The burden of proof in trade disputes and the environment

Henrik Horn a,b,c,*

a The Research Institute of Industrial Economics (IFN), Stockholm, Sweden
b Bruegel, Brussels, Belgium
c Centre for Economic Policy Research, London, United Kingdom

ARTICLE INFO

Article history:
Received 14 April 2009
Available online 9 April 2011

Keywords:
National Treatment
Burden of proof
Environment
GATT
WTO
Trade agreements

ABSTRACT

The WTO leaves discretion over environmental policies to its members, but requests that a fundamental non-discrimination principle is respected: National Treatment (NT). The provision seeks to prevent protectionist use of domestic policy instruments, requesting that when an imported product is sufficiently similar to a domestic product, they are treated identically. WTO adjudicators will often face severe informational problems in environmental disputes. Important for the practical implementation of NT is therefore the allocation of the burden of proof (BoP). This paper highlights basic implications of the BoP for the occurrence of judicial errors, for the environment and for welfare, using a setting where NT serves its intended role of supporting negotiated tariff liberalization. The paper suggests that NT may indeed constrain environmental policies, but that this may be desirable from an efficiency point of view. Also, BoP rules that benefit the environment may not benefit global welfare, and conversely.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

A basic restriction imposed by the World Trade Organization (WTO) on members’ environmental policies is the National Treatment (NT) clause in Art. III of the General Agreement on Tariffs and Trade (GATT). Similar provisions can also be found in a number of other agreements in the WTO. Art. III GATT covers virtually all domestic policies, environmental policies included, that “directly or indirectly” affect trade in goods, outlawing policies that treat imported products worse than “like” or “directly competitive or substitutable” domestic products. NT bans both policies that explicitly discriminate on the basis of origin, but also origin-neutral measures that for protectionist reasons fall more heavily on imported goods.1 The ambit of NT is therefore potentially, depending on interpretation, extremely wide.

The basic problem with regard to domestic instruments for the formation of a trade agreement is that they can substitute for border instruments. For instance, an environmental tax that is levied only on an imported product can largely replace an import tariff, rendering a tariff binding pointless. One possible solution would of course be to negotiate all internal instruments just like border instruments. But internal instruments can take an endless variety of forms, they can be applied in a huge number of different and changing circumstances, sometimes for what members would generally

---

1 NT disputes in the WTO almost always concern the latter, de facto type of discrimination. For instance, the ban on asbestos-containing products at issue in EC – Asbestos applied equally to domestic and imported products. The panel still found it to violate NT, since its effect fell more heavily on imported, asbestos-containing goods than on local, non-asbestos-containing “like” products. (The Appellate Body reversed this determination based on the controversial finding that the products were not “like”.) Environmental measures were also found to be de facto discriminatory in US – Gasoline and Brazil – Re-treaded Tyres.
consider as legitimate purposes, such as to protect the environment and sometimes for protectionist purposes. It would therefore be prohibitively costly to directly negotiate all internal instruments, if at all possible.

The GATT has chosen an alternative path. Members have discretion over domestic instruments, but these must be set in non-discriminatory fashion—they have to respect NT. The practical implementation of this principle has turned out to be difficult for at least two related reasons. One is the vagueness of the provision. For instance, it does not define central concepts such as “protection” and “like.” Fifty years of case law has failed to give any reasonable precision to such terms. A second difficulty is that adjudicators cannot directly observe the objectives that are being pursued through the contested policies; if they could, the agreement could stipulate that members should set internal policies so as to achieve some common goal. Adjudicators in the WTO thus typically face considerable uncertainty regarding both factual circumstances and the requirements of the law. They have significant investigative powers in principle, but they are restrained by the (non-ultra petita) norm that the judge should not make the case for the parties, and therefore must generally rely on evidence presented by the parties. Central to the evaluation of such evidence, and hence to the practical ambit of the provision, is the distribution of the burden of proof (BoP) between complainants and respondents. The allocation of the BoP is likely to affect the probability that measures will be deemed illegal. For instance, it is likely to make a considerable difference to the expected outcome of environmental disputes if complainants have to show that trade-restricting environmental measures are protectionist, or if importing countries have to prove that their measures do not amount to protection.

The primary purpose of the paper is to highlight some basic implications of the allocation of the BoP in environmental Art. III GATT disputes, in a setting where NT serves its intended role of supporting negotiated tariff liberalization. But the analysis may also shed light on other similar regulations of environmental instruments in the WTO Agreement, where the allocation of the BoP affects importing countries’ discretion over environmental policies. In particular, the Agreement on the Application of Sanitary and Phytosanitary (SPS) Measures requests members to use scientific evidence when making risk assessments in the SPS area; such issues were discussed in e.g. EC – Hormones, dealing with a ban on hormone-treated beef, and in EC – Approval and Marketing of Biotech Products, addressing a ban on genetically modified products. The SPS Agreement is in this respect an application of the NT principle, since the purpose of relying on scientific evidence is to prevent the use of protectionist measures that are motivated by flimsy claims that imported and domestic products should be treated differently.

The paper develops a two-country, partial equilibrium trade model. In a first stage, governments negotiate a contractually incomplete agreement that binds tariffs and imposes a NT-type restriction on taxes. Importing country governments may then privately observe environmental shocks associated with imports, after which they unilaterally determine their internal taxes on imported and locally produced goods; for instance, they may observe the detrimental environmental effects of the imported product, or the public’s understanding of such effects. Governments have standard reasons for restricting trade, but an environmental shock would add a reason for imposing a higher total tax on imports. If the shock is sufficiently severe, this will be desirable also from a global perspective. The unilaterally set taxes may in a final stage be challenged in a trade dispute, in which case an adjudicator determines their legality. The adjudicator cannot determine with certainty whether the measure is desirable from a global efficiency point of view, but the adjudicator is more likely to accept a measure when the importing country has been environmentally affected. Judicial mistakes will be committed in that governments will sometimes be allowed to differentiate their internal taxation between domestic and imported products, despite the fact that this reduces welfare from an international perspective, and in other instances governments have to remove differential taxation schemes even though they are globally desirable by protecting the environment. A central determinant of the propensity to make such mistakes is the allocation of the BoP. The paper analyzes how it affects negotiated tariffs, the nature of judicial mistakes, global welfare and environmental damage from imports.

Two general conclusions emerge from the analysis. First, the model suggests a tension between NT, at least as interpreted here, and environmental concerns. For instance, environmentally affected governments have stronger reasons to differentiate their taxation, since they have both environmental and protectionist reasons for doing this. As tariffs are reduced, NT will therefore first apply to governments facing environmental shocks, but not necessarily to purely protectionist governments. In this sense, NT tends to bind from the wrong side of the spectrum of government types.

Second, because of its rather intricate implications, the BoP is a rather blunt instrument to use for improving the environment and welfare. For instance, shifting the BoP toward complaining (exporting) countries will have a positive direct effect on the environment by reducing the number of disputes in which NT is wrongfully imposed. But it will also affect the negotiated tariff—this is indeed the purpose of NT in the first place. The direction of this change is not clear a priori. But if it lowers the tariff, imports will increase, thus at least partly offsetting the direct environmental gain. In addition, since the shift of the BoP also will affect complainants in disputes against purely protectionist governments, its welfare consequences are ambiguous.

Turning to the related literature, several studies consider the impact of some form of NT provision for the use of environmental standards, assuming that tariff levels are exogenous; recent examples are [4,5]. But there are very few
studies examining how NT affects the interaction between trade negotiations and domestic taxation, despite the fact that a main rationale for the provision is to provide incentives for countries to make tariff concessions.³ Closely related in certain respects is the study of the implications of a Most-Favored Nation clause in [9]; this clause restricts the extent to which products from different exporting countries can be treated differently, using border or domestic instruments. Somewhat related is also the literature on imperfect monitoring in trade agreements, analyzing how the design of dispute settlement institutions affects outcomes in trade agreements; for instance, [10] considers the impact of adjudication errors. There are also points of tangency with the discussion in [11] concerning whether trade agreements should include environmental provisions in cases where environmental externalities are not transboundary. Finally, there is a rapidly growing literature that considers various aspects of trade agreements in the presence of privately observed political pressures, often using more elaborate models of information transmission than is employed here; see e.g. [12–14].

The structure of the rest of the paper is as follows. Section 2 briefly discusses salient features of the BoP in NT disputes in the WTO. Section 3 presents the model of the economy and of adjudication. Section 4 derives unilaterally optimal taxes. Tariff negotiations and the resulting forms of judicial errors are considered in Section 5. Section 6 examines how the general stance on the allocation of the BoP affects the environment and welfare, and the ensuing section highlights the unconstrained (targeted) optimal allocation of the BoP. Section 8 briefly summarizes the main findings, and discusses several of the assumptions of the model. An Appendix available online analyzes a parametric version of the model, and discusses some of the underlying assumptions.

2. The BoP in GATT/WTO NT disputes

Art. III.1 GATT sets its sight on measures that are “applied so as to afford protection.” The more specific constraints with regard to fiscal instruments are contained in Art. III.2, which requests that imported products must not be taxed “in excess” of domestic “like” products.⁴ If an environmental measure is found to violate Art. III GATT, the respondent will almost certainly take recourse to Art. XX GATT, which contains general grounds for exceptions from any other provision in the agreement. For a measure to be eligible for an exception, it must not constitute “a disguised restriction on international trade.” Art. XX provides two main possibilities for an environmental dispute to be exonerated: Art. XX(b) refers to measures that are “necessary to protect human, animal or plant life or health,” and Art. XX(g) refers to measures “relating to the conservation of exhaustible natural resources if such measures are made effective in conjunction with restrictions on domestic production or consumption.”

There are two firmly established general principles for allocating the BoP in GATT/WTO disputes.⁵ One is that the party claiming a violation has responsibility for providing prima facie evidence for its claim; the other is that the party claiming an exception must demonstrate eligibility for the exception.⁶ The concept of prima facie evidence seems to have limited practical implications, however. For instance, there is no initial stage where the adjudicating body determines whether a complaint will be heard. Instead, courts typically seem to only weigh the totality of the evidence presented before them.

Standing Art. III.2 GATT case law requests of complainants to show that taxation is “in excess,” but the burden of persuasion for this has been set quite low. If an importing country seeks an Art. XX exception, it must show in particular that the measure is not “disguised protection,” and possibly that it is “necessary.” The burden of persuasion for this is likely to be rather high, even though courts seem to be more deferential when health and perhaps also environmental values are at stake.

A fundamental issue for the allocation of the BoP in NT disputes is to balance the incentives for countries to reduce their impediments to trade against the possibility to pursue legitimate regulations even if they result in higher taxation of foreign than of domestic products. Since the parties will not be able to provide adjudicators with full information concerning the circumstances of the respective disputes, judicial errors will inevitably be committed, and it would seem as if some of these errors have the potential to lead to significant environmental damage. The design of the BoP system must therefore weigh the costs of false positive findings of violations (Type I errors) against the costs of erroneous acquittals (Type II errors) while, at the same time, taking account of the implications for tariff liberalization. The analysis to follow is intended to shed light on these trade-offs.⁷

---

³ To the best of our knowledge, this is only done in elementary form in [6], in much more detail in [7], and in a framework where the structure of the agreement (including NT) is endogenously determined, in [8].

⁴ This description of the text and the case law is extremely rudimentary. Let it just be noted that there is also a second category of product pairs, with somewhat different rules, which are disregarded here—“directly competitive or substitutable” products. Ref. [1], and in particular [2], discuss more thorough Art. III.2 GATT and its case law, and [15] examines the permissible reach of national environmental policies under the WTO Agreement.

⁵ Ref. [16] analyzes the treatment of the BoP in NT disputes, as well as in other disputes of relevance to environmental regulations.

⁶ The two afore-mentioned principles hence specify the burden of production: the party that will lose the dispute if no evidence is brought forward. It is in principle distinct from the other main aspect of the burden of proof, the burden of persuasion, which specifies the amount of evidence required to fulfill the burden of production.

⁷ The Law and Economics literature also points to several other aspects of the BoP. A common concern is the impact on legal costs. As briefly argued in Section 8, and developed in some further detail in the Appendix available online, such costs could have a significant qualitative impact on the interaction between (potential) complainants and respondents. A second recurrent theme in the literature is that the optimal distribution of the BoP should reflect the distribution of information across parties. In particular, it is commonly found that the BoP should be put on the better informed party; see e.g. [17] for a formal analysis along these lines, and [18] for a dissenting view.
3. The model

Consider a two-sector, partial equilibrium model in which two symmetric countries, Home and Foreign, trade. In one sector, the good is produced in both Home and Foreign, but for simplicity there is no Foreign consumption of the product, so the product is imported by Home. In the other sector Home exports to Foreign, but since this sector is a complete mirror image, it need not be explicitly analyzed. In its import sector, Home levies an import tariff $t$, an internal tax $r$ on the domestic product, and an internal tax $s$ on the imported product, all of which are non-negative; countries do not have access to export taxes. It is assumed that the resulting Home import volume $M(r,t)$ increases in the tax on the domestic product, falls in the total tax $t+s+r$ levied on imports, and falls if both total taxes increase by the same amount: $M_t < 0$, $M_r > 0$, and $M_s + M_t < 0$. It is also assumed that Foreign’s export profits in the sector, denoted $\Pi^f(r,t)$, increases in Home’s tax on the competing domestic product ($\Pi^f_t > 0$), falls in the total tax on imports ($\Pi^f_t < 0$), and that an equal increase in $r = s$ reduces export profits $(d/dz)\Pi^f(z,z+r) < 0$.

The welfare of the Home government stems from three sources. The first is the possibly “politically” weighted sum of consumer surplus, producer surplus, and government revenue that is created in Home’s import sector, denoted $Y(r,t)$. The second source is the possible welfare loss from environmental damage caused by imports in this sector. The extent of such damage will depend on the level of imports $M(r,t)$, as well as on the intensity of the environmental shock, captured by a stochastic parameter $\theta \geq 0$. The total Home welfare derived from this sector, denoted $V(r,t,\theta)$, is

$$V(r,t,\theta) \equiv Y(r,t) - \theta M(r,t)$$

(1)

It is assumed that the environmental shock is only experienced in the importing country. The most natural interpretation is that the environmental problem stems from a negative consumption externality in the importing country. An example might be the importation of asbestos-containing construction materials, with the uncertainty in $\theta$ reflecting the government’s and/or the public’s understanding of the health effects of asbestos. The externalities may also arise in connection with the transportation of the product. For instance, oil imported on ships may expose the environment to hazards that do not arise when oil is transported through pipelines from local wells, or the packaging of imported hazardous products may differ from those of locally produced similar goods.\(^8\) The assumption that the exporting country is not affected by the realization of $\theta$ is compatible with the assumption that there is no consumption of the product in Foreign, but it is still maintained mainly for analytical convenience. Allowing for environmental impact of the shock also in Foreign would imply that the governments’ taxation problems become interdependent, which would complicate the analysis, but which does not seem to be of first-order interest for the issues at stake here.

For expositional convenience, $\theta$ takes on one of two values $\theta_L$ and $\theta_H$, where $\theta_H > \theta_L \geq 0$. The discussion will for the most part focus on the case where $\theta_L = 0$, i.e., where a “type” $L$ government does not experience any environmental shock from imports.\(^9\) The only reason for such a government to tax imports is protectionism, and it will therefore be referred to as “purely protectionist”. Governments experiencing $\theta = \theta_H$ will be referred to as “environmentally affected.” They have the same protectionist incentives as the purely protectionist governments, but also want to restrict imports due to their adverse environmental impact. Type $H$ governments are “affected” in the sense that their imports give rise to environmental damage, but the extent of such damage will also depend on the extent to which they restrict imports (which in turn will depend on whether the trade agreement so permits).\(^10\)

The third component of the Home government’s welfare is exporting firms’ profits; alternatively, it could also be the employment generated from exports. The total welfare of a Home government of type $i = L,H$ is thus the welfare from the import-competing sector $V(r,t,\theta_i)$ plus the profits from exports.

Given the symmetry of the two countries, it is natural to let global welfare be represented by the sum of the two governments’ welfare. Due to the partial equilibrium analysis and the symmetry between the sectors, we can focus on the industry in which Home is an importer; an analysis of the other sector would be identical. Letting $W$ denote aggregate government welfare created in the Home market,

$$W(r,s + \tau,\theta_i) \equiv V(r,s + \tau,\theta_i) + \Pi^f(r,s + \tau)$$

(2)

It is assumed that both $V$ and $W$ are strictly concave in $(r,t)$; a couple of further assumptions will be made below.

The sequence of events is as follows:

1. the parties agree on a contractually incomplete trade agreement that puts a rigid upper bound on the permissible tariff $t$, and also includes a NT provision that may require equal taxation $(r = s)$;
2. governments privately observe their environmental shocks;
3. governments set taxes; and

---

8 A recent actual example may serve as yet another illustration: it is alleged that oil drilling to take place in Cuban waters may expose the US to larger environmental risks than such extraction in US waters, since the drilling in Cuban waters may use less safe technology. $\theta_H$ would then correspond to the choice of such inferior technology.

9 We use the term “type” although governments are \textit{ex ante} identical.

10 The “purely protectionist” label for type $L$ governments is less adequate for cases where $\theta_L > 0$, since such governments would then have also an environmental motive for taxing imports, albeit a weaker environmental motive than type $H$ governments.
4. if an exporting country litigate, alleging a NT violation, there is adjudication.

If a dispute arises concerning whether a pair of taxes violates NT, the judge can either exonerate the importing country, in which case it maintains its disputed taxes, or let the complainant prevail, in which case the respondent must change its taxation such that \( r \geq s \). The judge can determine without error whether \( r \geq s \). But the judge will have imperfect information concerning the details of the situation, and will consequently not know whether it would be better from a global efficiency point of view to allow the differential taxation. We will not model the details of this decision making, which often can be expected to be highly complex in environmental disputes. Instead, we assume that the outcome of the adjudication is random, but that it is more likely that the complainant will prevail against a purely protectionist government than when the importing government experiences an environmental shock. Let \( \gamma_i \) denote the probability that the complainant (the exporting country) prevails against a type \( i \) government, which is assumed to have the following properties:

\[
\begin{align*}
\gamma_L = \gamma_H = 0 & \quad \text{for } r \geq s \\
0 < \gamma_H < \gamma_L < 1 & \quad \text{for } r < s
\end{align*}
\]

A central factor determining the magnitude of \( \gamma_i \) is the allocation of the BoP: the more the burden is put on complainants in disputes against type \( i \) governments, the lower is \( \gamma_i \). A general shift of the BoP toward complainants is therefore captured by reductions in both \( \gamma_H \) and \( \gamma_L \). Our main concern is to determine how the allocation of the BoP, by affecting the occurrence of adjudication errors, affects global welfare and the environment.

Finally, we will occasionally refer to a parametric version of the model, which assumes linear demand, linear supplies with Home and Foreign firms using the same production technology, perfect competition, and social welfare-maximizing governments. JEEM’s online repository (http://www2.econ.iastate.edu/jeem/supplement.htm) holds an Appendix providing derivations of the main results of the paper for this model.

4. Tax setting

We now turn to the two alternative situations in which taxes may be set. In the first case the importing country is unconstrained by NT, and in the second it has to respect NT.

4.1. Unilaterally optimal total taxation

If unconstrained by NT, government \( i \) would choose the taxes \( \hat{t}(\theta_i) \) and \( \hat{\theta}_i \) solving

\[
\max_{\hat{t}, \hat{\theta}_i} V(r; t; \theta_i)
\]

It is assumed that Home governments have unilateral incentives to tax imports higher than the domestic product absent environmental problems, since NT would otherwise serve no role: \( 0 \leq \hat{t}(\theta_i) < \hat{\theta}_i \). It is also assumed that the larger the damage from imports, the more Home governments want to differentiate the total taxation: \( \hat{t}_i \leq 0 \) and \( \hat{\theta}_i > 0 \).

When unilaterally determining total taxes, Home governments disregard the impact on foreign profits. They therefore levy a too high total tax on the imported product, regardless of whether it causes environmental damage. Let \( (t^i(\theta_i), t^o(\theta_i)) \)

\[\begin{align*}
\gamma_L = \gamma_H = 0 & \quad \text{for } r \geq s \\
0 < \gamma_H < \gamma_L < 1 & \quad \text{for } r < s
\end{align*}\]

11 This is not a self-evident depiction of Art. III.2 GATT, which as noted above also applies to measures that “indirectly” give domestic products more favorable treatment.

12 The information available to the adjudicator could easily be given a more explicit description also within the current approach. For instance, let \( p \) be the adjudicator’s probability assessment after the evidence has been presented that a pair \( r < s \) is applied so as to afford protection. The burden of persuasion requires that \( p \geq b \) in order for the adjudicator to determine in favor of the complainant. Let \( p_0 \) be the adjudicator’s prior. Since the burden of production rests with the complainant, \( p_0 \) is sufficiently low relative to \( b \) that should no evidence at all be presented in the dispute proceeding, so that there is no updating of the prior and thus \( p = p_0 \), the complainant loses: \( p_0 < b \). Let \( p_0 + \nu(\theta) \) be the probability assessment after the parties have presented their arguments, where \( \nu(\theta) \in [0, -a] \), \( 0 < a < 1 \). It is not necessary to impose any particular assumptions on the distribution of this stochastic variable, except that \( E(\nu(\theta_i)) < E(\nu(\theta_i)) \). Naturally, for \( \nu > 0 \), \( p = \min(1, p_0 + \nu) \), and for \( \nu < 0 \), \( p = \max(0, p_0 + \nu) \). With this formalization, we would have that \( \gamma_i = \Pr(p_0 + \nu(\theta_i) \geq b) \). It appears as if such a formulation of the model would mainly add notational complexity, as long as the realization of \( \nu \) is exogenous.

13 This formalization of the adjudication process rests on two additional, highly simplifying assumptions. One is that the probability a complainant wins a case where \( r < s \) is independent of the absolute magnitude \( s - r \). The second assumption is that there are no litigation costs. These assumptions are briefly discussed in Section 8, and a somewhat longer discussion can be found in the downloadable Appendix.

14 It follows from the first-order conditions for interior unilaterally optimal taxes, \( V_i(r, t, \theta_i) = 0 \) and \( V_i(r, t, \theta_i) = 0 \), that

\[
\begin{align*}
\hat{t}_o &= \frac{V_{t}\hat{V}_{t} + V_{t}\hat{V}_{o}}{V_{t}\hat{V}_{t} - V_{t}\hat{V}_{o}} = \frac{V_{t}\hat{M}_t - V_{t}\hat{M}_o}{V_{t}\hat{V}_{t} - V_{t}\hat{V}_{o}} \\
\hat{\theta}_o &= \frac{-V_{\theta}\hat{V}_{t} + V_{\theta}\hat{V}_{o}}{V_{\theta}\hat{V}_{t} - V_{\theta}\hat{V}_{o}} = \frac{V_{\theta}\hat{M}_t - V_{\theta}\hat{M}_o}{V_{\theta}\hat{V}_{t} - V_{\theta}\hat{V}_{o}}
\end{align*}
\]

The denominator is negative due to the concavity of \( V \) and \( V_{t}\hat{M}_t < 0 \), \( V_{t}\hat{M}_o > 0 \). A sufficient condition for \( \hat{t}_o < 0 \) and \( \hat{\theta}_o > 0 \) is that \( V_{t} \hat{V}_{t} < 0 \). These inequalities are also fulfilled for \( V_{t} > 0 \), provided it is not too large.
be the first-best total taxation, which solves

$$\max_{\tau, t} W(\tau, t, \theta_i)$$

(5)

We thus have

$$\hat{t}(\theta_i) > t^*(\theta_i) \text{ and } \hat{t}(\theta_i) \leq t^*(\theta_i) = 0 \text{ for } i = H, L$$

(6)

where the weak inequality allows for a corner solution.

Importantly for what follows, let $\hat{t}(\theta_i)$ be the highest tariff level at which the discretionary taxes just fulfill the NT rule; formally, $\hat{t}(\theta_i) = \hat{t}(\theta_i) - \hat{t}(\theta_i)$. For an agreed tariff $\tau \geq \hat{t}(\theta_i)$, the government can implement $(\hat{r}(\theta_i), \hat{t}(\theta_i))$ by setting $r = \hat{r}(\theta_i)$, $t = 0$, and $\tau = \hat{t}(\theta_i) - \tau$. But for $\tau < \hat{t}(\theta_i)$, the tariff alone yields less total taxation of imports than government $i$ prefers, and it is unilaterally optimal to use the domestic tax on imports to make up for the difference: $\hat{t}(\tau, \theta_i) = \hat{t}(\theta_i) - \tau$. Note that $\hat{t}(\theta_i) < \hat{t}(\theta_H)$; environmentally affected governments will need to impose an internal tax on the imported product for a larger range of tariffs.

The resulting Home and global welfare from the domestic market in the case of discretionary tax setting are thus, respectively:

$$\hat{V}(\theta_i) \equiv V(\hat{r}(\theta_i), \hat{t}(\theta_i), \theta_i); \quad \hat{W}(\theta_i) \equiv W(\hat{r}(\theta_i), \hat{t}(\theta_i), \theta_i)$$

(7)

4.2. Taxation under the NT constraint

If constrained to set equal internal taxes, government $i$ would choose the tax level

$$\pi(\tau, \theta_i) = \arg\max_{\tau} V(\pi, z + \tau, \theta_i)$$

with the associated first-order condition

$$V_t(z, z + \tau, \theta_i) + V_z(z, z + \tau, \theta_i) = 0$$

(9)

where $V_t < 0$ and $V_z > 0$ in optimum. We assume that an increase in the tariff induces the importing country to reduce its taxation of both the imported and the domestic product, but that the total taxation of imports increases as a result of the tariff increase:\footnote{Sufficient for this to hold, as well as for the second-order condition – which requires the denominator to be negative – is that $V_t$ is either negative, or not too large, if positive.}

$$-1 < \bar{z} = -\frac{V_{tt} + V_{zz}}{V_{tt} + 2V_{tt} + V_z} < 0$$

The domestically generated welfare under the NT regime for government $i$ is

$$\pi(\theta_i, \tau) \equiv V(\pi(\tau, \theta_i), \pi(\tau, \theta_i) + \tau, \theta_i)$$

(10)

where $\pi(\tau, \theta_i) < \hat{V}(\theta_i)$ when NT binds.

The resulting global welfare with a binding NT restriction is

$$\hat{W}(\tau, \theta_i) \equiv W(\pi(\tau, \theta_i), \pi(\tau, \theta_i) + \tau, \theta_i)$$

(11)

It follows from $\hat{W}(\hat{\tau}(\theta_i), \theta_i) = \hat{W}(\theta_i)$ that\footnote{Note that

$$d \pi = \pi^* \pi + (1 + \pi^*) \pi^*$$

where $\pi^* > 0, \pi > 0, 1 + \pi > 0$, and $\pi^* < 0$.}

$$\hat{W}(\hat{\tau}(\theta_i), \theta_i) = \frac{d}{d\tau} W(\pi(\tau, \hat{\tau}(\theta_i)), \pi(\tau, \hat{\tau}(\theta_i)) + \hat{\tau}(\theta_i), \theta_i) = \frac{d}{d\tau} \hat{\pi} < 0$$

(12)

That is, in a situation where the NT constraint is just binding for government $i$, a marginal reduction of the tariff enhances global welfare with such a government in Home. Denote the optimal tariff with NT binding and with government $i$ as\footnote{A variable with an upper bar pertains to a situation where NT binds, and superscript $G$ denotes variable levels or functions that are optimal from a global point of view.}

$$\hat{\tau}(\theta_i) = \arg\max_{\tau} \hat{W}(\tau, \theta_i)$$

It thus follows that $\hat{\tau}(\theta_i) < \hat{\tau}(\theta_i)$, and that $\hat{W}(\hat{\tau}(\theta_i), \theta_i) > \hat{W}(\theta_i)$. That is, it is always optimal to constrain government $i$ if the optimal tariff for this government type can be imposed.

Fig. 1 illustrates the relationship between expected welfare and the tariff for three levels of $\theta_i$: $\theta_L < \theta_H < \theta_H$. The uppermost curve pertains to the case with $\theta_i = \theta_H > 0$; for $\theta_i = 0$ we would have $\hat{\tau}(0) = 0$. For $\tau \geq \hat{\tau}(\theta_i)$, welfare is $\hat{W}(\theta_i)$, and then increases as the tariff is reduced until it reaches the optimal level for this type of government, $\hat{\tau}(\theta_i)$. The other curves have similar shape, but correspond to situations with environmental problems.
In order for the adjudication problem to be non-trivial, it is assumed that the difference between \( y_H \) and \( y_L \) is sufficiently large that environmentally affected governments can be erroneously constrained by NT. Sufficient for this to hold is that it is preferable to not constrain environmentally affected governments through NT at the tariff level that is optimal for purely protectionist governments:

\[
w(\tau, y_L) > w(\tau, y_H)\]

Finally, there are in general two values of \( \tau \) at which \( w(\tau, y_i) = \hat{w}(y_H) \). The higher value was defined above as \( \hat{\tau}(y_H) \). The lower tariff level at which this equality holds is denoted \( \tau_{\min}^{G}(y_H) \); it is illustrated in Fig. 1 for \( y_i = y_H \).

4.3. The choice of taxation strategy

We have so far derived Home’s optimal discretionary taxes, and the optimal tax given a binding NT constraint. It can directly be seen that for \( \tau < \hat{\tau}(y_i) \) it will be optimal for government \( i \) to set \( \hat{r}(\tau, y_i) \) and \( \hat{s}(\tau, y_i) \) regardless of whether it will face a complaint. This extreme feature of the model follows from two assumptions. First, there are no costs associated with taking part in, or losing, a dispute. Second, neither the probability of litigation, nor its outcome, is affected by the extent to which the two taxes diverge. The implications of relaxing these assumptions are further discussed in Section 8.

The equilibrium behavior of government \( i \) when facing a \( \tau \) at which \( w(\tau, y_i) = \hat{w}(y_H) \). The higher value was defined above as \( \hat{\tau}(y_H) \). The lower tariff level at which this equality holds is denoted \( \tau_{\min}^{G}(y_H) \); it is illustrated in Fig. 1 for \( y_i = y_H \).

5. Tariff negotiations and the nature of judicial errors

Governments cannot condition tariffs on environmental shocks, but negotiate an \textit{ex ante} efficient tariff agreement. The negotiations also involve Foreign’s import tariff in the background sector. Equal tariffs in the two sectors yield an equal division of the \textit{ex ante} gains from the agreement (the \textit{ex post} welfare depends on the realizations of the environmental shocks in the two countries). Due to the symmetry of the model, we can focus on the bargaining outcome that maximizes \( W \).

Trade negotiators will take into consideration the implications of any agreement for subsequent tax setting, including the possibility that judicial errors will be committed. \textit{A priori}, one can identify four types of errors, since NT may be incorrectly imposed (Type I error) or incorrectly not imposed (Type II error) for each type of government. Of particular interest is the case where environmentally affected governments are incorrectly required to respect NT, since this scenario seems to capture an often heard critique of the WTO in the policy debate. This will be referred to as an “innocent bystander” scenario, although environmentally affected governments have the same protectionist incentive as the other
The expected welfare in this sector with Home government of type $i$ is $\hat{w}(\theta_i)$ for $\tau \geq \hat{\tau}(\theta_i)$, and $\gamma_i \mathbb{W}(\tau, \theta_i) + (1-\gamma_i)\hat{w}(\theta_i)$ for $\tau < \hat{\tau}(\theta_i)$. It will always be optimal to choose a tariff below $\hat{\tau}(\theta_i)$. But how much below? There are two possibilities. The first is to set the tariff sufficiently high that only environmentally affected governments are constrained by NT, that is, $\hat{\tau}(\theta_i) \leq \tau < \hat{\tau}(\theta_i)$. The optimal tariff is then $\tau_k = \tau^*(\theta_i)$, which maximizes welfare specifically for type $H$ governments.

Letting $\phi_i$ denote the fraction of governments of type $i$ (i.e., the probability of an environmental threat), the resulting expected welfare is

$$w_k(\tau_k) = \phi_i \hat{w}(\theta_i) + \phi_H \gamma_H \mathbb{W}(\tau^*(\theta_i), \theta_H) + \phi_H (1-\gamma_H)\hat{w}(\theta_H)$$

Since $\mathbb{W}(\tau^*(\theta_i), \theta_H) > \hat{w}(\theta_H)$, judicial errors are only of Type II, i.e., erroneous acquittals.

The other possibility is to set the tariff low enough that NT constrains both government types, i.e., $\tau < \hat{\tau}(\theta_i)$, in which case the maximand is

$$w^E(\tau, \cdot) = \sum_{i=L,H} \phi_i [\gamma_i \mathbb{W}(\tau, \theta_i) + (1-\gamma_i)\hat{w}(\theta_i)]$$

The optimal tariff is then

$$\tau_k(\cdot) = \text{argmax}_{\tau < \hat{\tau}(\theta_i)} \phi_i \gamma_i \mathbb{W}(\tau, \theta_i) + \phi_H \gamma_H \mathbb{W}(\tau^*(\theta_i), \theta_H)$$

and it balances the commercial gains from a low tariff against the environmental costs, should an environmental shock occur. As $\phi_H$ increases from 0 to 1, the optimal tariff increases from $\tau^*(\theta_i)$ to $\tau^*(\theta_H)$.

Four types of situations may arise when NT constrains both types of governments:

$$w_k(\tau_k, \cdot) = \phi_i \gamma_i \mathbb{W}(\tau_k, \cdot, \theta_i) + \phi_H \gamma_H \mathbb{W}(\tau_k, \cdot, \theta_H)$$

The first two terms capture cases with purely protectionist governments. Recalling that $\mathbb{W}(\tau, \theta_i) > \hat{w}(\theta_i)$ for a range $\tau \leq \hat{\tau}(\theta_i)$, the first term pertains to cases where such governments correctly lose their disputes, and the second term captures situations with false acquittals (i.e., Type II errors). The third and fourth terms pertain to environmentally affected governments. The third term captures situations where the importing country is required to impose equal taxation. This will be desirable from a global point of view if $\tau > \tau^*(\theta_i)$, since the tax difference is then larger than it should be from an international point of view given the relatively high tariff, but it will be a false positive finding (i.e., a Type I error) in the opposite case. Finally, the fourth term reflects situations where environmentally affected governments are allowed to continue with differential taxation. This may be desirable, but will be a Type II error if $\tau < \tau^*(\theta_i)$. Hence, judicial mistakes will be committed for both types of governments. While in the case of purely protectionist governments the error can only be to accept measures that should have been struck down, in the case of environmentally affected governments, the qualitative nature of the errors will depend on the extent of trade liberalization. With limited liberalization, the adjudicator may erroneously accept differential taxation when NT should be imposed, and in case of more substantial liberalization, may also err by imposing NT when differential taxation should be permitted.18

Which equilibrium will arise? It will be optimal for tariff negotiators to set the tariff such that both types of governments are constrained by NT when $w_k(\tau_k, \cdot) > w_k(\tau_b, \cdot)$. A sufficient condition for this to hold is that $\tau^*(\theta_H) \leq \hat{\tau}(\theta_i)$, since type $L$ governments will then be constrained at the tariff level that it is optimal specifically for environmentally affected governments; this case is illustrated in Fig. 1 with type $H$ welfare being represented by the middle curve. In the case where $\mathbb{W}(\tau_k, \theta_i) > \hat{w}(\theta_i)$ it is ambiguous whether it is optimal to constrain only one or both types of governments.

The difference in expected welfare is

$$w_k(\tau_b, \cdot) - w_k(\tau_k, \cdot) = \phi_i \gamma_i [\mathbb{W}(\tau_b, \cdot, \theta_i) - \hat{w}(\theta_i)] + \phi_H \gamma_H [\mathbb{W}(\tau_b, \cdot, \theta_H) - \mathbb{W}(\tau^*(\theta_H), \theta_H)]$$

Taking into account the envelope properties with respect to $\tau_b$ and $\tau_k$, it is straightforward to see that it is more likely to be optimal to constrain both governments:

(a) the lower is the potential environmental damage $\theta_H$, since targeting type $H$ governments will then be less beneficial:

$$\frac{d}{d\theta_H} [w_k(\tau_b, \cdot) - w_k(\tau_k, \cdot)] = \phi_H \gamma_H [\mathbb{W}(\theta_H) - \mathbb{W}(\tau_k(\theta_H), \theta_H)] - \phi_H \gamma_H [\mathbb{W}(\tau_b(\theta_H)) - \mathbb{W}(\tau^*(\theta_H), \theta_H)] < 0$$

(b) the lower is the probability $\gamma_H$ that environmentally affected countries must respect NT, since the tariff will not be optimally chosen for instances when both types are constrained;

Note that the term judicial "error" does not refer to whether the court correctly pronounces on the existence of the environmental shock, but to whether the decision to allow the taxes yields higher welfare than disallowing it, at the negotiated tariff. The term reflects the notion that the primary role of the judge is to resolve disputes in a manner that jointly benefits the parties, rather than to find facts.
(c) the higher is the probability \( \gamma_L \) that purely protectionist governments must respect NT, since the alternative of letting them set tariffs discretionary is worse.\(^{19}\)

The following observation stems directly from the above:

**Proposition 1.** NT always constrains governments that are exposed to environmental shocks, but not necessarily governments that differentiate taxes solely for protectionist purposes.

NT thus seems environmentally unfriendly in a rather fundamental sense, in that environmentally affected governments always risk being constrained by NT, while the same is not true for purely protectionist governments.\(^{20}\) From an environmental point of view, NT hence binds from the “wrong” side of the spectrum of government types.

Before turning to the role of the BoP, we make a slight digression. It might intuitively seem that both the welfare cost and the environmental damage from erroneous impositions of NT should be larger, the more severe is the environmental shock \( \theta_H \).\(^{21}\) Assume, however, that there exists a common tax level \( \xi(t) \) such that imports are choked off for higher NT-compatible taxes; this level will be lower the higher is the tariff. Then, even if constrained by NT, Home can always achieve the autarky welfare level \( Y(\zeta(t),\xi(t) + \tau) \) in its import sector, regardless of the magnitude of \( \theta_H \), by setting \( r = s = \xi(t) \). This outcome can of course also be achieved if Home is unconstrained by NT. Consequently, the global cost of wrongful imposition NT (as defined above) as \( \theta_H \) becomes large, is bounded in this case:

**Proposition 2.** Assume that \( M(\xi(t),\xi(t) + \tau) = 0 \) for \( z \geq \xi(t) \), for any given \( \tau \). The global welfare cost of an erroneous imposition of NT is then bounded above as \( \theta_H \) becomes very large, and for sufficiently large \( \theta_H \) there will be no imports, and no environmental damage.

Then, what does it take to shut out imports? As long as the tariff is strictly positive, \( r = s = \xi(t) \) implies that the total tax on imports is higher than the tax on the domestic product. Hence, if production technologies are the same in both countries and trade is driven by differences in demand, there will be production and consumption of the domestic product even with environmental damage being completely prevented. But if trade is driven by lower production costs for exporters, it may be necessary to shut down domestic production and consumption as well. In practice, the importing country government may find this too costly. But if so, this is the government’s choice, it is not forced upon the country government by the WTO Agreement. Put differently, NT contains a form of safety valve that allows importing country governments to fully achieve their environmental objectives, regardless of the magnitude of environmental shocks, provided they are prepared to pay the price in terms of reduced commercial surplus.\(^{22}\)

6. **A general shift of the BoP toward complainants**

It is occasionally suggested that to prevent the WTO from unduly restricting members’ ability to pursue legitimate regulatory objectives, a rather high burden of persuasion should be imposed on complainants. This section examines the appropriateness of such recommendations within the model developed above, by examining how the environment and welfare is affected by a general shift of the BoP toward complainants, that is, a shift that would make it harder for complainants to prevail against both types of governments. The case where the governments can be separately targeted will be examined in the next section.

6.1. **NT only constrains environmentally affected governments**

To determine the role of the BoP for environmental damage, we need first to determine how it affects the negotiated tariff. If only type \( H \) governments are constrained by NT, the negotiated tariff will effectively maximize \( W(\tau, \theta_H) \), and will consequently be unaffected by the allocation of the BoP. A general shift in the burden toward complainants – \( d\gamma_H < 0 \) – will then reduce environmental damage by reducing the frequency of instances where the environmentally constrained government will have to respect NT. As can be seen from (13), the welfare impact of such a change will be unambiguously negative – it will be given by \( \phi_H(W(\tau(\theta_H, \theta_H)) - W(\theta_H)) d\gamma_H < 0 \), since \( W(\tau(\theta_H, \theta_H)) > W(\theta_H) \). Hence:

**Proposition 3.** If NT only constrains environmentally affected governments, a general shift of the BoP toward complainants unambiguously improves the environment, but unambiguously reduces global welfare.

---

\(^{19}\) As shown in the downloadable Appendix, in the parametric model, NT constrains both governments if and only if \( \theta_H < \frac{\pi}{2\sqrt{1 + (\phi_H c_H / \phi_H c_H)}} \).

\(^{20}\) In a more general model where governments are exposed to varying degrees of environmental problems and where not all of them are constrained by NT, the constrained governments would all be more environmentally affected than the governments that are not constrained.

\(^{21}\) The direct welfare cost of an erroneous imposition of NT is \( W(\theta_H) - W(\xi(t) + \tau, \theta_H) \). It is not obvious how the parameter \( \theta_H \) affects the magnitude of this cost, although it increases in \( \theta_H \) in the linear model.

\(^{22}\) There are in practice other means of resolving disputes in which significant costs are imposed on importing countries as a result of successful complaints, to be briefly discussed in Section 8.
6.2. NT constrains both types of governments

More ambiguous is the case where both types of governments are constrained by NT. The first-order condition associated with (15) is then

\[ w^E(t, \gamma_L, \gamma_H) = \phi_H \gamma_L W_{t0}(t, \theta_L) + \phi_H \gamma_H W_{t1}(t, \theta_H) = 0 \]  

(17)

Since \( W_{t0}(t^E, \theta_H) = -M, t > 0, W_{t1}(t^E, \theta_L) < 0 < W_{t1}(t^E, \theta_H). \) The first term in (17) represents the gain from reducing the tariff in case of purely protectionist governments that are requested to respect NT. The second term captures the cost of forcing environmentally affected governments further from the total taxation of imports that would maximize global welfare for this type. Differentiation of (17) yields

\[ d\tau_B(\gamma_L, \gamma_H) = \frac{1}{W_{t1}} \phi_H \gamma_H W_{t1}(t, \theta_H) \left[ \frac{d\gamma_L}{\gamma_L} - \frac{d\gamma_H}{\gamma_H} \right] \]  

(18)

where the second-order condition for an optimal tariff level requires \( W_{t1}^E < 0 \). The sign of \( d\tau_B \) is hence the opposite to that of the term in brackets. As can be seen from (18), a reduction in \( \gamma_L \) increases \( \tau \); intuitively, with lower \( \gamma_L \) there are fewer instances where NT is imposed on type \( L \) governments, thus reducing the expected benefit from a marginal tariff reduction. Consequently, it becomes optimal to increase the tariff in order to give environmentally affected governments some extra protection. A reduction in \( \gamma_L \) has of course the opposite effect; there would be fewer instances where these latter type of governments would have to impose NT, and there is therefore less reason to protect them \emph{ex ante} through a tariff. It follows from (18) that a general change in the stance on BoP would have an ambiguous effect on the negotiated tariff level:

\textbf{Lemma 1.} A small reallocation of the BoP such that \( d\gamma_L/\gamma_L > d\gamma_H/\gamma_H (d\gamma_L/\gamma_L < d\gamma_H/\gamma_H) \) will reduce (increase) the negotiated tariff.

Hence, the tariff response to a general change in the BoP does not depend on whether the burden is shifted toward complainants or respondents, but on how the ratio \( \gamma_L/\gamma_H \) is affected.

6.2.1. Environmental effects

We can now see how a general shift in the BoP affects environmental damage. Since there is only damage in case of a type \( H \) government, the expected level of environmental damage \( K \) is

\[ K(\gamma_H, \tau) = \phi_H \theta_H[\gamma_H M(\tau, \theta_H) + (1 - \gamma_H) M(\theta_H)] \]  

(19)

where \( M(\tau, \theta_H) = M(\tau(\theta_H), \tau(\theta_H), \tau(\theta_H) + \tau) \) and \( M(\theta_H) = M(\theta(H), \theta(H), \theta(H)). \) As can be seen, the expected environmental damage in general depends both directly on the allocation of the BoP, and indirectly through the negotiated tariff. Differentiating (19),

\[ \frac{1}{\phi_H \theta_H} dK = [\tilde{M}(\tau, \theta_H) - M(\theta_H)] d\gamma_H + \gamma_H \tilde{M}(\tau, \theta_H) d\tau_B(\gamma_L, \gamma_H) \]  

(20)

There are hence both direct effects for a given tariff, and induced effects via the change in the tariff. The first term in brackets is positive—imports will be higher when NT binds than with discretionary tax setting. A reduced probability for the complainant to win in cases where the government is environmentally affected will therefore tend to improve the environment. The second term captures the change in the import volume and thus environmental damage stemming from the change in the negotiated tariff in response to the change in the BoP.\(^{23}\) If the tariff increases, environmental damage is unambiguously reduced, since the induced increase in the tariff reinforces the direct effect from having a larger fraction of the environmentally affected governments prevailing in disputes. This requires the shift in the BoP to have the property that \( d\gamma_L/\gamma_L < d\gamma_H/\gamma_H < 0.\)\(^{24}\) On the other hand, the induced effect via the change in the tariff will tend to counteract the direct effect if \( d\gamma_H/\gamma_H < d\gamma_L/\gamma_L < 0. \) We have not been able to identify any clear circumstances under which this indirect effect dominates the direct effect, nor can we generally exclude that it can occur, except for that it does not occur in the parametric model. The point is not to argue that a shift of the BoP toward complainants can have such a counter-intuitive implication, however. The point is instead that there are induced effects on negotiated tariffs from a general change in the BoP that needs to be taken into consideration when assessing the total impact of the BoP allocation.

6.2.2. Welfare effects

Turning to the welfare implication of a general shift of the BoP toward complainants when both types of governments are constrained, it is by (16) given by

\[ dw^E(\tau_B, \gamma_L, \gamma_H) = \phi_H [\tilde{M}(\tau_B(\gamma_L, \gamma_H), \theta_L) - \tilde{M}(\theta_L)] d\gamma_L + \phi_H [\tilde{M}(\tau_B(\gamma_L, \gamma_H), \theta_H) - \tilde{M}(\theta_H)] d\gamma_H \]  

(21)

\(^{23}\) The sign follows directly from

\[ \tilde{M}_t = M_{z_1} + M_t (1 + z_1) < 0 \]

where \( M_t > 0, M_t < 0, \) and \(-1 < z_1 < 0.\)

\(^{24}\) It would not suffice that \( d\gamma_L = d\gamma_H < 0, \) since the tariff would then fall due to \( \gamma_H < \gamma_L.\)
Note that there are again in general two sources of effects: the direct effect for a given tariff, and the induced effect via the change in the tariff. However, there is no first-order welfare impact from the change in the tariff here, since it is optimally set. Despite this simplifying feature, and despite the extremely simple setting more generally, the welfare impact from the proposed allocation of the BoP is unclear. On the one hand, the shift of the BoP will allow more purely protectionist governments to continue with their protectionism. The adverse welfare impact from this is captured by the first bracketed term in (21). On the other hand, the change will allow a larger share of environmentally affected governments to escape the constraint imposed by NT, and the welfare impact of this will depend on the precise circumstances. But if it leads to more acquittals of environmentally affected governments that should be constrained by NT from a global point of view, the suggested shift of the BoP will unambiguously reduce welfare, increasing the number of false acquittals for both types of governments. It will have unclear consequences in the opposite case:

**Proposition 4.** When NT constrains both environmentally affected and purely protectionist governments, a general shift of the burden of proof toward complainants, will unambiguously reduce global welfare when \( w(\tau_H, o_H) > w(\theta_H) \).

To conclude, when allowing for judicial mistakes, it is possible to identify an “innocent bystander” outcome, where environmentally affected governments are occasionally constrained by NT, despite the fact that welfare would be higher if they had discretion to set taxes, and where there is environmental damage as a result. But it is no panacea to shift the BoP onto complainants. First, it might not be a very effective environmental policy: it will increase the number disputes where environmentally affected governments can impose their discretionary taxes, and thus reduce environmental damage. But the shift of the BoP will also induce negotiators to settle for a lower tariff, and this tends to at least partly offset the direct impact. Second, it is not clear that it enhances global welfare to allow environmentally affected countries more discretion to differentiate their taxes.

7. The optimal targeted BoP

We finally turn to the case where the probabilities \( \gamma_i \) can be perfectly and independently controlled, corresponding to a scenario where the tariff is still negotiated before the realization of the environmental shock and cannot be conditioned on the outcome, but where there is full information concerning government types after the realization, thus making it possible to apply NT in a state-specific manner.

7.1. Environmental effects

It might intuitively be thought that the optimal allocation of the BoP from an environmental point of view is to target purely protectionist governments fully, since there is no environmental damage at stake in such disputes, and to always let environmentally affected governments differentiate their taxation \( (\gamma_L = 1, \gamma_H = 0) \). As can be seen from (18) and (20), matters are not quite as simple. \( \gamma_H = 0 \) would indeed be optimal from an environmental point of view since it would give the importing country a free hand to set its discretionary tax on imports. There would still be environmental damage, however, as long as the importing country government does not find it optimal to choke off imports completely. If \( \gamma_H \) can be set exactly to zero, the magnitude of \( \gamma_L \) would be irrelevant.

But the magnitude of \( \gamma_L \) will affect the environment whenever \( \gamma_H > 0 \). Note, however, that it would in such an instance be desirable from an environmental point of view to make it impossible for complainants to prevail also against purely protectionist governments \( (\gamma_L = 0) \); the benefit from maintaining a low tariff – which stems from the instances where NT restrains purely protectionist governments – would then be smaller, which would increase the negotiated tariff, and thus reduce imports.

**Proposition 5.** Assume NT constrains both environmentally affected and purely protectionist governments, and that \( \gamma_H > 0 \) and \( \gamma_L > 0 \). (i) A targeted shift of the BoP \( \delta \gamma_H < 0 \) then improves the environment by reducing the frequency of disputes where the importing country has to respect NT, but this improvement can be partly offset by an increase in imports. (ii) A targeted shift \( \delta \gamma_L < 0 \) then improves the environment despite the fact there is no environmental damage at stake in the directly affected disputes.

7.2. Welfare effects

We finally examine the optimal targeted choices of \( \gamma_L \) and \( \gamma_H \) from a global welfare point of view. It can be seen from (13) and (16) that the welfare-maximizing \( \gamma_i \) will be either zero or one, taking into account the envelope property with regard to the tariff. There are hence in principle four possible choices. But the situation that perhaps seemed to be most promising from an environmental point of view, where NT constrains neither type of government \( (\gamma_L = \gamma_H = 0) \), is clearly dominated by e.g. the outcome where purely protectionist governments are constrained \( (\gamma_L = 1, \gamma_H = 0) \), leaving the remaining three cases to be considered.

A first possibility is to only target environmentally affected governments with NT \( (\gamma_L = 0, \gamma_H = 1) \). A necessary condition for this to be optimal is that \( \tau^*(\theta_H) > \tau(\theta_L) \), since it would otherwise improve welfare to increase \( \gamma_L \) for a constant tariff \( \tau^*(\theta_H) \), to constrain more purely protectionist governments. The expected welfare when only targeting type \( H \)
governments is
\[ w_h^{\text{max}} = \phi_l \tilde{w}(\theta_l) + \phi_H \tilde{w}(\tau_H(\theta_H), \theta_H) \]

A second option is to let NT constrain both types of governments (\( \gamma_l = 1, \gamma_H = 1 \)). The resulting expected welfare would be
\[ w_b^{\text{max}} = \phi_l \tilde{w}(\tau_b(\cdot), \theta_l) + \phi_H \tilde{w}(\tau_H(\cdot), \theta_H) \]

A necessary condition for this to be an equilibrium is that \( \tilde{w}(\tau_b(\cdot), \theta_H) > \tilde{w}(\theta_H) \), since in the opposite case it would be better at a given tariff to set \( \gamma_H = 0 \).

The third possibility is to only target purely protectionist governments with NT (\( \gamma_l = 1, \gamma_H = 0 \)), while setting the tariff at \( \tau_b(\theta_l) \). This makes it possible to impose a tariff that is low enough to discipline purely protectionist governments in the desirable way, without adversely affecting environmentally affected governments. The expected welfare would be
\[ w_l^{\text{max}} = \phi_l \tilde{w}(\tau^L(\theta_L), \theta_L) + \phi_H \tilde{w}(\theta_H) \]

It is straightforward to see that it is more likely to be optimal to target both types of governments, the smaller is the potential environmental damage:
\[
\frac{d}{d\theta_H} (w_b^{\text{max}} - w_l^{\text{max}}) = \phi_H [\tilde{w}(\tau_b(\cdot), \theta_H) - \tilde{w}(\theta_H)] = \phi_H [\tilde{M}(\theta_H) - \tilde{M}(\tau_b(\cdot), \theta_H)] < 0
\]
\[
\frac{d}{d\theta_H} (w_b^{\text{max}} - w_h^{\text{max}}) = \phi_H [\tilde{w}(\tau_b(\cdot), \theta_H) - \tilde{w}(\tau^C(\theta_H), \theta_H)] = \phi_H [\tilde{M}(\tau^C(\theta_H), \theta_H) - \tilde{M}(\tau_b(\cdot), \theta_H)] < 0
\]

To shed further light on the optimal pattern, Fig. 2 illustrates for the parametric version of the model, how the outcome depends on the parameters \( \phi_H \) and \( \theta_H \) (see the Appendix for the derivations). There are hence three regions. The optimal agreement will (i) set the tariff at the level \( \tau^C(\theta_l) \) that is fully optimal for type \( L \) governments, and will ensure that NT only constrains such governments when \( (\phi_H, \theta_H) \) is in the non-shaded region; (ii) it will set the tariff at the level \( \tau^L(\theta_L) \) that is optimal for type \( H \) governments, thus accepting that purely protectionist governments escape the NT constraint, and make sure that NT is imposed on all type \( H \) governments when \( (\phi_H, \theta_H) \) is in the darkly shaded region; and (iii) it will set the tariff \( \tau_b(\cdot) \) such that both types of governments are constrained by NT when \( (\phi_H, \theta_H) \) is in the lightly shaded region.

**Proposition 6.** If the probabilities for complainers to prevail in disputes can be fully controlled, and set separately for the two types of governments, it might be optimal to either target governments of type \( H \), of type \( L \), or both types of governments. It is more likely to be optimal to target both types, the smaller is \( \theta_H \).

The proposition applies not only to the parametric version of the model, but also the main model. But the figure allows us to make a couple of additional observations that are specific to the parametric model. First, it is for natural reasons more likely to be optimal to constrain environmentally affected governments, either jointly with type \( L \) governments, or alone, the larger is the fraction of type \( H \) governments. Second, a combination of a high \( \theta_H \) and high \( \phi_H \) makes it likely to be optimal to target only type \( H \) governments. The reason is of course that if these governments are threatened by severe damage, and constitute a very small proportion of governments, there are double reasons not to target them; and conversely, if such governments are very common, and do not differ much from purely protectionist governments, there is again no reason to target them. Such governments only should be targeted when there are enough of them to importantly affect expected welfare, and when constraining them by NT is very costly (\( \theta_H \) is high).

![Fig. 2. The optimal pattern of NT under full ex post information.](image)
This case of perfect \textit{ex post} targeting points to a more general observation. The “innocent bystander” scenario does not arise here, since whenever an environmentally affected government is constrained by NT, this is better from a welfare perspective than to let it set taxes discretionary. Hence, this scenario seems to require judicial mistakes. But it cannot be inferred that an innocent bystander has been affected whenever NT is imposed on an environmentally affected government. This may be optimal to do also in the case of perfect targeting of the BoP, because of the basic construction of the trade agreement, which gives governments discretion over domestic instruments.

8. Concluding discussion

Unilateral environmental policies cause a fundamental problem for the design of trade agreements: it is often very difficult for adjudicators to determine the true rationale for policies that are allegedly pursued to protect the environment, but that simultaneously protect commercial interests. The first line of defense in the WTO against protectionist use of internal instruments is the NT provision in Art. III GATT, which jointly with Art. XX GATT shapes the basic scope for unilateral environmental policies in the WTO. The practical ambit of these provisions is importantly affected by the allocation of the BoP. The purpose of this paper is to shed light on the likely functioning of NT in environmental trade disputes, focusing on the interaction between trade liberalization and the design of the NT provision, and in particular on the role of the allocation of the BoP.

The paper draws two broad conclusions. The first is that there is a tension between NT, as interpreted here, and environmental concerns. NT tends to apply in particular to governments facing environmental shocks. NT may also be erroneously imposed when both environmental and welfare concerns speak in favor of letting the importing country set its preferred taxes, similarly to what is claimed in the “innocent bystander” argument.

The second more general conclusion is that the BoP has rather complex implications for the environment and welfare, and that it is a rather blunt instrument to achieve these objectives. First, the BoP does not only affect environmental damage directly, by influencing the propensity for environmentally affected governments to prevail in NT disputes. There are also indirect effects stemming from changes in negotiated tariffs, and these effects in turn depend on the exact properties of the BoP. Second, there is no simple relationship between the general stance on the BoP and global welfare. Whether environmentally affected and/or purely protectionist governments should be constrained by NT is ambiguous even in the case where the probabilities for the parties to prevail in disputes can be fully and independently set.

The analysis has relied on a number of strong assumptions. First, the assumed absence of litigation costs is highly simplifying, but also means that some of the properties of the equilibrium will be extreme.\footnote{See \cite{19} for an analysis of the role of litigation costs in trade disputes.} In particular, an exporting country loses nothing from complaining whenever \( r < s \), so in all such situations there will be a complaint, and the importing country stands nothing to lose from setting such taxes and thereby possibly losing a resulting dispute. The online Appendix discusses the implication of assuming that participation in disputes is costly. This may reflect the direct administrative costs of participating in dispute settlement proceedings, or possibly costs arising from erosion of the parties’ confidence in the agreement. In such a case, governments’ decisions concerning dispute settlement become strategically interlinked. Importing countries may then find it optimal to set taxes so as to avoid provoking complaints, and exporting countries do not complain in each instance where NT is violated. Such costs may also explain why respondents choose not to contest complaints. Another important feature that arises is that NT may have an impact on the outcome even in cases where it is never invoked in a dispute. Hence, the \textit{off-equilibrium threat} of being requested to impose NT may deter pronounced differential taxation.

Second, GATT/WTO adjudicators have occasionally interpreted a significant difference in taxes as a sufficient verification of a violation of Art. III GATT. A natural extension of the model would therefore be to assume that \( \gamma_L \) increases in the magnitude \( s - r \). This would create an incentive for the importing country government to limit tax differentiation. Similarly to the case of litigation costs, it would imply that NT has an impact even if it is not imposed in equilibrium through dispute settlement. A similar phenomenon would arise if the difference between \( s \) and \( r \) affected the probability that the differential taxation is detected by the exporting country, as seems likely to be the case.

Third, the analysis above did not model how adjudicators form their views concerning the nature of the preferences of importing country governments. It thus excluded strategic use of information asymmetries by the parties and adjudicators. An alternative approach would be to model the interaction as an imperfect information signaling game, in which adjudicators can rationally deduce information from the observed actions by the governments. The standard approach would assume that the adjudicator has full information concerning the structure of the welfare function, the relative frequency of different types of governments, etc., but it does not observe the realization of \( \theta_L \) for the incumbent government. With its knowledge of welfare functions, the adjudicator could then potentially infer the value of \( \theta_L \) that must have generated the observed taxes \( r \) and \( s \). Since governments of type \( L \) would lose from being identified, they would have incentives to try to mimic the behavior of type \( H \) governments, and there may potentially be a pooling equilibrium with all governments taxing in the same manner in equilibrium. But governments of type \( H \) will want to distinguish themselves from type \( L \) governments, and may under certain circumstances be able to signal their identity in a separating equilibrium. Etc.
Such an approach would have advantages and disadvantages as compared to the one employed here. An advantage of the approach we have used, besides being analytically simpler, is that the outcome of the decision process is less sensitive to the assumed details of the interaction. Related, it is not a trivial matter to capture salient features of actual environmental dispute settlement using such an approach. For instance, in *EC – Approval and Marketing of Biotech Products* the panel report contained more than 1000 pages. The Table of Contents of the panel report runs almost 40 pages. The parties submitted over 3100 documents, some of which containing more than 100 pages. The dispute only partly dealt with Art. III GATT, but seems indicative of the potential complexity of environmental disputes. It is far from clear how the essence of such a dispute could be distilled to the degree that it can be meaningfully analyzed within a formalized game of imperfect information. What seems plausible is that the outcome of such a dispute partly depends on the true motives for the regulation and partly on the general stance on the BoP; hence the approach taken here. This is by no means to deny the possibility that analysis of environmental NT disputes as games of asymmetric information might bring important insights that cannot be achieved within the current framework.

Finally, the outcome in the “innocent bystanders” scenario above is globally inefficient when the importing country government is environmentally affected. One might therefore expect the parties in such disputes to try to escape the imposed restriction, and in practice there are several avenues for this. One possibility is to renegotiate the tariff, or alternatively to negotiate a side payment to the complainant for dropping the complaint.\(^{26}\) If renegotiations or side payments were fully allowed, there would be no global welfare loss, nor any environmental impact, from Type I errors. It would not directly affect the outcome of disputes in which there are erroneous dismissals, but it would affect the optimal tariff, which would now be set at the level that is optimal for purely protectionist governments. It would then be optimal to constrain both types of governments, and purely protectionist governments would be correctly found to violate NT in the fraction \(\phi_L \gamma_L\) of all cases, they would be erroneously acquitted in the fraction \(\phi_L (1 - \gamma_L)\) of all cases, and in the fraction \(1 - \phi_L\) of cases NT would correctly not constrain environmentally affected governments. Such an extension would raise some new issues, however. A distinguishing feature of actual disputes is the fact that the contested matter has already been negotiated unsuccessfully before the court proceeding is launched. Lacking a plausible explanation for why adjudication is necessary, it does not seem unreasonable to assume that the same factors that prevent the parties from reaching an agreement before the delivery of a determination by the judge, are also at play afterwards. One may also wonder, if governments can negotiate the total taxation once the state of the environment is revealed, what is the role of the first stage tariff agreement? Could not all decisions concerning taxation be left for ex post negotiation?\(^{27}\)

Yet another possibility for losing respondents to escape the NT constraint would be to refuse implementation. Strictly speaking, this has not occurred in the WTO so far. In the model above, if a losing environmentally affected foreign government refuses to comply, Home would have to impose retaliatory tariffs in its import sector. Conversely, whenever a type \(H\) government loses a dispute, it will face retaliatory tariffs imposed by Foreign in Home’s export sector. This would clearly be a less efficient form of escape compared to renegotiation, due to the retaliatory measures that are imposed.

To conclude, the assumptions concerning no renegotiation or side payment possibilities, and full compliance, are central to the existence of Type I errors. Neither of these assumptions is unproblematic as a depiction of the WTO. On the other hand, the opposite extreme assumptions of full-information renegotiation possibilities, fully efficient side payments, and a possibility to avoid implementation, are not very descriptive either.

Acknowledgments

I am very grateful for extremely helpful discussions with Petros C. Mavroidis. I would also like to thank Pär Holmberg, Erik Lindqvist, Lars Persson, Thomas Tangerás, and participants in the *Environment and Trade in a World of Interdependence* (ENTWINED) research consortium for useful comments, and Aron Berg and Christina Lönnblad for editorial assistance. Very helpful suggestions were also provided by an Associate Editor and three Reviewers. Financial support by the Marianne and Marcus Wallenberg Foundation, and Mistra (Stockholm) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jeem.2010.10.008.

References


\(^{26}\) Such renegotiation is analyzed in [20].
\(^{27}\) See [13] for an analysis of renegotiation. Of relevance is also the analysis of safeguards in [21], where safeguards are interpreted as efficiency-enhancing means of escaping bindings.