\}} - \Omega_{\{\emptyset\}}$ increases proportionally in α. Hence, in the case of $u = 0$, an increase in α increases proportionally the cost intervals over which each of the contracts $\{T = \bar{T}\}$ and $\{T = \bar{T}; S = \bar{S}\}$ is optimal, and in this sense does not affect the relative performance of the two contracts (i.e., does not effect c_2/c_3). As we will establish in the next section, however, this proportionality property is special to the $u = 0$ environment: with $u > 0$, changes in α alter the relative performance of the various contracts, and the implications of changes in α for the nature of the optimal contracts can then be interpreted through the export-supply-elasticity effect.

We close our discussion of the small uncertainty case by observing that Proposition 2 predicts $c_1 < c_2$, which is to say that the interval of contracting costs for which the contract $\{T = \bar{T}; S = \bar{S}\}$ is optimal can never shrink to zero. As we have observed above, this carries with it the important implication that, as contracting costs rise, the optimal contract first becomes increasingly rigid, and then it becomes increasingly discretionary, in the sense that it is always better to economize on contracting costs by first removing state variables (as in $\{T = T(\cdot); S = S(\cdot)\}$ and finally $\{T = \bar{T}; S = \bar{S}\}$) and only then to begin removing policy variables (as in $\{T = \bar{T}\}$). It is instructive to explore further why this must be so in the small uncertainty environment. There are two cases to consider. The first case is where $c_0 < c_1$. In this case, it is direct to show that $c_1 < c_2$ if and only if $k \cdot [\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} - \Omega^{\max}_{\{T=\bar{T}\}}] > 2 \cdot [\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} -$

[14] In our model, the foreign government is assumed to have no (export-sector) policy options, and so the home government's best-response import tax and the Nash equilibrium import tax T^{NE} are in fact one and the same. More generally, if the foreign government could also intervene in this sector, the foreign export supply elasticity relevant for gauging the cost of home-government discretion over S would be either (i) the Nash equilibrium foreign export supply elasticity, if the foreign policies were left to discretion in the contract, or (ii) the foreign export supply elasticity evaluated at the home-government best-response policy choices and the contracted foreign policy choices, if the foreign policies were specified in the contract.

$\Omega^{\max}_{\{T=T(\cdot);S=S(\cdot)\}}]$. The second case is where $c_0 = c_1$. In this case, it is direct to show that $c_1 < c_2$ if and only if $k \cdot [\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} - \Omega^{\max}_{\{T=\bar{T}\}}] > [\Omega_{\{EWFB\}} - \Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}]$. In each case, it can be seen that the magnitude on the right-hand-side of the inequality is zero when there is no uncertainty (when $u = 0$) and can be made arbitrarily close to zero when uncertainty is present but small (when $u > 0$ but small). Together with our other parameter restrictions, which ensure that the left-hand-side of each inequality is strictly positive, it follows that $c_1 < c_2$ for the small uncertainty case, as Proposition 2 indicates.

But it should now be clear that the conditions that ensure that $c_1 < c_2$ in the case of small uncertainty can be easily violated when uncertainty is large, and therefore that large uncertainty might create conditions in which the contract $\{T = \bar{T}; S = \bar{S}\}$ fails to be optimal for any level of contracting costs. Intuitively, the role of uncertainty in this regard can be understood by recalling the two possible methods of economizing on overall contracting costs as contracting costs rise, namely, removing states from the contract (hence increasing rigidity) or removing policy variables from the contract (hence increasing discretion). As we have observed previously, in the case of small uncertainty the cost of rigidity (only being right "on average") is very small compared to the cost of discretion (permitting terms-of-trade manipulation), and so the progression described above for economizing on overall contracting costs is clearly indicated. But with large uncertainty, it is easy to see that the cost of rigidity could become very large relative to the cost of discretion, raising the possibility that a contract such as $\{T = \bar{T}; S = \bar{S}\}$ could become disfavored while permitting discretion (over S) might instead become the most attractive way to begin to economize on contracting costs. As we demonstrate in the next section, this is indeed possible in a large uncertainty environment.

Large uncertainty

We next extend our investigation of the optimal contract to an environment in which the magnitude of uncertainty may be large. Without additional assumptions, this could become very tedious because of the large number of potential contracts that would need to be discussed and evaluated. To keep the characterization of the optimal contract manageable and to emphasize the key new features that arise in the large-uncertainty environment, we therefore work with a variant of the model considered in the previous section in which all state variables are deterministic except two: one state variable that is relevant for the EWFB contract, and one state variable that is not. We assume that each of the two uncertain state variables can in turn take on either of two values with equal (and independent) probability. For most of this section we consider a scenario in which the two uncertain state variables are the home-country demand intercept α and the consumption externality parameter γ. At the end of this section we will report results for an alternative scenario in which the uncertain state variables are the home-country demand intercept α and the production externality parameter σ. To summarize, the particular support of θ that we now consider is described by:

$$\mathbf{\Theta} = \{\alpha \in \{\bar{\alpha}-\Delta_\alpha, \bar{\alpha}+\Delta_\alpha\}, \alpha^* = \bar{\alpha}^*, \beta = \bar{\beta}, \beta^* = \bar{\beta}^*, \lambda = \bar{\lambda}, \lambda^* = \bar{\lambda}^*, \sigma = \bar{\sigma}, \gamma \in \{\bar{\gamma}-\Delta_\gamma, \bar{\gamma}+\Delta_\gamma\}\},$$

with $\Delta_\alpha > 0$ and $\Delta_\gamma > 0$ by assumption and possibly large.[15]

[15] In allowing Δ_α and Δ_γ to be large, we require only that their magnitudes remain below the levels that, in combination with the other parameter values, would result in a prohibitive Nash equilibrium.

In this setting, the EWFB contract is given by $\{T = \gamma; S = \bar{\sigma} - \gamma\}$, that is, a state-contingent import tax commitment and a state contingent effective subsidy, which costs $(4 + k)c$. For simplicity of exposition, we set $k = 1$ for the remainder of this section, though our main points are unaffected by this simplification. With this, the EWFB contract costs $5c$, and so the only additional contracts we need consider are those costing $4c$, $3c$ and $2c$ (recall that contracts costing c are dominated by the empty contract, owing to the perfect substitutability between τ and t_f and between s and t_h). Recalling from Lemma 1 that it can never be optimal to constrain S but not T, we then have three kinds of contracts to consider (in addition to the EWFB and the empty contract), corresponding to the three cost classes listed just above. In particular there are two contracts costing $4c$: $\{T = \bar{T}; S = \bar{S}\}$; and $\{T = f(\alpha, \gamma)\}$. There are two contracts costing $3c$: $\{T = f(\gamma)\}$; and $\{T = f(\alpha)\}$. And there is one contract costing $2c$: $\{T = \bar{T}\}$.

Our approach to characterizing the optimal contract is now straightforward. We begin by optimizing each contract. That is, we solve for the values of $\bar{T}$ and $\bar{S}$ (as functions of the parameters of the model) that yield the maximum expected gross global welfare under the contract $\{T = \bar{T}; S = \bar{S}\}$, and we perform a similar calculation for each of the five contracts listed above. For given parameter values, we then select the optimized contract in each cost class that yields the maximum expected gross global welfare. At this point, for the given parameter values and for each of the cost classes $4c$, $3c$ and $2c$, we have characterized the preferred contract and its associated expected gross global welfare. We also calculate the expected gross global welfare associated with the EWFB contract (which costs $5c$) and the empty contract (which costs zero). Then, for a level of contracting costs c, we identify the optimal contract by searching over the preferred contracts in each of the five cost classes ($5c$, $4c$, $3c$, $2c$, and zero) for the contract that yields the maximum expected net global welfare.

In analogy with Proposition 2 of the previous section, we wish to emphasize the way in which the optimal contract changes as contracting costs rise from zero (where the EWFB contract is optimal) to the prohibitive level (where the empty contract is optimal). Moreover, as in Remarks 1 and 2 of the previous section, we are interested in the impact that various underlying model parameter values have in favoring one contractual form over another. To this end, we exploit the additional structure that we have imposed in this section and report explicitly the way in which the nature of the optimal contract varies with both the level of contracting costs and the level of various parameters of interests. We do this using a series of figures, in which we plot on the vertical axis the level of contracting costs, normalized by the prohibitive level of contracting costs, and on the horizontal axis the value of the parameter of interest. We then report the optimal contractual forms over this entire parameter space. These figures are a useful way of summarizing how the fraction of non-prohibitive contracting costs for which a given contractual form is optimal is impacted by changes in the underlying parameter of interest. As in the previous section, we interpret an increase (drop) in this fraction as indicating that the associated contractual form is favored (disfavored) by the change in the underlying parameter.

We focus in sequence on four parameters of interest in order to emphasize the new features that arise in the large-uncertainty environment. In evaluating the impact of each parameter of interest, the remaining parameters are set at values which are convenient for illustrating these new features, but we emphasize that these results are not "fragile": the features we illustrate below arise for a wide range of the remaining parameters consistent with a large-uncertainty environment. The four parameters of interest are: (i) the expected value of the home-country demand intercept, $\bar{\alpha}$; (ii) the value of the home-country demand slope, $\bar{\beta}$; (iii) the level of uncertainty over the home-country consumption externality, as measured by the magnitude of Δ_γ; and (iv) the level of uncertainty over the home-country demand intercept, as measured by the magnitude of Δ_α. As we next demonstrate, changes in $\bar{\alpha}$ can illustrate the novel features of the export-supply-elasticity effect that arise when uncertainty is large (or simply non-zero), changes in $\bar{\beta}$ can illustrate the novel features of the instrument-substitutability effect that arise when uncertainty is large, and changes in Δ_γ and Δ_α can illustrate the distinctive impacts on optimal contractual form that are associated with large uncertainty over state variables which are, in the first instance, relevant for the EWFB contract, and in the second instance, not relevant for the EWFB contract.

To understand the potential impacts of large uncertainty on the optimal contractual form, it is useful to start from the small-uncertainty environment. If we were to restrict our attention here to the case of small uncertainty, in the sense that both Δ_γ and Δ_α were assumed small, then it is easy to see that, of the contracts costing $4c$, the contract $\{T = \bar{T}; S = \bar{S}\}$ would dominate the contract $\{T = f(\alpha, \gamma)\}$, and that each of the contracts costing $3c$ ($\{T = f(\gamma)\}$ and $\{T = f(\alpha)\}$) would be dominated by the single contract costing $2c$, namely $\{T = \bar{T}\}$. In this case, similarly to the more general treatment of the small-uncertainty case considered in the previous section, the optimal contract would switch from the EWFB contract to $\{T = \bar{T}; S = \bar{S}\}$ to $\{T = \bar{T}\}$ and finally to the empty contract as contracting costs rise from zero to the prohibitive level, with the possibility of "skipping over" the contract $\{T = \bar{T}\}$ as indicated (for example) in Remark 1. When uncertainty is instead large, however, the relative attractiveness of the various contracts is altered, and the structure of the optimal contract for various contracting cost levels can be altered as a result.

To see this, consider first the possibility that uncertainty over γ is large ($\Delta_\gamma >> 0$). Because the EWFB contract calls for T and S that are contingent on γ, a high degree of uncertainty over γ disfavors contracts that feature rigidity in either instrument ($\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$), raising a novel possibility: *either or both* of the contracts $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$ can become dominated for all contracting costs by contracts that feature greater state-contingency and/or greater discretion. As we illustrate below, this novel possibility can be interpreted through the export-supply-elasticity effect when $\bar{\alpha}$ is sufficiently low, and it can be interpreted through the instrument-substitutability effect when $\bar{\beta}$ is sufficiently high.

Consider next the possibility that uncertainty over α is large ($\Delta_\alpha >> 0$). Because the EWFB contract calls for T and S that are independent of α, it might be thought that the degree of uncertainty over α should be irrelevant for the relative attractiveness of the various contracts, and hence irrelevant for the structure of the optimal contract. However, this is not the case, and the reason can be understood to reflect the impact of uncertainty over α on the export-supply-elasticity effect. In particular, as α increases, the cost of discretion over S rises for the high-demand state and falls for the low-demand state according to the export-supply-elasticity effect, implying that it is then more valuable to have contracted over S in high-demand states but less-valuable to have done so in low-demand states. However, it can be shown that the value of contracting over S is convex in α, and so it follows that the expected cost of discretion over S is rising in Δ_α. Moreover, when S is left to discretion the value of permitting T to be contingent on γ when T is not also contingent on α falls as Δ_α rises, due again to the impact of uncertainty over α on the export-supply-elasticity effect. Hence, a higher Δ_α tends to favor the contracts $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$ over contracts that feature greater state-contingency and/or greater discretion. The upshot, then, is that a rising Δ_α has impacts on the optimal contractual form which are *essentially opposite* the impacts of a rising Δ_γ.

Having described the general nature of our novel large-uncertainty findings, we now turn to Figures 1a-1d, and describe these findings in some detail. Figure 1a depicts the impact of $\bar{\alpha}$ on the optimal contractual form in the presence of large uncertainty for $\bar{\alpha} \in [11.5, 100]$.[16] The lower bound on $\bar{\alpha}$ is determined by the requirement of positive Nash trade volume in each state. As discussed above and as Figure 1a confirms, reductions in $\bar{\alpha}$ disfavor the contracts $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$, and for sufficiently low levels of $\bar{\alpha}$ these contracts disappear as optimal contractual forms. This can interpreted from the perspective of the export-supply elasticity effect. In particular, for $\bar{\alpha}$ in this low range, the Nash equilibrium foreign export supply elasticity is very high and hence the cost of discretion over S is very low: for this reason, as contracting costs rise and it becomes optimal to switch from the EWFB contract, it is never optimal (for any contracting costs, in a large uncertainty environment) to switch to a contract that rigidly binds S, and so the contract $\{T = \bar{T}; S = \bar{S}\}$ disappears as an optimal contractual form. Instead, owing to the very low cost of discretion over S, overall contracting costs can be reduced with little loss in gross global welfare by switching from the EWFB to the contract $\{T = f(\gamma)\}$. And from here, as contracting costs rise further, the saving in overall contracting costs that could be achieved with a switch to the contract $\{T = \bar{T}\}$ is never (for any contracting costs, in a large uncertainty environment) worth the reduction in expected gross global welfare that this switch would entail, and so the contract $\{T = \bar{T}\}$ disappears as an optimal contractual form as well. Instead, eventually contracting costs rise to the prohibitive level where the empty contract becomes optimal. We note as well that the finding of Remark 2 for the case of no uncertainty is not borne out by Figure 1a when uncertainty is large (and indeed is not borne out for any $\Delta_\gamma > 0$), once $\bar{\alpha}$ has reached a level where both $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$ appear as optimal contracts: as $\bar{\alpha}$ rises beyond this level, the cost interval over which the contract $\{T = \bar{T}\}$ is optimal shrinks monotonically relative to the cost interval over which the contract $\{T = \bar{T}; S = \bar{S}\}$ is optimal.

More broadly, the finding that, for low $\bar{\alpha}$, it is never optimal to incur the extra contracting cost in order to impose a rigid constraint on the subsidy, reflects a basic trade-off across rigidity and discretion as cost-saving contracting options regarding the effective production subsidy. On the one hand, the rigid subsidy contract can restrain the opportunistic (terms-of-trade manipulating) setting of the subsidy, but it cannot facilitate the state-contingent nature of the first-best Pigouvian subsidy. On the other hand, leaving the subsidy to the discretion of the home government does not succeed in restraining the opportunistic setting of the subsidy, but it does allow the subsidy to be state-contingent, and thereby facilitates the state-contingent nature of the Pigouvian subsidy. In the case of low $\bar{\alpha}$, which is to say the case in which the

[16] To generate Figure 1a, the remaining parameters are set to the following values: $\Delta_\alpha = 3; \bar{\alpha}^* = 10; \bar{\beta} = 1; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 0; \bar{\gamma} = 1; \Delta_\gamma = 1$.

magnitude of the Nash equilibrium foreign export supply elasticity is high, we have observed that the home government has little unilateral incentive to distort its subsidy away from the Pigouvian level in an opportunistic (terms-of-trade manipulating) fashion. Hence, in this case, it is better to leave the subsidy to the discretion of the home government than to bind it rigidly. For this reason, once the magnitude of the elementary contracting cost c rises beyond a critical level, the EWFB contract with state-contingent import tax and effective production subsidy commitments is abandoned in favor of a state-contingent import tax commitment and no restraints on subsidies.

The second parameter of interest is the value of the home-country demand slope, $\bar{\beta}$. As pointed out in the previous section, in the case of small uncertainty a decrease in β raises the cost of discretion over S and therefore works against the optimality of the contract $\{T = \bar{T}\}$ through the instrument-substitutability effect. We now show in the case of large uncertainty that the cost of discretion over S can also be reduced through the instrument-substitutability effect to a sufficient degree that the contracts $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$ disappear as optimal contractual forms. Figure 1b depicts the impact of $\bar{\beta}$ on the optimal contractual form in the presence of large uncertainty for $\bar{\beta} \in [0, 1.5]$.[17] The upper bound on $\bar{\beta}$ is determined by the requirement of positive Nash trade volume in each state. Observe first that Figure 1b is consistent with the predictions of Remark 1: for sufficiently low values of $\bar{\beta}$, the cost of discretion over S is sufficiently high that the cost interval over which the contract $\{T = \bar{T}\}$ is optimal eventually disappears in favor of the contract $\{T = \bar{T}; S = \bar{S}\}$. On the other hand, Figure 1b also demonstrates that when $\bar{\beta}$ is sufficiently high in a large-uncertainty environment, the cost of discretion over S is reduced to the point where: first, the contract $\{T = \bar{T}; S = \bar{S}\}$ disappears as an optimal contract; second, the contract $\{T = f(\gamma)\}$ appears as an optimal contract for a range of intermediate contracting costs; and third, the contract $\{T = \bar{T}\}$ disappears as an optimal contract. More broadly, as Figure 1b indicates, when the substitutability across instruments is sufficiently low (when $\bar{\beta}$ is sufficiently high), it is never optimal (for any contracting costs) in a large-uncertainty environment to depart from the EWFB contract with a contract that rigidly binds S, because in this circumstance the costs of discretion over S are low while the costs of rigidity over S are high, and so any departure from the EWFB should simply leave S to discretion.

The last two parameters of interest are Δ_γ and Δ_α, which determine respectively the levels of uncertainty over the home-country consumption externality and demand intercept. Figure 1c characterizes the impact of Δ_γ on the optimal contractual forms for $\Delta_\gamma \in [0, 3.5]$, while Figure 1d characterizes the impact of Δ_α on the optimal contractual forms for $\Delta_\alpha \in [0, 8.5]$. In each case, the upper bound on the parameter of interest is determined by the requirement of positive Nash trade volume in each state.[18] As can be seen from a comparison across Figures 1c and 1d, the two forms of uncertainty have opposing effects on the desirability of the contracts $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$. As described above and as Figure 1c indicates, because the

[17] To generate Figure 1b, the values of $\bar{\gamma}$ and Δ_γ are set as described in the text, while the remaining parameters are set to the following values: $\bar{\alpha} = 20; \Delta_\alpha = 0; \bar{\alpha}^* = 10; \bar{\beta} = 1; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 0$.

[18] To generate Figure 1c, the remaining parameters are set to the following values: $\bar{\alpha} = 18; \Delta_\alpha = 3; \bar{\alpha}^* = 10; \bar{\beta} = 1; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 0; \bar{\gamma} = 4$. To generate Figure 1d, the remaining parameters are set to the following values: $\bar{\alpha} = 18; \bar{\alpha}^* = 10; \bar{\beta} = 1; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 0; \bar{\gamma} = 1; \Delta_\gamma = 1$.

EWFB contract calls for T and S that are contingent on γ, increasing uncertainty over γ tends to favor contracts that feature greater state-contingency and/or greater discretion, namely, the EWFB contract and the contract $\{T = f(\gamma)\}$, and to disfavor the contracts that feature rigidity in one or both instruments, namely, $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$. By contrast, as Figure 1d indicates, increasing uncertainty over α tends to favor contracts that feature rigidity in one or both instruments ($\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$) over contracts that feature greater state-contingency and/or greater discretion (the EWFB contract and $\{T = f(\gamma)\}$), and this can be interpreted as reflecting the impact of uncertainty over α on the export-supply-elasticity effect, as we have discussed above.

More generally, Figures 1c and 1d indicate that contracting in an environment of substantial uncertainty can lead to very different contracting outcomes depending on the nature of that uncertainty. When the uncertainty concerns the levels of state variables that are themselves directly relevant to the EWFB contract, the preferred contracting options will tend to feature state-contingency and/or discretion, due essentially to the direct impact that this kind of uncertainty has on the relative performance of the various contracting options. However, when the uncertainty concerns the levels of state variables which are themselves not directly relevant to the EWFB contract, the preferred contracts may ironically feature greater rigidity than would otherwise be the case, due to the above-described indirect effects of this kind of uncertainty on the relative performance of the various contracting options.

We now briefly describe results for an alternative scenario in which the uncertain state variables are the home-country demand intercept α and the production externality parameter σ. The particular support of θ that we now consider is described by:

$$\mathbf{\Theta} = \{\alpha \in \{\bar{\alpha}-\Delta_{\alpha}, \bar{\alpha}+\Delta_{\alpha}\}, \alpha^* = \bar{\alpha}^*, \beta = \bar{\beta}, \beta^* = \bar{\beta}^*, \lambda = \bar{\lambda}, \lambda^* = \bar{\lambda}^*, \sigma \in \{\bar{\sigma}-\Delta_{\sigma}, \bar{\sigma}+\Delta_{\sigma}\}, \gamma = \bar{\gamma}\},$$

with $\Delta_{\alpha} > 0$ and $\Delta_{\gamma} > 0$ by assumption and possibly large. The EWFB contract is now given by $\{T = \bar{\gamma}; S = \sigma - \bar{\gamma}\}$, that is, a rigid import tax commitment and a state contingent effective subsidy, which costs $5c$.

Under this alternative scenario, we generate figures analogous to Figures 1a-1d, with σ replacing the role of γ. However, rather than report these additional figures, we simply describe the qualitative differences relative to Figures 1a-1d. This is simple, because each of the qualitative features of Figures 1a-1d are preserved under our alternative scenario with one exception: under the alternative scenario, the contract $\{T = f(\gamma)\}$ is never optimal. This difference can be easily understood to reflect the distinction between the EWFB contracts across the two scenarios, namely, that the EWFB contract under our first scenario features a state-contingent import tax commitment while the EWFB contract under our alternative scenario features a rigid import tax commitment. Hence, under our alternative scenario, the reason for the attractiveness of the contract $\{T = f(\gamma)\}$ is simply absent. Aside from this difference, all statements above describing and interpreting the first scenario continue to apply under our alternative scenario.

We now take stock of our findings thus far, and consider at a very coarse level their possible implications for interpreting the contracting approaches taken by the GATT and now the WTO. An interesting feature of the approaches taken by the GATT and the WTO is that these approaches can be viewed as somewhat distinct and evolving. In particular, a possible interpretation of the GATT contracting approach is that it applied a rigid rule for tariffs but left internal measures such as subsidies in effect to the unilateral discretion of each government (we will offer a more nuanced interpretation in a later section). By contrast, it may be argued that the WTO has taken on the task of contracting rigid rules over certain internal measures, most notably subsidies.

At a broad level, our theory could account for this evolution in two ways: either as reflecting a drop in the magnitude of the elementary contracting costs c, or as reflecting changes in the underlying parameters that served to (i) exacerbate the incentives for distorting the choice of internal measures for terms-of-trade gain, (ii) increase the substitutability between import taxes and remaining internal measures, and/or (iii) decrease the uncertainty over the level of externalities. The first interpretation corresponds to a vertical downward movement in any of Figures 1a-1d, and can be confirmed by observing that, for any fixed parameter values, the range of contracting costs for which $\{T = \bar{T}; S = \bar{S}\}$ is optimal always lies strictly below the range of contracting costs for which $\{T = \bar{T}\}$ is optimal. The second interpretation corresponds to a horizontal movement in Figures 1a-1d, and can be confirmed respectively by observing that the contract $\{T = \bar{T}\}$ is disfavored relative to the contract $\{T = \bar{T}; S = \bar{S}\}$ with (i) a leftward movement in Figures 1a and 1d, (ii) a leftward movement in Figure 1b, and (iii) a rightward movement in Figure 1c.[19]

It is also noteworthy to observe that each of the factors which we have identified as increasing the cost of discretion and hence favoring the contract $\{T = \bar{T}; S = \bar{S}\}$ over $\{T = \bar{T}\}$ – the instrument-substitutability effect and the export-supply-elasticity effect – could arguably be seen to apply most aptly to relatively large developed countries. The essence of the instrument-substitutability effect is that a government has access to a rich array of internal measures which it can use to carry on terms-of-trade manipulation should its border measures be constrained through an international trade agreement. This is arguably more likely to be true of developed countries than developing countries. And the essence of the export-supply-elasticity effect is that a country is large in world markets, so that it faces foreign export supply that is far from perfectly elastic. Again, this is arguably more likely to be true of large developed countries. When viewed in light of our discussion above, this observation is potentially interesting, because it suggests that the attractiveness of contracting over internal measures (such as those embodied in S in our formal model) may be different for large developed countries than for small developing countries, and in particular that optimal contractual design might entail small developing countries negotiating trade agreements of the form $\{T = \bar{T}\}$ while large developed countries negotiate trade agreements of the form $\{T = \bar{T}; S = \bar{S}\}$. While our two-country model cannot really address this issue formally, our results are at least suggestive of the possible benefits of a kind of "special and differential treatment" for small/developing countries when it comes to contracting over internal measures (as opposed to border measures).[20]

We will return to these possible interpretive observations at several later points in the paper.

[19] By "horizontal movement" we mean tracking how a change in the parameter on the horizontal axis affects the cost intervals over which the various contracts are optimal.

[20] We thank Robert Lawrence for first bringing this implication to our attention.

4. The Role of the National Treatment Clause

A distinguishing feature of the GATT/WTO (albeit increasingly less of the WTO) is that it combines rigid bindings on tariffs with a significant amount of discretion over internal measures. This discretion is not complete, however. There are two legal instruments that can be invoked by Members in order to attack internal measures: the National Treatment (NT) provision in GATT Article III and the Non-Violation (NV) nullification-or-impairment provisions in GATT Article XXIII.1(b). These instruments are very likely core features of the GATT/WTO from an incomplete contracts perspective. In this section we therefore evaluate the usefulness of the NT clause as a means to economize on contracting costs. In the next section we perform an analogous evaluation of the NV clause.

A complete evaluation might seek to define a meaningfully wider class of contracts which would include these legal instruments, and then ask whether there are conditions under which these instruments might be part of an optimal contract within this wider class. We do not provide such a complete evaluation here. Instead, we rely more heavily on the institutional motivation for studying these particular clauses, and attempt something more modest: we seek to understand the conditions under which adding the NT clause (in this section) and the NV clause (in the next section) to the contracting options embodied in the class of contracts $\mathcal{K}_0$ would be valuable to governments, and to understand the way that these additional contracting options would alter the optimal contractual forms.

In our setting, the NT clause is simply the constraint $t_h = t_f$.[21] A literal interpretation of our assumptions on contracting costs would imply that the NT clause costs $2c_p$, that is the same as specifying numbers for t_h and t_f, for example as in $\{t_h = 3, t_f = 5\}$. We believe however that it is more realistic to assume that the NT clause costs less than specifying numbers for two policy instruments. This would be especially compelling if we had many sectors, in which case specifying numbers for two policy instruments in all sectors would be vastly more complicated than specifying a blanket NT rule. We try to capture this type of consideration in our single-sector model by assuming that the NT clause costs only c_p, but we note that our qualitative results would hold more generally as long as the NT clause costs strictly less than $2c_p$.

We refer to a contract that includes the NT clause as an "NT-based" contract. As indicated above, in this section we focus on an extended set of contracts that includes the class considered in the previous section ($\mathcal{K}_0$) plus the class of NT-based contracts. We label this extended set of contracts $\mathcal{K}_1$.

We begin by observing that the relationships between price wedges and policies are different for non-NT contracts and NT-based contracts. For non-NT contracts, these relationships are the same as in the previous section:

$$p = p^* + T, \text{ and} \tag{4.1}$$

$$q = p^* + T + S. \tag{4.2}$$

Within this class, as we argued previously, we can focus on contracts that tie down S and/or T.

[21] In fact, the "treatment no less favorable" language of the NT clause indicates that a literal interpretation within our model would be embodied in the inequality constraint $t_h \geq t_f$. However, for simplicity we abstract here and more generally from inequality constraints until section 6.

However, for NT-based contracts, the relationships between price wedges and policies become

$$\begin{aligned} p &= p^* + \tau + t, \text{ and} \\ q &= p^* + \tau + s. \end{aligned} \tag{4.3}$$

Within this class, we can focus on contracts that tie down some or all of the instruments (τ, t, s).

In order to evaluate NT-based contracts, it is useful to derive expressions for the equilibrium prices and trade volume under NT. Equilibrium prices are easily derived to be

$$p(\tau, t, s) = [\alpha + \alpha^* + (\beta^* + \lambda^*)\tau - \lambda s + (\beta^* + \lambda + \lambda^*)t]/A,$$

$$q(\tau, t, s) = [\alpha + \alpha^* + (\beta^* + \lambda^*)\tau + (\beta + \beta^* + \lambda^*)s - \beta t]/A, \text{ and}$$
$$p^*(\tau, t, s) = q^* = [\alpha + \alpha^* - (\beta + \lambda)\tau - \lambda s - \beta t]/A,$$

where as before, $A \equiv \lambda + \lambda^* + \beta + \beta^*$, and market-clearing home import volume/foreign export volume is given by

$$M(\tau, t, s) = E^*(\tau, t, s) = [\alpha(\beta^*+\lambda^*) - \alpha^*(\beta+\lambda) - (\beta+\lambda)(\beta^*+\lambda^*)\tau - \beta(\beta^*+\lambda^*)t - \lambda(\beta^*+\lambda^*)s]/A.$$

Notice that the NT clause by itself has no real effect of any kind, because as we have observed previously τ and t_f are perfectly substitutable policy instruments, and so any constraints placed on t_f – such as the NT requirement that t_f remain equal to t_h – if applied in the absence of further constraints can always be undone by appropriate changes in τ. This point carries with it the important implication that the efficient outcome derived in the previous section in the absence of NT can be implemented also with policies that conform to the NT clause. In particular, it is direct to establish that the efficient policies under NT are given by:

$$\begin{aligned} \tau^{eff} &= 0, \\ t^{eff} &= \gamma, \text{ and} \\ s^{eff} &= \sigma. \end{aligned}$$

Finally, it is useful also to recall that the Nash policies can be equivalently expressed in a form that complies with the NT constraint, namely, $\{\tau^{NE} = \frac{p^*}{\eta^*}, t_h^{NE} = t_f^{NE} \equiv t^{NE} = \gamma, s^{NE} = \sigma\}$, and it is useful as well to record the expression for the best-response level of t for any τ and s:

$$t^R(\tau, s) = t^{eff} + \frac{A^2 - (\beta + \lambda)^2}{A^2 - \beta(\beta + \lambda)}(\tau^{NE} - \tau) - \frac{\lambda(\beta + \lambda)}{A^2 - \beta(\beta + \lambda)}(s^{eff} - s).$$

These expressions will be helpful in interpreting our results below.

There are potentially many kinds of NT-based contracts, but we can reduce the number that must be considered by recalling that our goal here is to characterize NT-based contracts that would have value to governments relative to the non-NT contracting options contained in the class of contracts $\mathcal{K}_0$. For this reason, we will focus only on NT-based contracts that are strictly optimal for some range of contracting costs. In this regard, we first observe that any NT-based contract that ties down only one policy instrument is empty; this follows immediately from the fact that there is a degree of redundancy in the NT price relationships (4.3). Moreover, any NT-based contract that ties down the level of the common consumption tax t, e.g. with a clause $t = \bar{t}$, must be equivalent to or strictly dominated by a non-NT contract in $\mathcal{K}_0$. To see why, note first that under our assumptions, specifying the NT clause and the clause tying down t would cost $2c_p$ (the NT clause costs c_p and the constraint on t costs an additional c_p).[22] But as we have just observed, from (4.3) it follows that, for the contract not to be empty, it would have to tie down at least one other instrument. If the additional instrument is τ, the contract will cost $3c_p$ and will be dominated by a non-NT contract that ties down T, which costs less and can implement the same outcome. If the additional instrument is s, again the contract will cost $3c_p$ and will be dominated by a non-NT contract that ties down S, since this costs less and can implement the same outcome. If the additional instruments are both τ and s, the contract will be equivalent to a non-NT contract that ties down T and S. The same argument applies if the constraint on t is contingent.

From the above discussion it follows that the only kind of NT-based contracts that could have value to governments relative to the non-NT contracting options in the class of contracts $\mathcal{K}_0$ are NT-based contracts that tie down τ and s, leaving the government free to choose the common level of the consumption tax t. Recalling now our parametrization of contracting costs ($c_p = c$, $c_s = k \cdot c$), and that under this parameterization the EWFB contract costs $(4 + 2k) \cdot c$, our focus with regard to the NT clause can therefore be restricted to NT-based contracts that leave t to discretion and constrain τ and s (in a rigid or contingent fashion), and to whether and under what conditions such NT-based contracts would be chosen over the non-NT contracts contained in $\mathcal{K}_0$.

As before, we start by deriving strong analytical results for the case of small uncertainty, and then turn to the case of large uncertainty where we use both analytical and numerical methods to gain further insights.

Small uncertainty

As we have explained above, we may restrict our focus with regard to the NT clause to NT-based contracts that leave t to discretion and constrain τ and s. Of course, this still leaves a very large number of potential NT-based contracts that in general must be evaluated. However, when we restrict attention to the case of small uncertainty, our problem becomes much simpler. In fact, when uncertainty is small, there is only *one* NT-based contract that has potential value relative to the non-NT contracts in $\mathcal{K}_0$: the NT-based contract that imposes rigid constraints on τ and s (costing $3c$), which we denote by $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$. To see why no NT-based contract with state-contingent constraints could be valuable relative to the non-NT contracts in $\mathcal{K}_0$ when the level of uncertainty is small, observe that any NT-based contract that leaves t to discretion yields a level of expected gross global welfare which is strictly lower than that associated with the EWFB. This means that there exists $\hat{c} > 0$ such that for $c \leq \hat{c}$ these contracts are dominated by the EWFB contract. And for $c > \hat{c}$ these contracts are dominated

[22] Note that when we assign cost $2c_p$ to the pair of clauses $(NT, t = \bar{t})$ we are not diverging from our baseline assumptions, since this is equivalent to $(t_h = \bar{t}, t_f = \bar{t})$, which according to our baseline assumptions costs $2c_p$. The only kind of contract for which we are amending our baseline assumptions is a contract that includes NT but not a constraint on t; in this case, according to our baseline assumptions the additional cost of the NT clause would be $2c_p$, but we are amending it to c_p.

by the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$, because the level of uncertainty is small and hence the cost of introducing contingencies in the contract exceeds the gain in expected gross global welfare relative to the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$.

Consider, then, the potential value of the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ relative to the non-NT contracts contained in $\mathcal{K}_0$. As we have already seen from Proposition 2, the non-NT contracts that can be optimal within the class of contracts defined by $\mathcal{K}_0$ are: (i) the EWFB contract $\{T = \gamma; S = \sigma - \gamma\}$, which costs $(4 + 2k) \cdot c$; (ii) contracts of the form $\{T = T(\cdot); S = S(\cdot)\}$, where T and S are contingent on a single state variable, either γ or σ, which cost $(4 + k) \cdot c$; the contract $\{T = \bar{T}; S = \bar{S}\}$, which costs $4c$; the contract $\{T = \bar{T}\}$, which costs $2c$; and the empty contract. Our question now is simple. Can the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ be optimal for some c? Observing that the cost of the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is $3c$, the results of Proposition 2 permit a simple answer to this question: Yes, if and only if $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ outperforms the contract $\{T = \bar{T}; S = \bar{S}\}$ at the contracting cost level c_2 defined in Proposition 2, i.e., iff $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}} - \Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} > c_2$.

To understand this condition, recall from Proposition 2 that c_2 is defined as that level of contracting cost at which the contract $\{T = \bar{T}; S = \bar{S}\}$ yields the same level of expected net global welfare as either (i) the contract $\{T = \bar{T}\}$, when $c_2 < c_3$, or (ii) the empty contract, when $c_2 = c_3$. Consider first case (i), where $c_2 < c_3$. Since the cost of $\{T = \bar{T}; S = \bar{S}\}$ is $4c$ while the cost of $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is $3c$ and the cost of $\{T = \bar{T}\}$ is $2c$, the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ will never beat both $\{T = \bar{T}; S = \bar{S}\}$ and $\{T = \bar{T}\}$ if it does not do so at c_2, while if it does so at c_2 then it is the optimal contract in the extended set of contracts $\mathcal{K}_1$ for the contracting cost level c_2. Consider next case (ii), where $c_2 = c_3$. Since the cost of the empty contract is zero, the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ will never beat both $\{T = \bar{T}; S = \bar{S}\}$ and the empty contract if it does not do so at c_2, while if it does so at c_2 then it is the optimal contract in the extended set of contracts $\mathcal{K}_1$ for the contracting cost level c_2.

Collecting these points, we may thus state the following:

Proposition 3. *Consider the extended contract class $\mathcal{K}_1$. There exists $\hat{u} > 0$ (function of other parameters) such that, if $u < \hat{u}$, then:*
The contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is optimal for $c \in (\underline{c}_{NT}, \bar{c}_{NT})$, where $\underline{c}_{NT} \leq c_2 \leq \bar{c}_{NT}$. The interval $(\underline{c}_{NT}, \bar{c}_{NT})$ is nonempty (i.e. $\underline{c}_{NT} < \bar{c}_{NT}$) iff $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}} - \Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} > c_2$.

To further interpret the conditions under which an NT-based contract is optimal in the small-uncertainty case, notice that the expected gross global welfare under $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ must be lower than the expected gross global welfare under $\{T = \bar{T}; S = \bar{S}\}$, i.e., that $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}} < \Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}$, since the NT-based contract leaves t to discretion while no instruments are left to discretion in the contract $\{T = \bar{T}; S = \bar{S}\}$. Hence, if the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is to be preferred to the contract $\{T = \bar{T}; S = \bar{S}\}$, it must be because $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}}$ is sufficiently close to $\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}$ to warrant the savings in cost $(1c)$ that can be enjoyed in switching from $\{T = \bar{T}; S = \bar{S}\}$ to $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$. But notice that $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}}$ will be close to $\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}$ when *the cost of discretion over t is low*. On the other hand, if $c_2 < c_3$ and the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is to be preferred to the contract $\{T = \bar{T}\}$, it must be because $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}}$ is sufficiently far above $\Omega^{\max}_{\{T=\bar{T}\}}$ to warrant the extra cost $(1c)$ that must be payed in switching from $\{T = \bar{T}\}$ to $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$. But notice that $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}}$ will be far above $\Omega^{\max}_{\{T=\bar{T}\}}$ when *the cost of discretion over S is high relative to the cost of discretion over t.* Finally, if $c_2 = c_3$ and the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is to be preferred to the empty contract, it must be because $\Omega^{\max}_{\{NT,\tau=\bar{\tau},s=\bar{s}\}}$ is sufficiently far above $\Omega_{\{\emptyset\}}$to warrant the extra cost ($3c$) that must be payed in switching from the empty contract to $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$, which again is more likely when the cost of discretion over t is low. Hence, it is clear that a key determinant of the conditions under which an NT-based contract is optimal involve an evaluation of the costs of discretion over t and S and a comparison between them.

To evaluate the impact of changes in the underlying parameters α, β and λ on the relative performance of the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$, we proceed as before and consider the case of no uncertainty ($u = 0$), report our results in a series of Remarks (which we prove with Maple calculations available on request), and then develop an interpretation of our findings. Under the assumption that $u = 0$, the contract $\{T = \bar{T}; S = \bar{S}\}$ is the EWFB contract and is optimal for c in $(0, \underline{c}_{NT})$, and the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is optimal for c in $(\underline{c}_{NT}, \bar{c}_{NT})$. The remarks below describe how the NT-based contract fares relative to the EWFB contract as the underlying model parameters are altered:

Remark 3. *Assume* $u = 0$. *Then:*
$\underline{c}_{NT}/\bar{c}_{NT}$ *is weakly increasing in* β, *with* $\underline{c}_{NT}/\bar{c}_{NT} < 1/2$ *for* β *sufficiently small;*
$\underline{c}_{NT}/\bar{c}_{NT}$ *is weakly decreasing in* λ, *with* $\underline{c}_{NT}/\bar{c}_{NT} = 1$ *for* λ *sufficiently small.*

Remark 4. *Assume* $u = 0$. *Then:*
$\underline{c}_{NT}/\bar{c}_{NT}$ *is independent of* α*;*

According to Remark 3, the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is favored relative the contract $\{T = \bar{T}; S = \bar{S}\}$ as β falls, and is sure to be an optimal contract for a greater range of contracting costs than $\{T = \bar{T}; S = \bar{S}\}$ if β is sufficiently small. On the other hand, $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ is disfavored relative to $\{T = \bar{T}; S = \bar{S}\}$ as λ falls, and disappears completely as an optimal contract (for any contracting costs) if λ is sufficiently small. Finally, according to Remark 4, α has no bearing on the relative attractiveness of the contracts $\{T = \bar{T}; S = \bar{S}\}$ is favored relative the contract $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$.

To interpret the findings of Remarks 3 and 4, we consider the degree of substitutability between τ and t and the implications for the cost of discretion over t,. We can define an index of substitutability between τ and t in a similar way as we defined the index of substitutability between T and S in the previous section. Namely, we consider the ratio $\frac{d\Omega(\tau,t^R(\tau),s)/d\tau}{\partial\Omega(\tau,t^R(\tau),s)/\partial\tau}$ evaluated at the Nash equilibrium. This yields

$$\left(\frac{d\Omega(\tau, t^R(\tau), s)/d\tau}{\partial\Omega(\tau, t^R(\tau), s)/\partial\tau}\right)_{\tau=\tau^{NE}, s=s^{NE}} = \frac{\lambda A^2}{(\beta + \lambda)[A^2 - \beta(\beta + \lambda)]}$$

It can be shown that this index takes on a value of one when $\beta = 0$, a value of zero when $\lambda = 0$, and is monotonically decreasing in β and increasing in λ and independent of α. Accordingly, we may say that the degree of substitutability between τ and t is increasing in β, decreasing in λ, and is independent of α.[23]

Consider next the relevance of the degree of substitutability between τ and t for the cost of discretion over t. Clearly, if τ and t are highly substitutable, then any constraints placed on τ (and s) through an NT-based contract can be largely undone if t is left to discretion, just as with S and T and non-NT contracts. Notice, though, that the underlying parameter conditions that cause τ and t to be highly substitutable (high β and/or low λ) – and hence lead to a high cost of discretion over t through the instrument-substitutability effect – are *different* from those parameter conditions that cause S and T to be highly substitutable (low β and/or high λ).

Gathering these observations together, we may draw three conclusions. First, through the instrument-substitutability effect a low level of β tends to raise the cost of discretion over S and lower the cost of discretion over t, and hence through this effect tends to favor $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ and disfavor $\{T = \bar{T}; S = \bar{S}\}$ (i.e., reductions in β tend to reduce $\underline{c}_{NT}/\bar{c}_{NT}$). Second, through the instrument-substitutability effect a low level of λ tends to lower the cost of discretion over S and raise the cost of discretion over t, and hence through this effect tends to disfavor $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ and favor $\{T = \bar{T}; S = \bar{S}\}$ (i.e., reductions in λ tend to increase $\underline{c}_{NT}/\bar{c}_{NT}$). And finally, changes in α have no impact on the relative performance of the contracts $\{NT; \tau = \bar{\tau}; s = \bar{s}\}$ and $\{T = \bar{T}; S = \bar{S}\}$ through the instrument-substitutability effect (i.e., changes in α leave $\underline{c}_{NT}/\bar{c}_{NT}$ unchanged)

As is evident from this discussion, the degree of substitutability between policy instruments forms the basis of a useful interpretation of the findings in Remark 3 and 4. And as before, the export-supply elasticity effect only becomes relevant for understanding the changing nature of the optimal contracts as underlying parameters change once strictly positive uncertainty ($u > 0$) is permitted. Nevertheless, before turning to the large-uncertainty environment of the next section, we briefly describe here the nature of the export-supply elasticity effect in the presence of NT, since it will be useful for interpreting the large-uncertainty results. To understand, consider first the relevance of the magnitude of the Nash equilibrium foreign export supply elasticity for the cost of discretion over t. Using the expression given above for the best-response level of t, $t^R(\tau, s)$, and noting that $\tau^{NE} = \frac{p^*}{\eta^*}$, it can be seen that the distance between $t^R(\tau^{eff}, s^{eff})$ and t^{eff} is proportional to $\frac{p^*}{\eta^*}$, with the factor of proportionality independent of α. Hence, when α is high and the magnitude of the Nash equilibrium foreign export supply elasticity is low, it follows that $t^R(\tau^{eff}, s^{eff})$ is far from t^{eff} and the cost of discretion over t is high. We now turn to the large-uncertainty environment, and utilize the instrument-substitutability effect and export-supply elasticity effect to help interpret our findings.

Large uncertainty

We next proceed to explore the role of the NT clause in the case of large uncertainty. We maintain our focus on NT-based contracts that strictly dominate optimal non-NT contracts in $\mathcal{K}_0$. Our purpose here is twofold. First, we illustrate how the insights regarding the performance of NT-based contracts which are recorded in Remark 3 for the case of small uncertainty are borne out as well in our large-uncertainty setting, while the predictions of Remark 4 no longer

[23]Intuitively, as with the substitutability between T and S, the impact of α, β and λ on the degree of substitutabity between τ and t follows from inspection of the expressions for $p(\tau, t, s)$, $q(\tau, t, s)$ and $p^*(\tau, t, s)$.

apply. And second, we consider how the performance of NT-based contracts are affected by increasing levels of uncertainty.

Continuing with the two scenarios from our previous large-uncertainty analysis, we again focus initially on our original scenario, in which contracting takes place in the presence of uncertainty over the home-country demand intercept α and the home-country consumption externality parameter γ, and then describe briefly results under the alternative scenario in which the uncertain state variables are the home-country demand intercept α and the production externality parameter σ.

As our earlier discussion makes clear, whether uncertainty is small or large, the only NT-based contracts that can be strictly optimal are those that bind both s and τ but not t. In the case of small uncertainty, we were able to argue as well that the only such contract that is undominated is $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$, which costs $3c$. However, with large uncertainty we cannot rely on this argument. With our assumption of uncertainty over the two state variables α and γ, this implies that there are two additional NT-based contracts which we must consider, each of which costs $4c$: $\{NT, \tau = \bar{\tau}, s = f(\alpha)\}$ and $\{NT, \tau = \bar{\tau}, s = f(\gamma)\}$.[24] As it turns out, we are unable to find any values of the underlying parameters of our model for which either of the NT-based contracts costing $4c$ becomes the dominant $4c$ contract in the large-uncertainty environment.[25] And so, in line with our small-uncertainty findings, we find that the NT-based contract costing $3c$, $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$, is the only NT-based contract that is ever strictly optimal for a range of contracting costs c.

To explore the conditions that are most favorable to $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ as an optimal contract, we report in Figures 2a-2c results from three calculations: changes in the mean of the home-country demand intercept, $\bar{\alpha}$ (Figure 2a); changes in the slope of the home-country demand curve, $\bar{\beta}$ (Figure 2b); and changes in the level of uncertainty over the home-country consumption externality parameter, Δ_γ (Figure 2c).[26]

Figure 2a illustrates how changes in $\bar{\alpha}$ alter the usefulness of the NT-based contract. Specifically, through the export-supply-elasticity effect, higher values of $\bar{\alpha}$ induce greater unilateral incentives to distort unconstrained internal measures away from their Pigouvian first-best levels when tariffs are contractually constrained to efficient levels. This in turn disfavors the discretionary treatment of t, and hence disfavors the $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ contract, a result that does not arise in the case of zero uncertainty as Remark 4 indicates (while we do not report them, analogous results occur with regard to the demand uncertainty parameter Δ_α, and for analo-

[24] As can be seen from the NT pricing relationships in (5), of the remaining NT contracts costing $4c$ under our assumption of uncertainty over α and γ, $\{NT, \tau = f(\alpha), s = \bar{s}\}$ is identical to $\{NT, \tau = \bar{\tau}, s = f(\alpha)\}$, while $\{NT, \tau = f(\gamma), s = \bar{s}\}$ is identical to $\{NT, \tau = \bar{\tau}, s = f(\gamma)\}$.

[25] We suspect that this feature holds generally (for all parameter values consistent with the scenario under consideration), but we have not yet been able to prove it.

[26] Our parameter values were chosen to allow the strictly optimal NT contract $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ identified in Proposition 3 to arise for some range of contracting costs. To generate Figure 2a, the remaining parameters are set to the following values: $\Delta_\alpha = 3; \bar{\alpha}^* = 10; \bar{\beta} = 0.15; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 0; \bar{\gamma} = 10; \Delta_\gamma = 3$. To generate Figure 2b, the remaining parameters are set to the following values: $\bar{\alpha} = 20; \Delta_\alpha = 0.5; \bar{\alpha}^* = 10; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 0; \bar{\gamma} = 1; \Delta_\gamma = 1$.To generate Figure 2c, the remaining parameters are set to the following values: $\bar{\alpha} = 20; \Delta_\alpha = 0.5; \bar{\alpha}^* = 10; \bar{\beta} = 0.15; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 0; \bar{\gamma} = 10$.

gous reasons). Figure 2b confirms for the large-uncertainty case what Remark 3 establishes for the case of small uncertainty, namely that, through the instrument-substitutability effect, the NT-based contract $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ is most useful as a device for economizing on contracting costs when $\bar{\beta}$ is very low, so that the consumption tax (t) is ineffective at manipulating the terms of trade, thereby favoring the approach of the $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ contract, which leaves t to discretion. And finally Figure 2c shows that greater uncertainty over the home-country consumption externality parameter γ (higher Δ_γ) tends to favor the $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ contract because, in light of the fact that $t^{eff} = \gamma$, the discretionary treatment of t which permits state-contingency is then more valuable.

Turning briefly to our alternative scenario, in which contracting takes place in the presence of uncertainty over the home-country demand intercept α and the home-country production externality parameter σ, we use Figure 2d to illustrate the impact of changes in the level of uncertainty over the home-country production externality parameter σ.[27] Intuitively, in light of the fact that $s^{eff} = \sigma$ and that $s = \bar{s}$ in the $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ contract, it might be expected that increases in uncertainty over σ (higher Δ_σ) should disfavor the NT contract. As Figure 2d demonstrates, this is borne out, with the $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ contract disappearing as on optimal contract for σ sufficiently high.

Overall, then, we may conclude from this discussion that the conditions that favor the NT clause as a means of economizing on contracting costs are: (i) relatively high elementary contracting costs c (because NT contracts do not approximate very closely the EWFB contract in terms of expected gross global welfare, but offer relatively dramatic reductions in contracting complexity); (ii) relatively low $\bar{\alpha}$ and Δ_α and hence relatively low and stable demand (because this gives rise to low incentives for distorting the choice of internal measures for terms-of-trade gains, thereby favoring discretionary treatment of t); (iii) relatively low $\bar{\beta}$ and hence relatively inelastic demand (because the consumption tax (t) is then ineffective at manipulating the terms of trade, thereby favoring discretionary treatment of t); (iv) relatively high Δ_γ and hence uncertainty over the level of the consumption externality γ (because in light of the fact that $t^{eff} = \gamma$ this favors the discretionary treatment of t); and (v) relatively low Δ_σ and hence uncertainty over the level of the production externality σ (because in light of the fact that $s^{eff} = \sigma$ this does not penalize the rigid treatment of s).

5. The Role of Non-Violation Complaints

We now turn to an evaluation of the usefulness of the Non-Violation (NV) nullification-or-impairment provisions in GATT Article XXIII.1(b) as a means to economize on contracting costs. As we noted in the previous section, together with the NT clause, it is very likely that the role of the NV clause in the GATT/WTO may be best understood from an incomplete contracts perspective.

To develop this understanding, we begin by noting that tariff commitments in the GATT/WTO are interpreted as implying something beyond simple tariff obligations: in particular, they are

[27] To generate Figure 2d, the remaining parameters are set to the following values: $\bar{\alpha} = 20; \Delta_\alpha = 0.5; \bar{\alpha}^* = 10; \bar{\beta} = 1; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 4; \bar{\gamma} = 0$.

seen as guaranteeing a level of "market access." Market access, in turn, is interpreted as reflecting the "conditions of competition" between domestic and foreign producers, something which is clearly related to but not synonymous with import volume. Evidently, the market access guarantee that accompanies a tariff binding in the GATT/WTO is a subtle concept: it is a promise not to alter future government policies in a way that would upset the conditions of competition that are established at the time of the original negotiations, but it is not a guarantee against changes in market conditions and hence not a guarantee of trade volumes. In the GATT/WTO, it is the NV clause that serves the role of protecting the market access expectations of governments against changes in policies by their trading partners – even when these policies are not contracted over – which would have the effect of upsetting the market access that a government could have reasonably expected based on a prior GATT/WTO negotiation.

We model the NV clause in a very rudimentary way. In particular, we suppose that the NV clause imposes a constraint on the home government of the form: "Subsequent to negotiations in which the home government binds its tariff, the home government must change its tariff to maintain the implied level of market access if it changes its domestic policies." Formally, arguing along the lines of Bagwell and Staiger (2001), it can be shown in the present setting that this kind of constraint can be captured by the following restriction, which applies only if the home tariff is contractually constrained (we let $\tau(\cdot)$ denote the possibly-state-contingent contracted constraint on the home tariff):

$$p^*(\cdot) = p^*(\tau = \tau(\cdot), t_f^0, s^0, t_h^0; \boldsymbol{\theta}).$$

The left-hand side of this restriction denotes the market-clearing terms of trade evaluated at the ex-post chosen policies and at the realized state $\boldsymbol{\theta}$; more concretely, this is the value of p^* that is observed ex-post. The right-hand side of this restriction denotes the market-clearing terms of trade evaluated at the contracted tariff and the home internal policies in place at the time of the original negotiation, given the realized state $\boldsymbol{\theta}$; this is the value of p^* that should be observed ex-post provided that the expected level of market access is delivered. The NV clause requires that the observed value of p^* (the left-hand side) be equal to the value of p^* that should be observed ex-post provided that the expected level of market access is delivered (the right-hand side).

When coupled with a tariff binding, the constraint imposed by the NV clause serves to grant the home government a degree of unilateral discretion over its internal policies, but only up to a point. Specifically, the exercise of this discretion cannot undermine the implied level of market access that the tariff binding has granted to the foreign government. It is direct to establish within our model that this limited degree of discretion has a very attractive feature: it guarantees that the home government will always set its internal policies equal to the Pigouvian levels, and then adjust τ to satisfy the NV requirement.[28] Intuitively, as we have argued previously, the only reason that the home government might distort its internal measures away from their Pigouvian levels is for the purpose of manipulating the terms of trade

[28] More accurately, what is guaranteed under the NV clause is that $S = \sigma - \gamma$, with T then adjusted to meet the NV requirement. However, recalling that $T \equiv \tau + t_f$ and $S \equiv s - t_h$, it can be seen than this is implied if the home government sets its internal policies equal to the Pigouvian levels, so that $t_f = \gamma, s = \sigma$, and $t_h = \gamma$, and then adjusts τ to satisfy the NV requirement.

(p^*), when it is constrained from using its tariff for this purpose. But as the NV clause prevents the home government's unilateral choice of policies from having any effect on p^*, this distorting motive is eliminated. From this discussion, we may therefore conclude that any contract that includes the NV clause (which we denote NV) and places a constraint on the tariff will deliver Pigouvian levels of the home government's internal measures (that is, $t_f = \gamma, s = \sigma, t_h = \gamma$).

To explore the potential value of the NV clause, we first consider whether it might be useful as part of a first-best contract. Since any contract that includes the NV clause and places a constraint on the tariff will deliver Pigouvian levels of the home government's internal measures, the first best is assured provided only that the contracted tariff level $\tau(\cdot)$ is determined so that the terms of trade p^* (and therefore the market-clearing trade volume) is maintained at its efficient level for each realized state $\boldsymbol{\theta}$, or $p^*(\tau = \tau(\cdot), t_f^0, s^0, t_h^0; \boldsymbol{\theta}) = p^*(\tau = 0, t_f = \gamma, s = \sigma, t_h = \gamma; \boldsymbol{\theta})$. Using the expression for p^*, it is direct to show that the efficient level of gross global welfare can be achieved under the NV clause when the contracted tariff level is made contingent on the externalities and the other state variables. That is, a contract of the form $\{NV; \tau = \tau(\theta)\}$ can achieve the first-best expected gross global welfare, i.e., $\Omega^{\max}_{\{NV;\tau=\tau(\theta)\}} = \Omega_{\{EWFB\}}$.[29] The immediate question is, Can a contract of the form $\{NV; \tau = \tau(\theta)\}$ achieve the first-best expected gross global welfare at a cost less than $4c_p + 2c_s$, the cost of the contract $\{T = \gamma; S = \sigma - \gamma\}$ which is the EWFB contract in $\mathcal{K}_0$?

At a general level, comparing the contracting costs of the alternative first-best contracts $\{NV; \tau = \tau(\theta)\}$ and $\{T = \gamma; S = \sigma - \gamma\}$ presents some clear trade-offs.[30] On the one hand, by utilizing the NV clause and not specifying any internal policy instruments directly, a contract of the form $\{NV; \tau = \tau(\theta)\}$ economizes on contracting costs associated with *policies* relative to the contract $\{T = \gamma; S = \sigma - \gamma\}$.[31] On the other hand, enforcement of the NV clause requires verification of the state of demand and supply (which is utilized to run the appropriate "but for" counter-factual), and so the contract $\{T = \gamma; S = \sigma - \gamma\}$ economizes on contracting costs associated with *states* relative to a contract of the form $\{NV; \tau = \tau(\theta)\}$. Hence, at a general level a contract of the form $\{NV; \tau = \tau(\theta)\}$ will be a relatively low-cost way of delivering the first best when k, the cost of including states relative to the cost of including policies in the contract, is low.

To quantify these trade-offs, we suppose that a contract of the form $\{NV; \tau = \tau(\theta)\}$ costs $c_p + 8c_s$: that is, it costs c_p for the tariff and an additional $8c_s$ for the state variables.[32] Recalling

[29] When we write $\tau = \tau(\theta)$ we mean that the tariff is contingent on all 8 state variables.

[30] We observe that the basic tradeoffs we describe here would be present in a comparison between any outcome-based contracts (of which the NV contract is a particular example) and instrument-based contracts. Hence, while our results in this section are derived within the context of an outcome-based contract which is particularly relevant for the GATT/WTO (the NV contract), they should apply more generally to other outcome-based contracts that might be of interest as well.

[31] A similar idea is expressed by Sykes (2003, p. 18) in the context of comparing the NV clause as a method to discipline domestic subsidies in relation to the more direct approach taken under the WTO Agreement on Subsidies and Countervailing Measures: "A nice feature of the nonviolation doctrine is the fact that it does not require subsidies to be carefully defined or measured. A complaining member need simply demonstrate that an unanticipated government program has improved the competitive position of domestic firms at the expense of their foreign competition. The administration of the doctrine is thus reasonably straightforward, and the fighting issue is likely to be whether the government policy in question was foreseen by trade negotiatiors."

[32] Note that enforcing the NV clause does not require the court to verify the level of the tariff τ, because the

our parametrization of contracting costs ($c_p = c$, $c_s = k \cdot c$), we may then state the condition under which a contract of the form $\{NV; \tau = \tau(\theta)\}$ (costing $(1+8k)\cdot c$) represents a less-costly way of delivering the first best than the contract $\{T = \gamma; S = \sigma - \gamma\}$ (costing $(4+2k)\cdot c$), namely $k < 1/2$.

To focus on the role of the NV clause in incomplete contracts, we consider a broader set of contracts that include contracts in $\mathcal{K}_0$ plus contracts which include the NV clause, which we henceforth refer to as "NV-based" contracts. We refer to this extended set of contracts as $\mathcal{K}_2$. Notice that, if the contract is based on a tariff binding and an NV clause, there is no role for any additional constraint on internal policies, since as we have observed internal policies will be set at their Pigouvian levels. For this reason, the class of contracts $\mathcal{K}_2$ includes the class $\mathcal{K}_0$, plus the contract $\{NV; \tau = \tau(\theta)\}$ (which costs $(1+8k)\cdot c$) and two kinds of remaining NV-based contracts that warrant consideration: (i) contracts of the form $\{NV; \tau = \tau_1(\cdot)\}$, where τ is contingent on a single externality state variable, γ or σ, and up to 6 non-externality state variables, which cost $(1+7k)\cdot c$; and contracts of the form $\{NV; \tau = \tau_0(\cdot)\}$, where τ is not contingent on any externality state variable, but is contingent on up to 6 non-externality state variables, which cost $(1+6k)\cdot c$.

Next we turn to a formal evaluation of the NV clause. Again, we first consider the small-uncertainty case, and then report our findings for a large-uncertainty environment.

Small uncertainty

As we have explained above, with $k \geq 1/2$ we may restrict our focus with regard to the NV clause to NV-based contracts of the form $\{NV; \tau = \tau_1(\cdot)\}$ and $\{NV; \tau = \tau_0(\cdot)\}$, which cost $(1+7k)\cdot c$ and $(1+6k)\cdot c$, respectively. In terms of expected gross global welfare, these two kinds of NV-based contracts present an interesting trade-off: they cannot achieve the first best as long as there is any uncertainty over the realized values of σ and γ, because they do not permit the needed sensitivity of p^* to σ and γ; but they get the internal policies to the Pigouvian levels.

Consider, then, the value of these NV-based contracts relative to two contracts that, according to Proposition 2, are optimal in the set $\mathcal{K}_0$ over some range of contracting costs in a small-uncertainty setting, namely, contracts of the form $\{T = T(\cdot); S = S(\cdot)\}$, where T and S are contingent on a single state variable, either γ or σ, and the contract $\{T = \bar{T}; S = \bar{S}\}$. Recalling that these contracts cost $(4+k)\cdot c$ and $4\cdot c$, respectively, it may now be seen that the contract $\{NV; \tau = \tau_1(\cdot)\}$ dominates $\{T = T(\cdot); S = S(\cdot)\}$ and the contract $\{NV; \tau = \tau_0(\cdot)\}$ dominates $\{T = \bar{T}; S = \bar{S}\}$ for $k \geq 1/2$ but sufficiently small. This follows because: (i) for $k = 1/2$, the total contracting costs of each contract are the same within each pair of contracts being compared; and (ii) within each pair of contracts being compared, the expected gross global welfare of the NV-based contract is strictly higher (i.e., $\Omega^{\max}_{\{NV;\tau=\tau_1(\cdot)\}} > \Omega^{\max}_{\{T=T(\cdot);S=S(\cdot)\}}$ and $\Omega^{\max}_{\{NV;\tau=\tau_0(\cdot)\}} > \Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}$), because the NV-based contract permits (at least) the same degree of state-sensitivity of p^* but also gets the internal policies to the Pigouvian levels.

Recalling from Proposition 2 that the optimal contract in $\mathcal{K}_0$ for $c \in (c_0, c_1)$ is of the form $\{T = T(\cdot); S = S(\cdot)\}$, while the optimal contract in $\mathcal{K}_0$ for $c \in (c_1, c_2)$ is of the form

level specified in the contract is only a reference value. Thus, the cost of enforcing the NV clause might be somewhat less than $c_p + 6c_s$. Our qualitative results would not change if we took this formally into account.

$\{T = \bar{T}; S = \bar{S}\}$, we may now state:

Proposition 4. *Consider the extended contract class $\mathcal{K}_2$. There exists $\hat{u} > 0$ (function of other parameters) such that, if $u < \hat{u}$, then:*
(i) If k is lower than a critical level $\hat{k} > 1/2$, the optimal contract is of the NV type for $c \in (\underline{c}^{NV}, \bar{c}^{NV})$, where $\underline{c}^{NV} < c_0 < \bar{c}^{NV}$.
(ii) If $k < 1/2$, the optimal contract is of the NV type for $c \in (0, \bar{c}^{NV})$.
(iii) If k is sufficiently close to zero, the optimal contract is either an NV-type contract or the empty contract.

We observe that the condition identified in the proposition under which the NV-based contracts will be optimal for some c reflects a basic feature of the nature of (possible) contract-cost saving that the NV clause affords. In effect, this clause removes the burden of specifying a detailed list of policies (see Sykes' (2003) discussion in note 12) but puts in its place the burden of verifying the appropriate state contingencies necessary to perform the "but for" counterfactual calculations. This economizes on contracting costs when the cost of specifying states is relatively low compared to the cost of specifying policies. In our formal model, there are a fixed number of policy instruments (4) and states relevant for the counterfactual calculations (6), and the relative cost of specifying states versus policies is then captured by the parameter k. However, it is easy to see that the cost of specifying states relative to policies would also be affected by the relative numbers of policy instruments and states, and would in particular fall for fixed k as the number of potential policy instruments grows relative to the number of relevant states. In this sense, Proposition 4 suggests that NV-based contracts might be most useful in settings (e.g., agriculture?) where the policy environment is especially detailed and complex relative to the economic environment.

To further interpret the conditions under which an NV-based contract is optimal in the small-uncertainty case, we next consider the determinants of the value of the partial discretion permitted by the NV clause. As we have observed above, this discretion ensures that the home government sets its internal measures at their Pigouvian levels (that is, $t_f = \gamma, s = \sigma, t_h = \gamma$), and then adjusts τ to satisfy the NV requirement, which implies equivalently that $S = \sigma - \gamma$ under the NV clause, with T then adjusted to meet the NV requirement. But the efficiency gain in being able to adjust S and T in this way depends on the magnitude of the home demand and supply slope parameters β and λ. To see this, note that the expressions for $p(T, S)$, $q(T, S)$ and $p^*(T, S)$ can be used to calculate expressions for dp/dS and dq/dS along an $iso - p^*$ locus (i.e., with T adjusted hold p^* fixed):

$$dp(T(S), S)/dS|_{dp^*=0} = -\lambda \cdot [(\beta^* + \lambda^*)/(\beta + \lambda) + 1]/A, \text{ and}$$
$$dq(T(S), S)/dS|_{dp^*=0} = [(\beta + \beta^* + \lambda^*) - \lambda(\beta^* + \lambda^*)/(\beta + \lambda)]/A.$$

But it may now be seen that, for fixed β, the efficiency gain that is associated with being able to adjust S and T along an $iso - p^*$ locus is higher the larger is the magnitude of $dp(T(S), S)/dS|_{dp^*=0}$, and therefore the larger is λ. Similarly, it may be seen that, for fixed λ, the efficiency gain that is associated with being able to adjust S and T along an $iso - p^*$ locus is higher the larger is the magnitude of $dq(T(S), S)/dS|_{dp^*=0}$, and therefore the larger is

β. As a result of these observations, we may expect that NV-based contracts will be favored in settings where β and/or λ are high, because the partial discretion granted under the NV clause is then most valuable. We will confirm this prediction in the large-uncertainty environment to which we now turn.

Large uncertainty

We now explore the role of the NV clause in the case of large uncertainty. Referring to the two scenarios from our previous large-uncertainty analysis, we focus this time on our alternative scenario, in which contracting takes place in the presence of uncertainty over the home-country demand intercept α and the home-country production externality parameter σ, and simply note that analogous results arise when it is the consumption externality parameter γ rather than the production externality σ that is uncertain.

In this setting, the NV-based contract $\{NV; \tau = \tau(\sigma)\}$ achieves the first best at a cost of $(1 + 2k) \cdot c$, while the contract $\{T = \bar{\gamma}; S = \sigma - \bar{\gamma}\}$ achieves the first best at a cost of $(4 + k) \cdot c$. Hence, in order for the contract $\{T = \bar{\gamma}; S = \sigma - \bar{\gamma}\}$ to remain the EWFB contract in the extended set of contracts $\mathcal{K}_2$, we require that $k \geq 3$. The remaining NV-based contracts of interest in this setting are then $\{NV; \tau = \tau(\alpha)\}$ and $\{NV; \tau = \bar{\tau}\}$, each of which costs $(1 + k) \cdot c$. It is direct to show in this setting, however, that there is no gain from allowing τ to be contingent on α, and so we may focus on the contract $\{NV; \tau = \bar{\tau}\}$.

For purposes of illustration we set $k = 4$, which implies that the NV contract $\{NV; \tau = \bar{\tau}\}$ costs $5c$. The important point is that this is more than the cost of the contract $\{T = \bar{T}; S = \bar{S}\}$ ($4c$), which then ensures that $\{T = \bar{T}; S = \bar{S}\}$ is not dominated by $\{NV; \tau = \bar{\tau}\}$ at all contracting costs, but less than the cost of the EWFB contract $\{T = \bar{\gamma}; S = \sigma - \bar{\gamma}\}$ ($8c$), which then ensures that $\{NV; \tau = \bar{\tau}\}$ is not dominated by the EWFB contract at all contacting costs. Finally, in order to contrast the conditions which favor NV-based contracts with the conditions favoring NT-based contracts, we calculate the optimal contracts over the extended set of contracts $\mathcal{K}_3 \equiv \mathcal{K}_1 \cup \mathcal{K}_2$, i.e., the class of contracts that include $\mathcal{K}_0$ plus NT-based contracts plus NV-based contracts.

In Figures 3a-3b, we explore the conditions that favor the NV contract.[33] Figure 3a depicts the impact of the demand slope parameter $\bar{\beta}$. Recall that low levels of $\bar{\beta}$ tend to favor the NT contract, because the consumption tax (t) is then ineffective at manipulating the terms of trade, thereby favoring the approach of the NT contract, which leaves t to discretion. This is reflected in Figure 3a. But the impact of low levels of $\bar{\beta}$ on the NV contract is just the opposite: as Figure 3a shows, the NV contract is favored by higher levels of $\bar{\beta}$. This reflects the point that, as we explained above, when β (and/or λ) is high, NV-based contracts will be favored because the partial discretion granted under the NV clause is then most valuable (while not reported, a similar result obtains with regard to λ). And in effect, when $\bar{\beta}$ is low, the extra flexibility afforded by the $\{NV; \tau = \bar{\tau}\}$ contract over the $\{T = \bar{T}; S = \bar{S}\}$ contract is more apparent than real, and the lower contracting costs associated with the latter contract ensures that it dominates the former contract.

[33] Our parameter values were chosen to allow the NV-based contract $\{NV, \tau = \bar{\tau}\}$ to arise for some range of contracting costs. To generate Figure 3a, the remaining parameters are set to the following values: $\bar{\alpha} = 20; \Delta_\alpha = 0.5; \bar{\alpha}^* = 10; \bar{\beta} = 1; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 1; \Delta_\sigma = 1; \bar{\gamma} = 0$. To generate Figure 3b, the remaining parameters are set to the following values: $\bar{\alpha} = 18; \Delta_\alpha = 3; \bar{\alpha}^* = 10; \bar{\beta} = 1; \bar{\beta}^* = 1; \bar{\lambda} = 1; \bar{\lambda}^* = 1; \bar{\sigma} = 9; \bar{\gamma} = 0$.

Figure 3b illustrates the impact of increases in the level of uncertainty over the production externality σ (higher Δ_σ) on the performance of the NV contract. As Figure 3b shows, the NV contract is dominated by the $\{T = \bar{T}; S = \bar{S}\}$ contract for very low levels of Δ_σ, reflecting the fact that when Δ_σ is very low there is little to gain from paying the extra contracting cost of the NV contract in order to facilitate state-contingent policies (and a q) that can vary with σ. On the other hand, when Δ_σ reaches a high enough level the NV contract is squeezed out because it is then better to economize on contracting costs once c reaches a critical level by moving from the EWFB contract directly to $\{T = \bar{T}\}$: the reason is that for high Δ_σ, the fixed p^* demanded by the NV contract becomes too costly, and the adjustments to p^* that occur under the added flexibility of the $\{T = \bar{T}\}$ contract become desirable.

Overall, then, we may conclude from this discussion that the conditions that favor the usefulness of the NV clause as a means of economizing on contracting costs are: (i) relatively low elementary contracting costs c (because the NV contract approximates very closely the EWFB contract in terms of expected gross global welfare, but does not offer very dramatic reductions in contracting complexity); (ii) relatively high $\bar{\beta}$ and/or $\bar{\lambda}$ and hence relatively elastic demand and/or supply (because the freedom that the NV clause provides to make policy changes that fix p^* but alter q and/or p is then more valuable); and (iii) a moderate level of uncertainty over the level of the production externality σ and the consumption externality γ (because this favors the moderate level of price flexibility – over p and q but not p^* – that the NV contract affords). These conditions are in addition to the one identified in Proposition 4, namely, that the relative cost of specifying states versus policies (as captured by the parameter k) be low.

Finally, it is interesting to observe that, while the GATT/WTO features both an NT clause and an NV clause, a strict interpretation of our findings would suggest that theses two clauses should never be found in the same contract. However, our finding that the NT clause works best when demand is relatively inelastic, while the NV clause works best when demand is relatively elastic, offers the possibility of a resolution of this puzzle in a broader model. In particular, if (i) a substantial portion of the costs associated with each clause is not in the writing of the clause but in its enforcement (verification costs), (ii) clauses are only enforced by the GATT/WTO when there is a claimant asserting that its rights have been nullified or impaired, and (iii) at least a portion of the burden of proof that a nullification or impairment has occurred falls on the claimant, then it could be optimal to write into the contract both an NT clause and an NV clause if the key elasticities were unknown at the time of contracting, so that ex post the most effective clause could be utilized by claimants given the elasticity conditions that prevail at that time.

6. The role of weak bindings

Thus far we have focused on contracts that impose equality constraints, such as $S = \bar{S}$. This of course would be without loss of generality in a world of costless contracting, for in this case the optimal contract would implement the first best outcome (e.g. $S = \sigma - \gamma$, $T = \gamma$), and hence there would be nothing to gain from using inequality constraints. In the presence of contracting costs, however, the optimal contract may not implement the first-best outcome, and intuition suggests that the next best solution may be to impose a ceiling of the kind $S \leq \bar{S}$, rather than imposing a rigid equality constraint ($S = \bar{S}$) or leaving the subsidy to discretion. This intuition can be conveyed with a simple example.

Suppose the only uncertain parameter is the production externality σ, which can take two values, σ_0 and σ_1, and set $\gamma \equiv 0$. Moreover, assume that c_s is very high, so that writing contingent contracts is too costly, and that $c_p = 0$, so that it is costless to write rigid contracts. Now suppose that $\Pr(\sigma_0)$ is small, so that, if one ignores the possibility of inequality constraints, the optimal contract is of the form ($T = \bar{T}, S = \bar{S}$), where $\bar{T}$ is close to zero and $\bar{S}$ is close to σ_1. Consider now an alternative contract that imposes only a ceiling on S: ($T = \bar{T}, S \leq \bar{S}$). Clearly, this contract is at least as good as the previous one, because given the constraint on T, the government's incentive to distort S is upwards: the government is tempted to "oversubsidize" production. Is this contract *strictly* preferable to the previous one? This depends crucially on the subsidy reaction function $S^R(T;\sigma)$ and on the support of σ. If $S^R(0;\sigma_0)$ is below σ_1, then in state σ_0 the contract ($T = \bar{T}, S \leq \bar{S}$) induces a more efficient outcome than the contract ($T = \bar{T}, S = \bar{S}$). Intuitively, if the government is free to go below the ceiling level $\bar{S}$, it will choose to do so if the externality σ is low enough, and this is beneficial for global welfare, even though the government will not decrease S all the way to the Pigouvian level.

The above reasoning suggests that inequality constraints should dominate equality constraints when the support of the state vector is large, so that imposing an equality constraint such as $S = \bar{S}$ creates opposite incentives to deviate in different states: an incentive to go above $\bar{S}$ in some states and an incentive to go below $\bar{S}$ in other states. The next proposition confirms and extend this intuition. We refer to a constraint of the kind $S \leq \bar{S}$ as a "rigid weak binding" and to a constraint of the kind $S = \bar{S}$ as a "rigid strong binding".

Proposition 5. *Rigid weak bindings are weakly preferable to rigid strong bindings. The preference is strict if the support of the state vector, Θ, is sufficiently large.*

It is interesting to note that a rigid weak binding combines rigidity and discretion, since the ceiling does not depend on the state of the world and a government has discretion to set the policy below the ceiling. Thus, our result highlights that *rigidity and discretion may be complementary ways to economize on contracting costs.*

Based on this result, one could characterize the optimal contract within the larger class of contracts that allows for weak bindings. It is not hard to show that all the qualitative results derived in the previous sections would continue to hold, with the only amendment that strong bindings are replaced by weak bindings.

In reality, it is reasonable to believe that the set of potential states of the world that are relevant for a trade agreement (what we call Θ in our model) is very large, and in this way our model predicts that the constraints imposed by trade agreements should predominantly take the form of weak bindings. This prediction is quite consistent with the observed nature of the GATT/WTO contract, where policy commitments are essentially all in the form of weak bindings.[34]

[34] We note here that this is not the only possible explanation for the use of weak bindings. Maggi and Rodriguez-Clare (2005) propose an alternative explanation based on political-economy considerations: their basic idea is that weak bindings allow governments to extract rents from lobbies after the agreement is signed,

7. Conclusion

The analysis above has highlighted how the interaction between contracting costs and uncertainty may help explain a number of salient features of the GATT. Our modeling approach has been extremely simple. We are confident that certain aspects of the model could be generalized without changing the picture qualitatively, although possibly at the cost of having to rely on numerical solutions. But there are several questions that are wide open and we would like to address in future work.

A first question concerns the modeling of contracting costs. This is clearly a highly complex issue. But it is easy to point to several aspects that our highly stylized framework does not capture. One such aspect is that one might expect the number of states and the number of policies that is contracted to interact in the determination of the complexity costs. One may also want to distinguish more clearly between the costs of "writing" contracts (including the cost of preparing and evaluating bargaining proposals), and the cost of implementing contracts. For instance, practitioners often allege that a main virtue of the MFN clause is that it dramatically reduces the complexity of the negotiations. Our present framework would not be able to capture such an implication of MFN.

A second limitation of the model is that it is based on social welfare maximization. It seems likely that the introduction of political considerations may give a larger role for provisions such as safeguard measures than in our welfare-maximizing setup. In some sense our model yields a form of safeguard measure, for the parameter region where the optimal contract includes a state-contingent total tax $T(\theta_k)$. But this agreement does not include an NT provision, and it is not based on some form of economic outcome, such as injury, like real world safeguards. The development of the model to include political economy considerations is also left for future work.

A third aspect that is largely missing from the analysis above is that of time. One means of saving on state verification costs would be to renegotiate the agreement very frequently. This would ensure that the agreement is well adapted to the existing economic and political situation without having to be highly state contingent. On the other hand, trade negotiations are likely to be associated with significant fixed costs, and frequent renegotiations would therefore give rise to substantial costs of this type. One should therefore expect that the duration of the agreement is partly determined by the trade-off between these different aspects of contracting costs.

since they allow a government to credibly threaten its domestic lobbies to lower the level of trade protection below the ceiling. We also note that the explanation proposed here is somewhat related to the one proposed in Bagwell and Staiger (2005), where weak bindings may be preferred to strong bindings in the presence of political-economy shocks that are privately observed by governments.

8. Appendix

Proof of Proposition 2: We start with the benchmark case of no uncertainty ($u = 0$). In this case there are four types of contract to consider: (i) $\{T = \bar{T}; S = \bar{S}\}$, which costs $4c$; (ii) $\{S = \bar{S}\}$, which costs $2c$; (iii) $\{T = \bar{T}\}$, which costs $2c$; (iv) the empty contract, $\{\emptyset\}$, which is costless.

The best contract of type (i) is clearly the EWFB contract, $\{T = \bar{\gamma}; S = \bar{\sigma} - \bar{\gamma}\}$.

Contracts of type (ii) cannot achieve any improvement over the Nash equilibrium, as Lemma 1 establishes.

What is the optimal level of $\bar{T}$ for a contract of type (iii)? This is the value of T for which the subsidy reaction function, $S^R(T)$, is tangent to an iso-Ω curve in (T, S) space. This tangency condition is given by

$$\frac{W_T + W_T^*}{W_S + W_S^*} = \frac{\bar{\beta}^* + \bar{\lambda}^*}{\bar{\beta} + \bar{\beta}^* + \bar{\lambda}^*}.$$

Rather than solving explicitly for the optimal $\bar{T}$, we just notice two properties of this contract. First, it cannot implement the first best, because the only points it can implement are those along the $S^R(T)$ curve, and this curve does not go through the first best point ($T = \bar{\gamma}; S = \bar{\sigma} - \bar{\gamma}$). Second, this contract sets $\bar{T} < T^{NE}$, and it strictly improves on the Nash equilibrium outcome. To see this, note that at the Nash equilibrium the LHS of the tangency condition is given by $\frac{W_T^*}{W_S^*} = \frac{\bar{\beta}+\bar{\lambda}}{\bar{\lambda}} > 1$ while the RHS is $\frac{dS^R(T)}{dT} = \frac{\bar{\beta}^*+\bar{\lambda}^*}{\bar{\beta}+\bar{\beta}^*+\bar{\lambda}^*} < 1$. Hence, at the Nash point the $S^R(T)$ curve is not tangent to the iso-Ω curve, and there is room for improvement over the Nash outcome when $\bar{T}$ is lowered below T^{NE}.

The empty contract $\{\emptyset\}$ of course yields the Nash surplus.

So far we have established a ranking of the three candidate contracts in terms of gross-of-contracting-costs surplus:

$$\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} > \Omega^{\max}_{\{T=\bar{T}\}} > \Omega_{\{\emptyset\}}.$$

Let $\mathcal{C}_{\{K\}}$ denote the interval of c (possibly empty) for which a contract of type K is optimal. Recalling that the costs of the contracts $\{T = \bar{T}; S = \bar{S}\}$, $\{T = \bar{T}\}$ and $\{\emptyset\}$ are respectively $4c$, $2c$ and 0, the above ranking immediately implies

$$\mathcal{C}_{\{T=\bar{T};S=\bar{S}\}} \preceq \mathcal{C}_{\{T=\bar{T}\}} \preceq \mathcal{C}_{\{\emptyset\}}$$

where $\mathcal{C}' \preceq \mathcal{C}''$ means that each element of $\mathcal{C}'$ is weakly lower than each element of $\mathcal{C}''$. Next we ask whether each of these intervals can be empty. Clearly, if c is sufficiently small then $\{T = \bar{T}; S = \bar{S}\}$ is optimal, hence $\mathcal{C}_{\{T=\bar{T};S=\bar{S}\}}$ is nonempty, and if c is sufficiently high then $\{\emptyset\}$ is optimal, hence $\mathcal{C}_{\{\emptyset\}}$ is nonempty. It remains to check whether $\mathcal{C}_{\{T=\bar{T}\}}$ is nonempty. Consider the level of c for which the contracts $\{T = \bar{T}; S = \bar{S}\}$ and $\{\emptyset\}$ yield the same net surplus. This is given by

$$c_a = (\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} - \Omega_{\{\emptyset\}})/4$$

Clearly, $\mathcal{C}_{\{T=\bar{T}\}}$ is empty if and only if $\{T = \bar{T}\}$ yields a lower net surplus than $\{\emptyset\}$ for $c = c_a$. This gives the condition

$$\Omega^{\max}_{\{T=\bar{T}\}} > \frac{1}{2}\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} + \frac{1}{2}\Omega_{\{\emptyset\}}$$

Let us now introduce a small amount of uncertainty. Let $u \equiv E(||\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}||)$, and suppose u is small but strictly positive. Now the unique EWFB is given by $\{T = \gamma; S = \sigma - \gamma\}$, which costs $(4 + 2k)c$. In addition to the EWFB, there is another type of contract that may be optimal in the presence of a small uncertainty, namely $\{T = T(\theta_k); S = S(\theta_k)\}$, where T and S are contingent on a single state variable θ_k; this costs $(4 + k)c$.

Are there other contracts that we should consider? The answer is no. In particular, contracts of the type $\{T = T(\cdot)\}$ or $\{S = S(\cdot)\}$ cannot be optimal if u is small. Consider for example $\{T = T(\cdot)\}$. Fix a level of c strictly inside the interval $\mathcal{C}_{\{T=\bar{T};S=\bar{S}\}}$, and call it $\hat{c}$. For $c \geq \hat{c}$, $\{T = T(\cdot)\}$ is dominated by $\{T = \bar{T}\}$, because the cost of adding contingencies exceeds the benefit (since u is small). For $c < \hat{c}$, $\{T = T(\cdot)\}$ is dominated by $\{T = \bar{T}; S = \bar{S}\}$. This is because for $c < \hat{c}$ the contract $\{T = \bar{T}; S = \bar{S}\}$ dominates the contract $\{T = \bar{T}\}$, and hence if u is small enough it will also dominate $\{T = T(\cdot)\}$.

For the five candidate contracts we have the following relationship in terms of gross surplus:

$$\Omega_{FB} > \Omega^{\max}_{\{T=T(\theta_k);S=S(\theta_k)\}} > \Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} > \Omega^{\max}_{\{T=\bar{T}\}} > \Omega_{\{\emptyset\}}$$

or, using the notation $\{T = T(\cdot); S = S(\cdot)\}$ to encompass contracts that are contingent on one or two state variables (as in the statement of the proposition), we have

$$\Omega^{\max}_{\{T=T(\cdot);S=S(\cdot)\}} > \Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} > \Omega^{\max}_{\{T=\bar{T}\}} > \Omega_{\{\emptyset\}}$$

This implies

$$\mathcal{C}_{\{T=T(\cdot);S=S(\cdot)\}} \preceq \mathcal{C}_{\{T=\bar{T};S=\bar{S}\}} \preceq \mathcal{C}_{\{T=\bar{T}\}} \preceq \mathcal{C}_{\{\emptyset\}}$$

Using similar arguments as the ones made above, one can show that $\mathcal{C}_{\{T=T(\cdot);S=S(\cdot)\}}$ and $\mathcal{C}_{\{\emptyset\}}$ are nonempty, and that $\mathcal{C}_{\{T=\bar{T}\}}$ is nonempty if and only if $\Omega^{\max}_{\{T=\bar{T}\}} > \frac{1}{2}\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}} + \frac{1}{2}\Omega_{\{\emptyset\}}$. It remains to argue that $\mathcal{C}_{\{T=\bar{T};S=\bar{S}\}}$ is nonempty.

Suppose by contradiction that $\mathcal{C}_{\{T=\bar{T};S=\bar{S}\}}$ is empty, and let c_1 denote the cost at which $\mathcal{C}_{\{T=T(\cdot);S=S(\cdot)\}}$ meets $\mathcal{C}_{\{T=\bar{T}\}}$. Then for $c = c_1$ the $\{T = \bar{T}\}$ contract yields the same net surplus as either the $EWFB$ contract or the $\{T = T(\theta_k); S = S(\theta_k)\}$ contract. Suppose it is the $EWFB$ contract; the argument is analogous if it is the other contract. Then c_1 is given by

$$\Omega^{\max}_{\{T=\bar{T}\}} - 2c_1 = \Omega_{FB} - (4 + 2k)c_1$$

that is

$$c_1 = \frac{\Omega_{FB} - \Omega^{\max}_{\{T=\bar{T}\}}}{2(1 + k)}$$

Notice that c_1 does not become small if uncertainty u approaches zero. Now we argue that for $c = c_1$ the contract $\{T = \bar{T}; S = \bar{S}\}$ dominates $EWFB$. This is because u is small, hence $\Omega^{\max}_{\{T=\bar{T};S=\bar{S}\}}$ is close to Ω_{FB}, while the cost of the former is discretely lower than the cost of the latter. But this contradicts the statement that $\mathcal{C}_{\{T=\bar{T};S=\bar{S}\}}$ is empty. **QED**

Proof of Proposition 4: We prove the three claims in reverse order:

(iii) If k is close to zero, the cost of the first-best NV contract $\{NV; \tau = \tau(\theta)\}$ is close to c. This is strictly less than any non-empty non-NV contract in class $\mathcal{K}_2$, hence the first-best NV contract dominates any non-empty non-NV contract.

(ii) Recall that the first-best NV contract $\{NV; \tau = \tau(\theta)\}$ costs $(1+8k)c$. If $k < 1/2$, this cost is lower than the cost of the EWFB in contract class $\mathcal{K}_1$, that is $\{T = \gamma; S = \sigma - \gamma\}$, therefore the EWFB in class $\mathcal{K}_2$ is $\{NV; \tau = \tau(\theta)\}$. One can then apply the same logic that we used in the proof of Proposition 2 to show that this contract must be optimal for a non-empty cost interval $c \in (0, \bar{c}^{NV})$.

(i) It suffices to show that, if $k > 1/2$ and k is sufficiently close to $1/2$, then an NV contract is optimal for a nonempty cost interval. As we argued in the text, the contract $\{NV; \tau = \tau_0(\cdot)\}$ dominates $\{T = \bar{T}; S = \bar{S}\}$ for k sufficiently close to $1/2$. This in turn implies that $\{NV; \tau = \tau_0(\cdot)\}$ is the cheapest contract among the ones that implement an outcome "close" to the first best. One can then use a similar argument as in the proof of Proposition 2 to establish that such a contract must be optimal for a nonempty cost interval. **QED**

Proof of Proposition 5: We will consider those contracts that include rigid strong bindings (RSBs) and are optimal for some c, and show that, if we replace RSBs with rigid weak bindings (RWBs), efficiency is weakly increased. We can focus on the following contracts: (a) the best $\{T = \bar{T}\}$ contract; (b) the best $\{T = \bar{T}; S = \bar{S}\}$ contract; and (c) the best $\{NT, \tau = \bar{\tau}, s = \bar{s}\}$ contract. Note that for the NV-based contract, the rigid binding on the tariff $\tau = \bar{\tau}$ serves only as a reference point, therefore nothing changes if it is replaced with $\tau \leq \bar{\tau}$.

Let us start with contract (a). Consider replacing $\{T = \bar{T}\}$ with $\{T \leq \bar{T}\}$. This can decrease efficiency only if in some state the government chooses $T < \bar{T}$ and this implies lower global welfare than $T = \bar{T}$. Let us argue that this cannot happen. If G goes below $\bar{T}$, it must be that the Nash import tax T^{NE} is below $\bar{T}$, in which case G will choose $T = T^{NE}$. Let us show that Ω increases as we move down from $\bar{T}$ to T^{NE}. Recalling that G chooses subsidy $S = S^R(T)$, we need to evaluate the derivative

$$\frac{d}{dT}\Omega(T, S^R(T)) = W_T(T, S^R(T)) + \frac{d}{dT}W^*(T, S^R(T))$$

We want to show that this derivative is negative for all $T > T^{NE}$. Clearly $W_T < 0$ for $T > T^{NE}$. Also, the sign of $\frac{d}{dT}W^*(T, S^R(T))$ is the same as the sign of $\frac{d}{dT}p^*(T, S^R(T))$. Simple algebra reveals that

$$\frac{d}{dT}p^*(T, S^R(T)) = \frac{\lambda(\beta^* + \lambda^*) - (\beta + \lambda)(\beta + \beta^* + \lambda^*)}{A(\beta + \beta^* + \lambda^*)} = -\frac{\beta(\beta + \beta^* + \lambda^*) + \lambda\beta}{A(\beta + \beta^* + \lambda^*)} < 0$$

This implies that switching to a weak binding cannot decrease Ω.

Next consider contract (b) and replace it with $\{T \leq \bar{T}; S \leq \bar{S}\}$. For a given state, the relevant possibilities are four: (i) G chooses $(T = \bar{T}, S = \bar{S})$; (ii) G chooses $(T = \bar{T}, S = S^R(\bar{T}))$; (iii) G chooses $(T = T^R(\bar{S}), S = \bar{S})$; (iv) G chooses $(T = T^{NE}, S = S^{NE})$. In case (i) of course there is no change in Ω. In case (ii) it must be that $\bar{S} > S^R(\bar{T})$. Let us evaluate $\Omega_S = W_S + W_S^*$.Clearly, $W_S < 0$ for $S > S^R(\bar{T})$, and $W_S^* < 0$, hence $\Omega_S < 0$ in this region, which in turn implies that switching to weak bindings increases Ω. In case (iii) it must be that $\bar{T} > T^R(\bar{S})$. Let us look at $\Omega_T = W_T + W_T^*$. Since $W_T < 0$ for $T > T^R(\bar{S})$, and $W_T^* < 0$, it follows that $\Omega_T < 0$ in this region, which ensures that switching to weak bindings increases Ω. In case (iv) the same result can be shown by combining the arguments we just made for cases (ii) and (iii).

Next consider contract (c). Since this contract fixes the wedge $q - p^*$ and leaves the wedge $p - p^*$ discretionary, it is convenient to re-define variables as follows:

$$\begin{aligned} p - p^* &\equiv z \\ q - p^* &\equiv v \end{aligned}$$

We can think of z and v as the policy instruments and of contract (c) as imposing a constraint $v = \bar{v}$. Also, it is useful to rewrite the world price as a function of v and z as

$$p^* = (\alpha + \alpha^* - \beta z - \lambda v)/A$$

and the reaction function for z as

$$z^R(v) = \frac{M + \lambda(v - \sigma)}{\beta^* + \lambda + \lambda^*} + \gamma$$

Let us now replace contract (c) with $\{NT, \tau \leq \bar{\tau}, s \leq \bar{s}\}$. In the new notation, this means replacing the constraint $v = \bar{v}$ with the constraint $v \leq \bar{v}$. We can apply a similar argument as for contract (a): it suffices to show that, for any given state, $\Omega(v, z^R(v))$ is decreasing in v for $v > v^{NE}$. Thus we need to study the derivative

$$\frac{d}{dv}\Omega(v, z^R(v)) = W_v(v, z^R(v)) + \frac{d}{dv}W^*(v, z^R(v))$$

Clearly, $W_v < 0$ for $v > v^{NE}$. Also, $\frac{d}{dv}W^*(v, z^R(v))$ has the same sign as $\frac{d}{dv}p^*(v, z^R(v))$. It can be checked that

$$\frac{d}{dv}p^*(v, z^R(v)) = -\frac{\lambda}{\beta + \beta^* + \lambda^*} < 0$$

This implies that switching to weak bindings cannot decrease Ω.

We have shown that replacing RSBs with RWBs cannot decrease the expected surplus. It remains to show that there is some state $\boldsymbol{\theta}$ for which replacing RSBs with RWBs increases Ω strictly. Applying the arguments developed above, we know that a sufficient condition for this to happen is that the noncooperative level for a policy is below the ceiling of that policy. It is easy to show that there exists a state of the world for which this is the case, and hence if $\boldsymbol{\Theta}$ is large enough it will encompass such state. **QED**

9. References

Bagwell, K. and Staiger, R., (2001), “Domestic Policies, National Sovereignty and International Economic Institutions,” *Quarterly Journal of Economics*, 116(2), pp 519-562.

Bagwell, K. and R. Staiger (2005), “Enforcement, Private Political Pressure and the GATT/WTO Escape Clause,” mimeo.

Battigalli, P. and G. Maggi (2002): “Rigidity, Discretion, and the Cost of Writing Contracts,” *American Economic Review*, 92, 798-817.

Battigalli, P. and G. Maggi (2003): "International Agreements on Product Standards: an Incomplete-Contracting Theory," NBER Working Paper no. 9533.

Copeland, B.R., "Strategic Interaction among Nations: Negotiable and Non-Negotiable Trade Barriers" *Canadian Journal of Economics*, 23, pp. 84-08.

Horn, H., "National Treatment in the GATT/WTO," mimeo.

Johnson, H. (1954), “Optimum Tariffs and Retaliation, *Review of Economic Studies*, 21(2), 142-153.

Maggi, G. and A. Rodríguez-Clare (2005), “A Political-Economy Theory of Trade Agreements", mimeo.

Figure 1a. Uncertainty over Demand and the Consumption Externality: Decreasing the Foreign Export Supply Elasticity by Increasing alphabar.

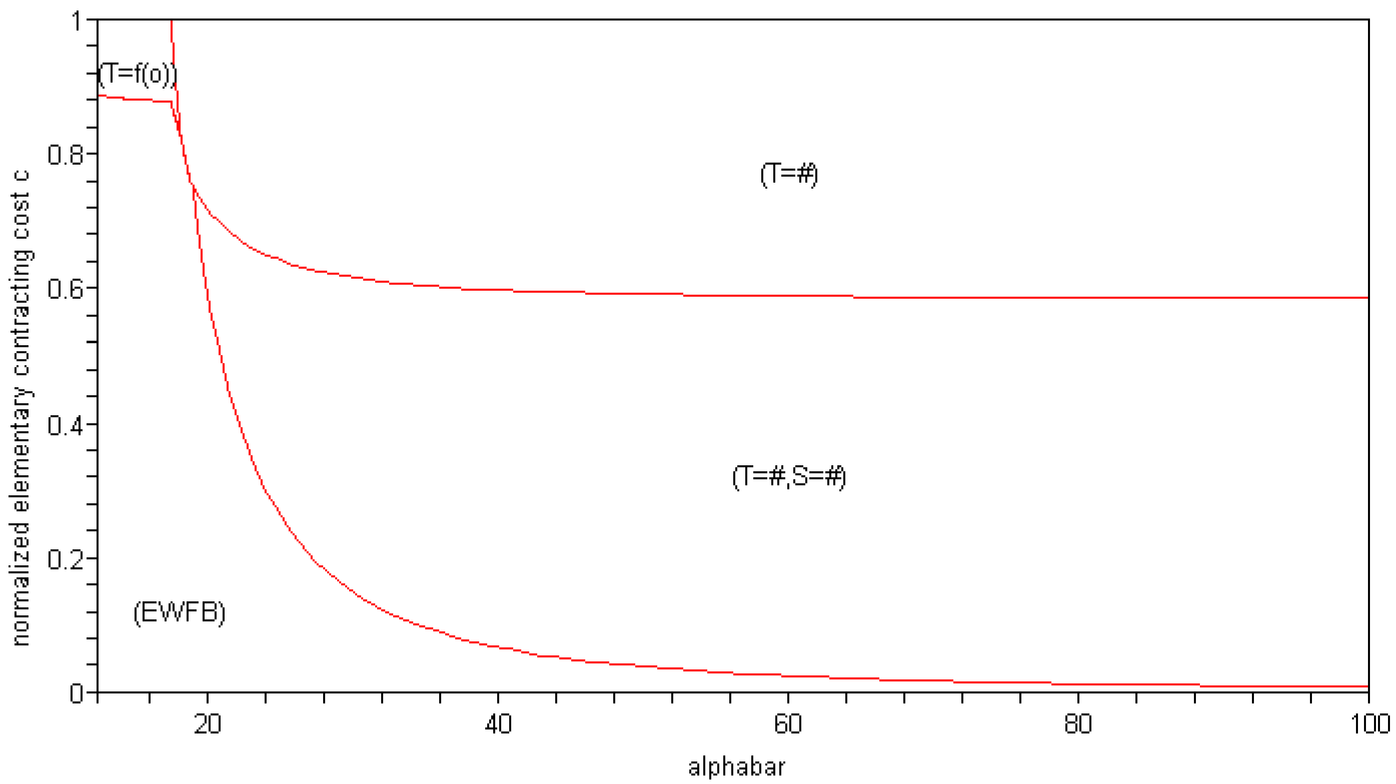

Figure 1b. Uncertainty over the Consumption Externality: Decreasing the Substitutability between Import Taxes and Domestic Policies by Increasing Beta.

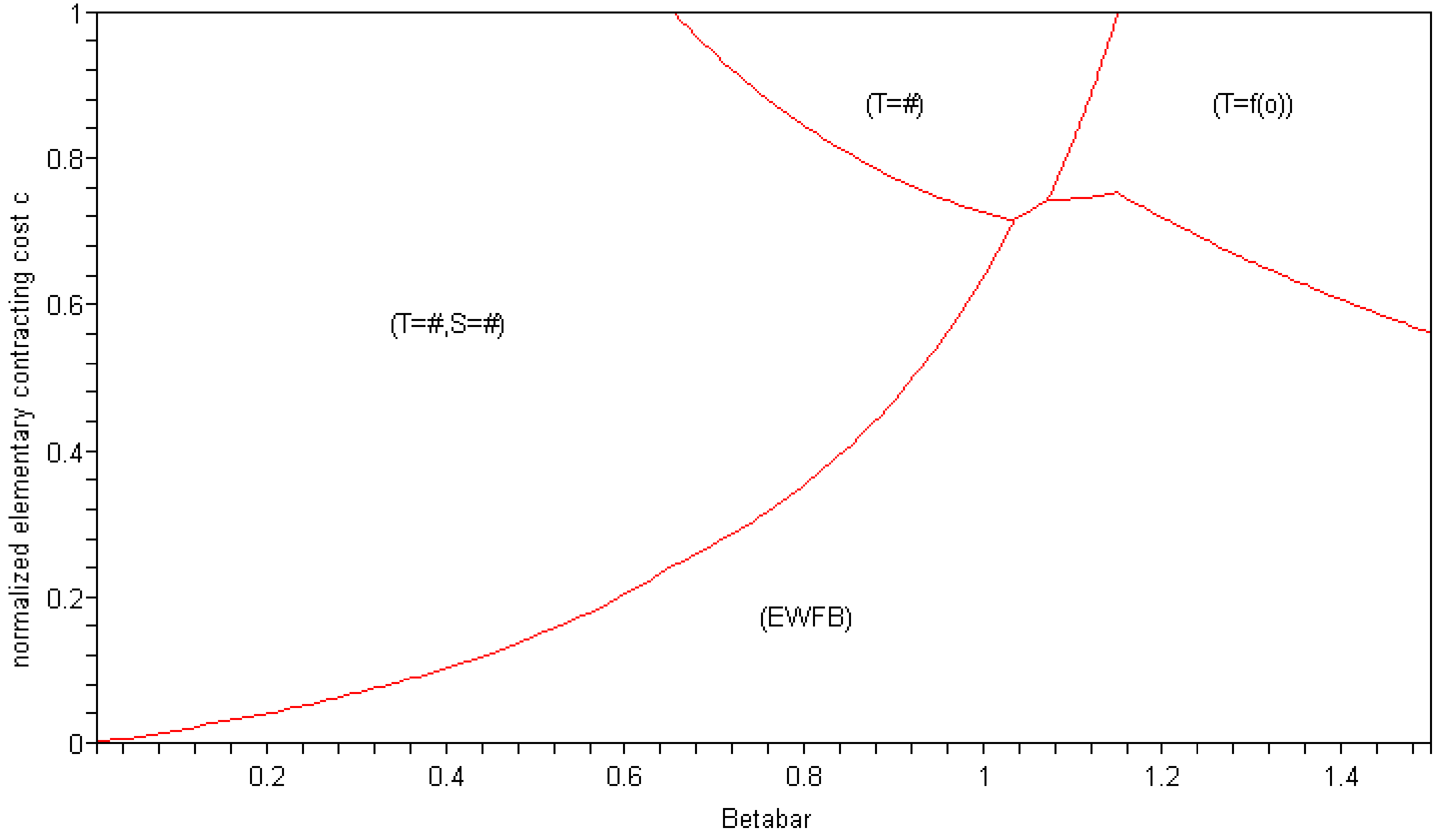

Figure 1c. Uncertainty over Demand and the Consumption Externality: Increasing Uncertainty over Demand.

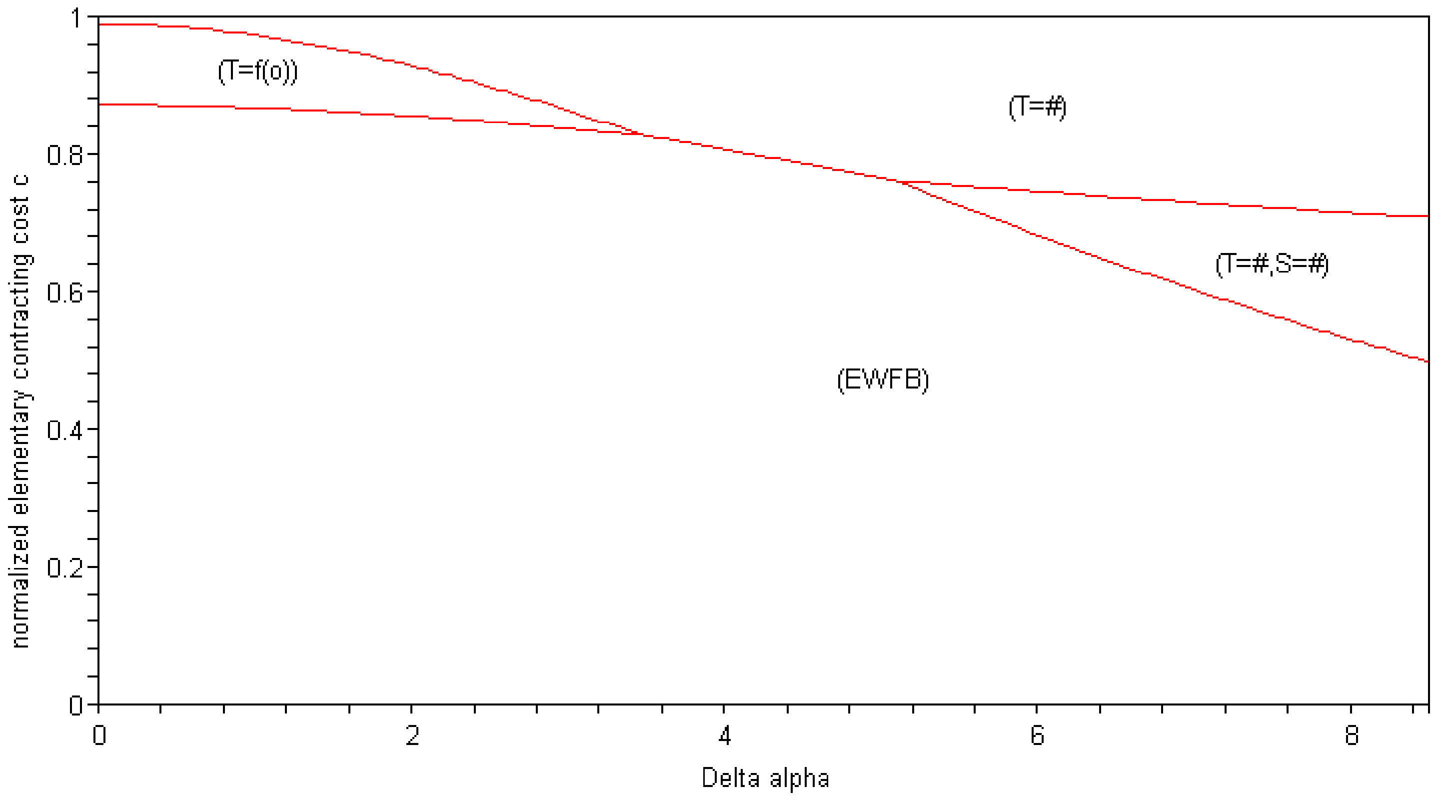

Figure 1d. Uncertainty over Demand and the Consumption Externality: Increasing Uncertainty over the Consumption Externality.

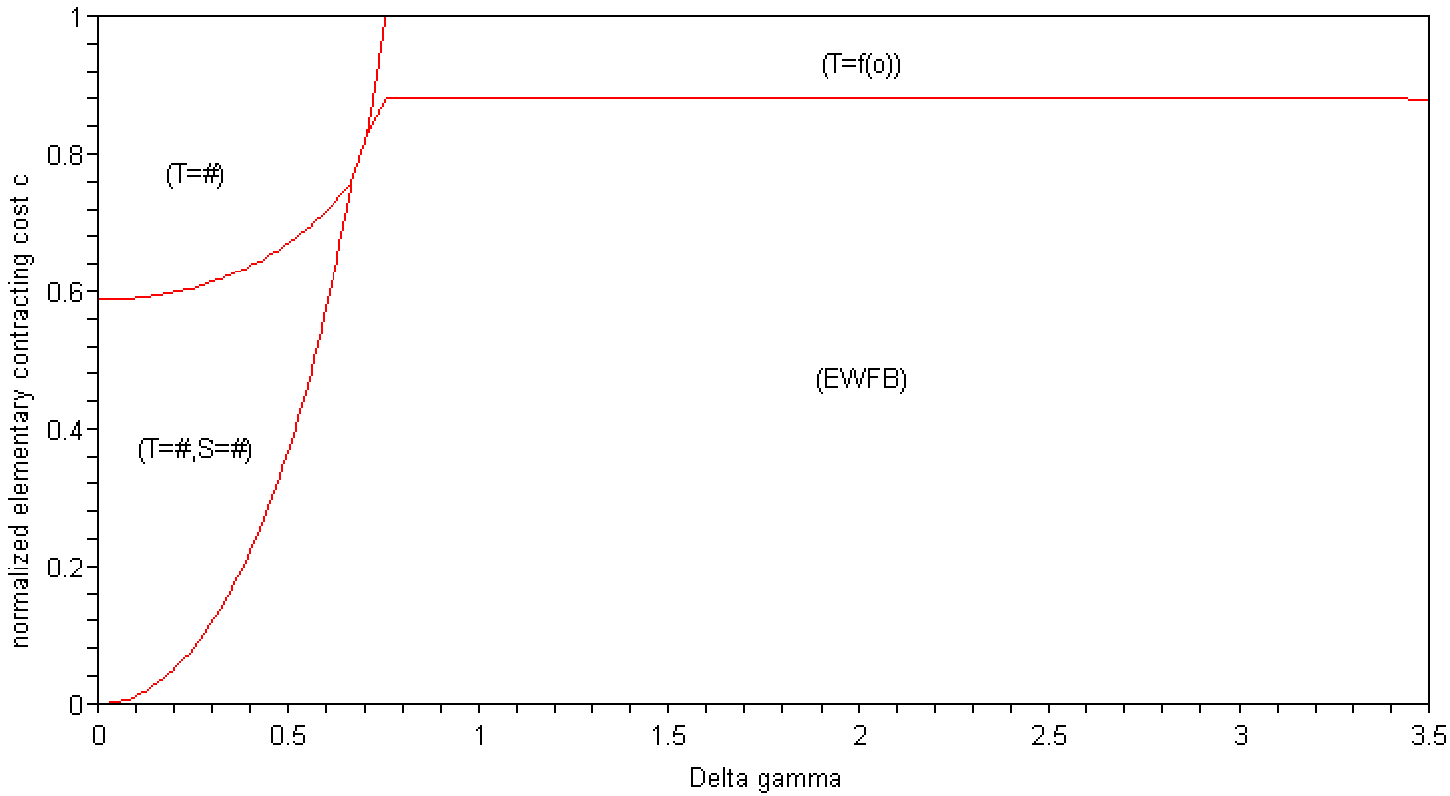

Figure 2a. NT with Uncertainty over Demand and the Consumption Externality: Decreasing the Foreign Export Supply Elasticity by Increasing alphabar.

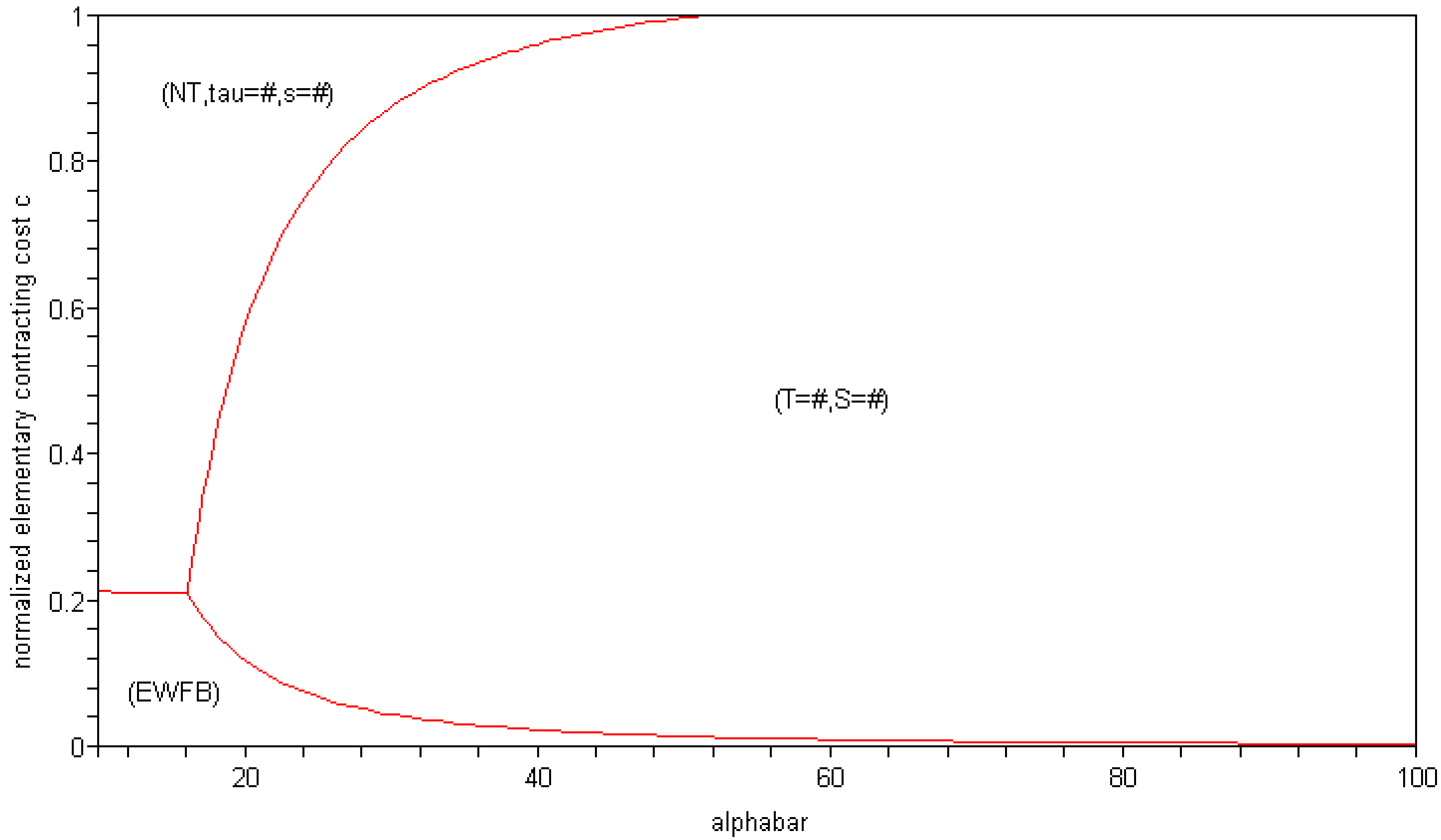

Figure 2b. NT with Uncertainty over Demand and the Consumption Externality: Increasing Beta.

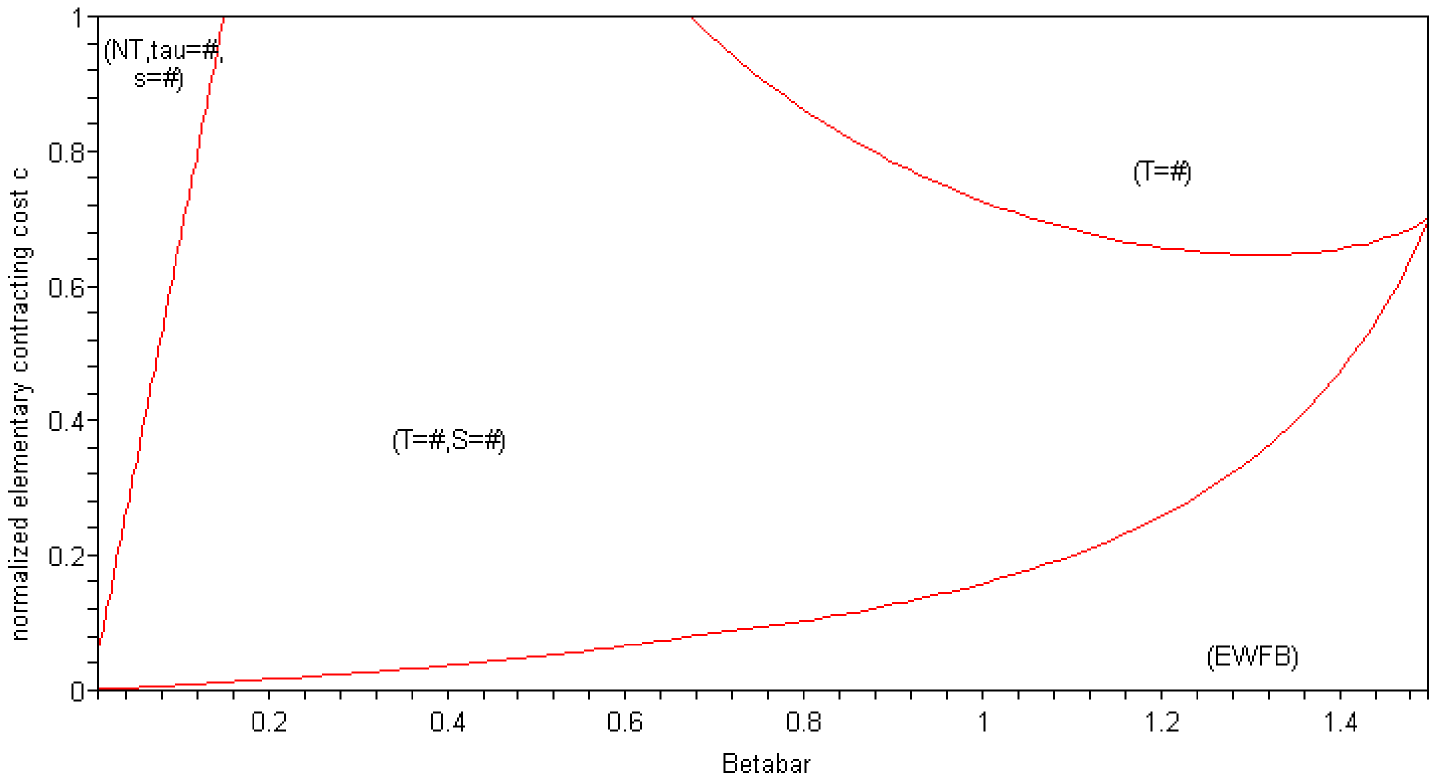

Figure 2c. NT with Uncertainty over Demand and the Consumption Externality: Increasing Uncertainty over the Consumption Externality.

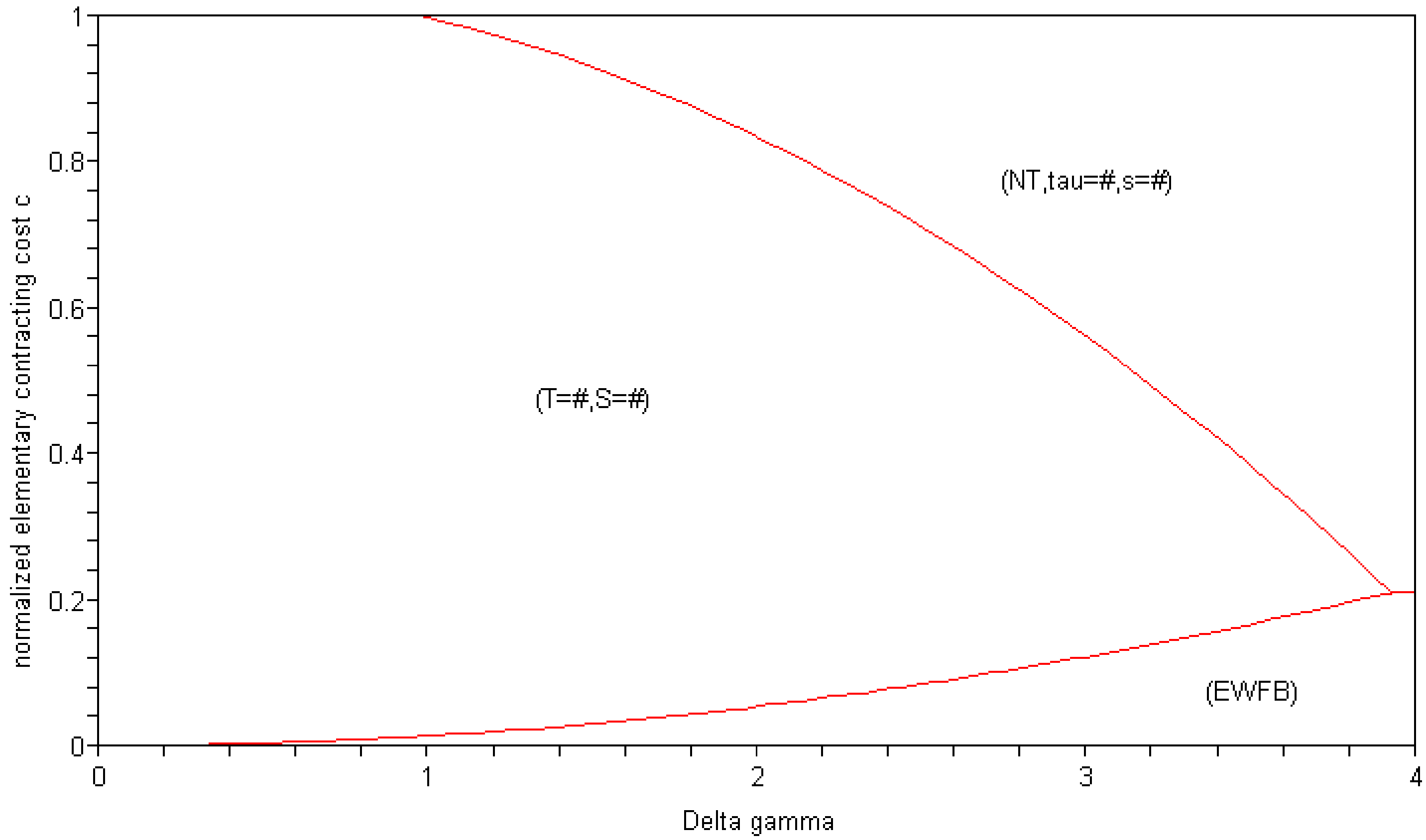

Figure 2d. NT with Uncertainty over Demand and the Production Externality: Increasing Uncertainty over the Production Externality.

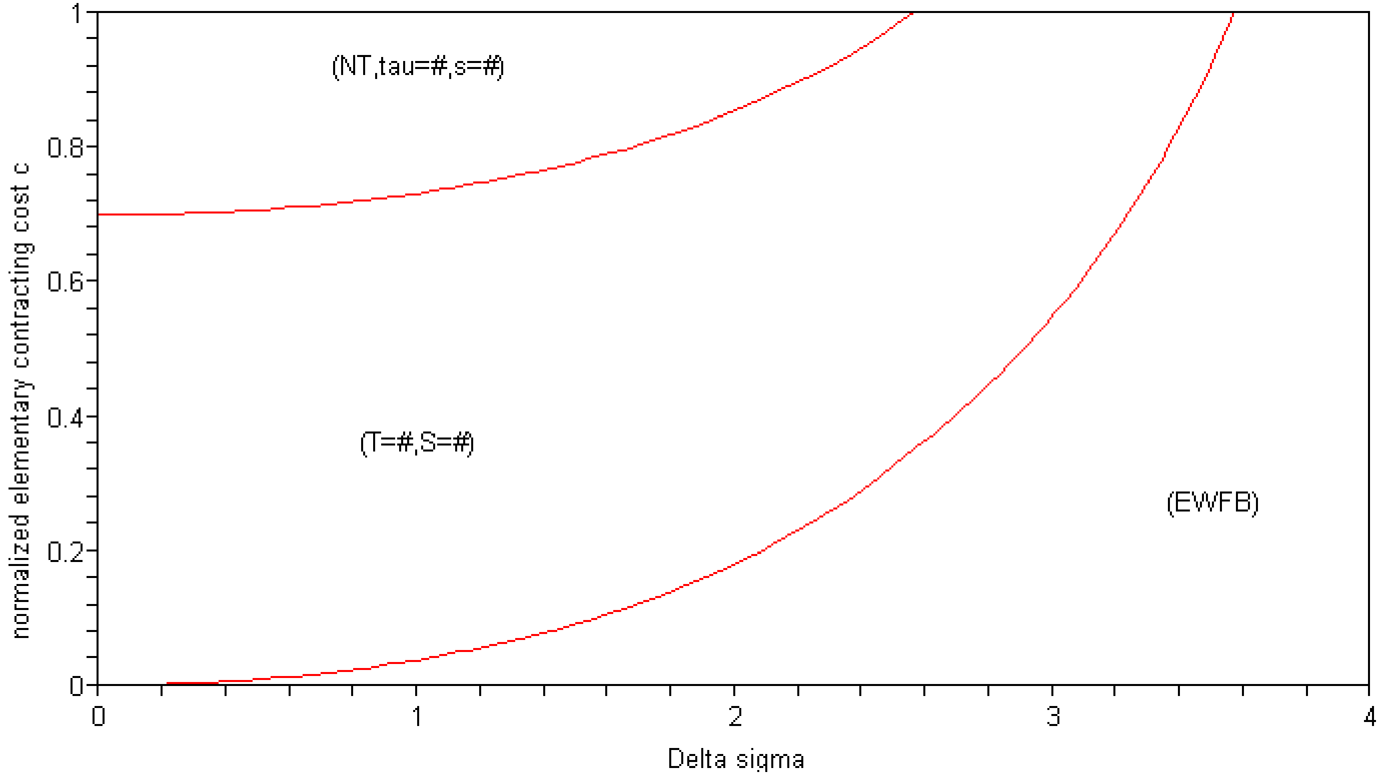

Figure 3a. NT and NV with Uncertainty over Demand and the Production Externality: Increasing Beta.

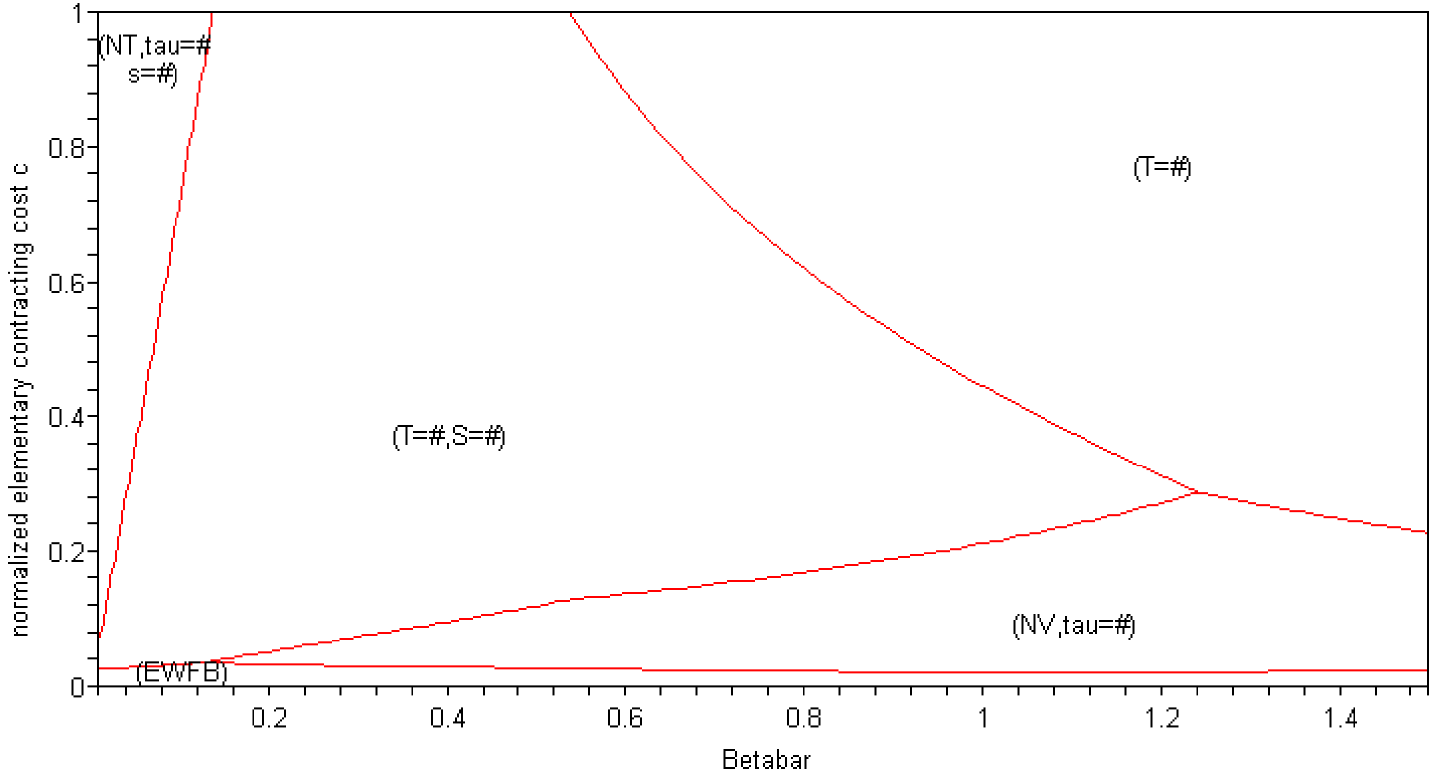

Figure 3b. NV with Uncertainty over Demand and the Production Externality: Increasing Uncertainty over the Production Externality.