# Optimal environmental border adjustments under the General Agreement on Tariffs and Trade*

Edward J. Balistreri†
Colorado School of Mines

Daniel T. Kaffine
University of Colorado Boulder

Hidemichi Yonezawa
ETH Zürich

March 2016

**Abstract**

A country's optimal environmental border policy includes a strategic component that is inconsistent with commitments under the General Agreement on Tariffs and Trade (GATT). We extend the theory to include GATT compliance. Theory supports optimal border adjustments on carbon content that are below the domestic carbon price, because price signals sent through border adjustments encourage consumption of emissions intensive goods in unregulated regions. The theory is supported in our applied numeric simulations. Countries imposing border adjustments at the domestic carbon price will be extracting rents from unregulated regions at the expense of efficient environmental policy and consistency with international trade law.

*JEL* **classifications: F13, F18, Q54, Q56**
**Keywords:** climate policy, border tax adjustments, carbon leakage, trade and carbon taxes.

*The authors thank Harrison Fell, James R. Markusen, Aaditya Mattoo, Thomas F. Rutherford and participants of seminars at the World Bank, and the Western Economic Association meetings for helpful comments and discussion. A portion of this research was completed while the authors were supported through a donation to the Colorado School of Mines by the Alcoa Foundation.

†Corresponding author: Division of Economics and Business, Colorado School of Mines, Golden, CO 80401-1887; email: ebalistr@mines.edu.

# 1 Introduction

An important challenge for economists and policymakers alike is that countries may coordinate on some issues while disagreeing on others. International law under the World Trade Organization (WTO) is designed to favor a cooperative trade outcome, where countries are punished if they attempt to use trade restrictions to extract rents from trade partners.[1] At the same time, however, there remains a substantial lack of coordination concerning global climate policy. The recent negotiations in Paris (United Nations Convention on Climate Change, Conference of the Parties 21, December 2015) did result in a set of pledges by participating nations to reduce emissions through their Intended Nationally Determined Contributions (INDCs) to keeping average temperature increases below 2 degrees Celsius, but did not indicate a coordinated path for achieving these goals. In fact, the resulting framework relies heavily on unilateral action and enforcement. In this context it is highly likely that a country or group of countries will consider trade restrictions such as border carbon tariffs designed to limit the *free-rider* incentives inherent in uncoordinated action (Nordhaus (2015)) and mitigate a reshuffling of emissions to non-regulated regions (carbon leakage).[2] These trade restrictions must, however, be reconciled with the commitments to cooperative trade. As such, we ask, what is the optimal border carbon tariff in the presence of cooperative trade agreements, and how does it compare to the domestic carbon price?

We modify the established theory on cross-border externalities to analytically derive optimal environmental trade distortions in the context of cooperative trade and non-cooperative

[1] Recent empirical evidence (Broda et al., 2008; Bagwell and Staiger, 2011; Ludema and Mayda, 2013) has shown that countries have a unilateral incentive to exploit terms-of-trade effects and that the WTO has been successful in neutralizing that incentive.

[2] The U.S. and the EU, in particular, have shown considerable propensity to combine any unilateral action on carbon emissions with some form of border adjustment. For example, as noted in Cosbey et al. (2012), the U.S. considered trade restrictions within climate legislation in 2009, and France also considered trade restrictions in the context of phase III of the EU Emissions Trading System. The EU allows for trade restrictions in the form of a (WTO compliant) "carbon equalization system" under Directive 2009/29/EC (see paragraph 25).

environmental policy. This is a departure from previous studies that focus on either fully cooperative or fully non-cooperative settings. Our results are reconciled with the prior literature, and provide salient insights for the current policy environment where trade is largely coordinated but environmental policy is not. We find an important general-equilibrium effect that indicates a divergence between a country's optimal domestic carbon price and its optimal pricing of carbon embodied in trade. The optimal border adjustment will not equal the domestic Pigouvian rate, even when the country's border adjustments are motivated purely by environmental concerns. The intuition is clear. While a carbon-based border tariff sends a price signal that discourages foreign emissions it also encourages foreign consumption of the more carbon intensive goods. The theory indicates that Pigouvian based border adjustments are likely to be too aggressive. Our empirical simulations support this finding, illustrating first-order differences between optimal tariffs and domestic prices. Specifically, the optimal import tariff on the carbon content of aluminum and other nonferrous metals is found to be 40% of the optimally set domestic carbon price.

Our results imply the current set of Pigouvian-based border adjustments being considered in the policy arena are sub-optimal (too aggressive) based on purely environmental concerns. And, because they are too aggressive from an environmental perspective, these adjustments fall outside of the environmental provisions granted in the WTO's General Agreement on Tariffs and Trade (GATT). As such, Pigouvian-based border adjustments are in jeopardy of being challenged on the grounds that they are, at least in part, de facto a beggar-thy-neighbor policy. Aggressive Pigouvian-based border adjustments that correct for different carbon prices across trade partners might, in fact, result in an indirect backpedaling from the Paris, COP 21, commitments on burden sharing to the extent that the border adjustments result in rent extraction from developing countries.

Our analytic approach focuses on the theory of optimal border carbon adjustments to inform consistent policy advice. We start from the two-good, two-country neoclassical general equilibrium theory of Markusen (1975) to establish the optimal unilateral domestic and trade instruments when facing a cross-border production externality in a non-cooperative trade setting (results which are echoed in Keen and Kotsogiannis (2014) and Hoel (1996)). The key is to refine the established theory by incorporating GATT consistency using the constraint proposed by Böhringer et al. (2014), which effectively eliminates any beggar-thy-neighbor incentives.[3] This constraint requires that trade partners be made no worse off by unilateral trade policies. With this constraint in place, we derive the optimal domestic and trade policies. Contrasting these optimal policies with Pigouvian-based border-adjustments shows that the Pigouvian prescription is too aggressive. Finally, to illustrate the magnitude of this wedge between the optimal and Pigouvian-based instruments, a numerical simulation of Annex-I carbon policy finds the optimal border tariff on the carbon content of aluminum and other nonferrous metals is substantially less than the domestic (Pigouvian) carbon price.

The tension between environmentally motivated border policies and the WTO's objective of cooperative trade is a topic of interest for both legal and economic scholars. While there have been attempts to reconcile carbon based tariffs as a tax adjustment under Articles II and III of the GATT (and Article XVI for carbon based export rebates), the general view is that carbon-based border policies would most easily be legitimized under the General Exceptions offered under Article XX. Cosbey et al. (2012), for example, argue that border carbon adjustments will violate the non-discriminatory provisions in the GATT because of differences in carbon intensities across regions.[4] Throughout the analysis in this paper we assume that border carbon adjustments would be implemented under an Article XX exception. In particular, a case can be made that border carbon adjustments are policy measures

[3]Böhringer et al. (2014) use their proposed constraint to decompose the environmental versus terms-of-trade incentives to impose different domestic carbon prices across sectors.

[4]In addition to the comprehensive look at prospects for border adjustments offered by Cosbey et al. (2012), there are several good reviews of legal issues related to border carbon adjustments. Tamiotti (2011), Pauwelyn (2013) and Horn and Mavroidis (2011) cover legal issues for carbon regulation in the US and/or Europe in general. van Asselt et al. (2009) focuses on the US Climate Security Act (Lieberman-Warner bill), whereas de Cendra (2006) focuses on the EU's Emissions Trading System.

covered under either paragraph (b): “necessary to protect human, animal or plant life or health,” or paragraph (g): “relating to the conservation of exhaustible natural resources if such measures are made effective in conjunction with restrictions on domestic production or consumption.”[5] While Article XX offers an opportunity to utilize border carbon adjustments as a compliment to subglobal action, its preamble clearly sets some limits. The policy measures cannot be “applied in a manner which would constitute a means of arbitrary or unjustifiable discrimination between countries” and cannot be a “disguised restriction on international trade.” In this context we argue that WTO consistent carbon adjustments should be limited to environmental objectives (as opposed to strategic rent-seeking objectives).

This study provides several important contributions. First, it extends the existing theoretical literature on optimal unilateral border policies in the presence of cross-border environmental damages to incorporate GATT consistency. Second, it shows that optimal border policy should be less aggressive than the Pigouvian prescription due to the general equilibrium response by consumers in unregulated regions. Third, our data-driven numerical simulations illustrate that there may be considerable differences between the optimal border tariff and the domestic carbon price. Finally, our study has important implications for policy. While there has been frequent policy advice to set border tariffs on embodied carbon based on the domestic carbon price (e.g. Barrett and Stavins (2003), Aldy and Stavins (2008), Cosbey et al. (2012), and Stiglitz (2013)), we show that countries following such advice will be extracting rents from unregulated regions at the expense of efficient environmental policy and consistency with international law.

We proceed with the paper as follows: Section 2 provides additional discussion of prior literature and sets the context for our theoretical and empirical analysis of border adjustments. Section 3 presents the economic theory of optimal border policy, in which we disentangle

[5] See https://www.wto.org/english/res_e/booksp_e/gatt_ai_e/art20_e.pdf for the full text of Article XX.

the strategic and environmental objectives. Section 4 presents a set of data driven numeric simulations that show the significance of our argument in the context of a model calibrated to data. Section 5 concludes.

## 2 Background

The formal theoretic literature on optimal environmental tariffs begins with Markusen (1975). Markusen establishes the optimal unilateral domestic and trade instruments when facing a cross-border production externality. Markusen illustrates the theory in a transparent two-good two-country neoclassical general equilibrium, and completes his analysis by considering a series of second-best responses. We choose to adopt Markusen's transparent model as the ideal setting in which we disentangle the strategic-trade and environmental incentives to distort trade. One useful feature of Markusen's setting is that it clearly highlights the role of relative international prices (the terms of trade) as a mechanism to signal foreign agents. A *small* country has neither a strategic nor an environmentally motivated incentive to distort trade because a lack of market power indicates an inability to affect foreign-agent behavior. Markusen's analysis is not specifically focused on carbon tariffs, but it is an essential starting point for any analysis of cross-border externalities.

An important theoretic examination of unilateral *carbon* policy is offered by Hoel (1996). Hoel's analysis achieves a set of conclusions on the first and second-best policy responses consistent with Markusen (1975) in the more general context of a model with any number of goods which may, or may not, be tradable. The central conclusion is that a country's carbon tax should be uniform across sectors if a set of trade distortions are available. Hoel's approach is slightly different than Markusen's, however, in that foreign carbon emissions are simply modeled as a function of net imports. The logic is clear that home-country imports change world prices and these world prices subsequently affect foreign emissions.

We emphasize the full chain, however, which includes the role of carbon tariffs in sending a price signal to foreign agents.[6] The theory established by Hoel is the foundation for much of the contemporary work on climate policy and carbon tariffs.

Both Hoel (1996) and Markusen (1975) establish, in a noncooperative trade setting, an optimal tariff which includes a strategic and additive environmental term, but the environmental term is inherently entwined with terms-of-trade adjustments. It is not clear, at least from our perspective, that the form of the environmental term will be preserved once we incorporate GATT consistency. Other examples of studies that focus on the general setting of non-cooperative trade with cross-border externalities include, Krutilla (1991), Ludema and Wooton (1994), Copeland (1996), and Jakob et al. (2013). Ludema and Wooton (1994) do consider the case of a cooperative trade restriction whereby a domestic environmental tax can be used to manipulate terms-of-trade in the absence of a tariff instrument. Copeland (1996) also shows that the rent shifting incentives to distort trade can be strengthened by foreign environmental regulation. To date, however, little effort has been focused on disentangling the unilateral environmental objective to distort trade, relative to beggar-thy-neighbor incentives.

Keen and Kotsogiannis (2014) offer a critical theoretic contribution by considering a setting of globally coordinated trade and environmental policy. This generalizes the partial-equilibrium analysis of Gros (2009) on the optimal instrument choice of an altruistic country seeking to maximize global welfare. The strategic incentive to distort trade is obviously neutralized in these settings. Keen and Kotsogiannis establish the conditions for globally Pareto efficient carbon tariffs in a general theoretic model that is familiar to trade economists. Under a set of constraints on internal policy (for a subset of countries) a set of carbon tariffs are

[6] Hoel (1996) argues (on page 25) that countries with little market power might still have significant carbon tariffs. His theory (consistent with Markusen (1975)) shows, however, that the optimal tariff must approach zero as international market power approaches zero. The distortion cannot be beneficial unless it changes foreign behavior.

prescribed to achieve efficiency. While the form of efficient border policy is generally complex, Keen and Kotsogiannis (2014) highlight a set of conditions under which the standard advice to set the carbon tariff at the difference between the home and foreign carbon tax is obtained. This special case backs out a set of restrictions on preferences and technologies such that the import adjustment just equals the difference between internal carbon taxes across the countries. Within our model we show that this restriction requires that consuming agents are unresponsive to price changes. We highlight the relaxation of the restrictive case offered by Keen and Kotsogiannis (2014) in both our analytical model and in our data driven numeric simulations.

Our analysis with cooperative trade and non-cooperative environmental policy might be cast as a special case of the fully cooperative model considered by Keen and Kotsogiannis (2014). In particular, we analyze one (particularly relevant) globally efficient allocation where the regulating country is maximizing welfare subject to holding welfare in the unregulated country fixed. This is a relevant allocation because it is consistent with the compensatory action that the unregulated country would be entitled to under international trade law. Keen and Kotsogiannis (2014), however, solve for the full set of efficient allocations under the restriction that some set of countries are not able to set their domestic environmental policies optimally. That is, Keen and Kotsogiannis (2014) consider the full constrained Pareto frontier. We consider a point on the constrained Pareto frontier that is consistent with cooperative trade (where trade policy cannot harm the unregulated country). Also adopting the fully cooperative setting Gros (2009), notably, comes to the same conclusion as us: the optimal border carbon adjustment is less than the optimally set domestic carbon price. Thus, we can place our analysis within the literature that looks at fully cooperative settings in that we generalize the partial equilibrium work of Gros (2009) and we look at a salient special case of Keen and Kotsogiannis (2014). In particular, our contribution relative to these studies is to reconcile the constrained Pareto allocation with the real world legal and political situation, and to show that in this situation the standard advice leads to border adjustments that are too aggressive.

The standard advice to establish a border tariff by applying the domestic carbon price to emissions embodied in imports, or equivalently requiring forfeiture of an emissions permit upon importing embodied carbon, is pervasive in the economic and policy literature. Examples of such advice include Stiglitz (2013), Cosbey et al. (2012), Aldy and Stavins (2008), and Barrett and Stavins (2003). Some authors consider such border policies as sanctions against non-participating countries [e.g., Böhringer et al. (2013b) and Aldy et al. (2001)]. There are two exceptions that reflect the theoretic results presented here that the optimal unilateral environmental border adjustment (under WTO) is below domestic carbon pricing: our earlier work (Yonezawa et al., 2012); and the recent paper by Böhringer et al. (2013a). In Böhringer et al. (2013a), a set of scenarios are considered in a Computable General Equilibrium model that approximate the optimal border adjustments. These are approximations because they use a set of reference scenarios to establish trade responses and do not explicitly include a valuation for the environment (which is endogenous to abatement). Our contribution is to clearly establish the theory for border adjustments free of strategic incentives, but in the policy relevant context of uncoordinated international environmental policy. Our numeric analysis is unique in demonstrating the operation of these border adjustments in a general equilibrium that includes environmental valuation in establishing the optimal adjustment.

We also consider, in our simulations, a *full border adjustment* policy. This policy advises that, in addition to imposing embodied-carbon tariffs, regulated countries would impose embodied-carbon subsidies on exports. That is, there would be a rebate of the accumulated value of carbon charges in the supply chain at the point of export. While these proposed policies appear in the literature and are often studied in numeric simulation, there is no clear theoretic justification for their adoption on efficiency grounds. Elliott et al. (2010) argue that in an open economy, full border adjustment effectively transforms a domestic *production* tax on carbon emissions into a *consumption* tax on embodied emissions (a result confirmed by Jakob et al. (2013) in a formal theoretical model), and that under some circumstances this may be desirable in terms of domestic welfare. Full border adjustment proposals also have some political economy advantages as they are favored by domestic producers of energy intensive goods, and consumption based policies might have broader normative or moral appeal. These are not, however, arguments that appeal to the efficiency properties.

In fact, Jakob et al. (2013) use a generalized version of Markusen's model to prove that full border adjustment is not optimal. They explain that optimal trade restrictions depend on the carbon-intensity differential between the foreign country's export and non-export sectors—not the differential between home and foreign export sectors (which is the tax basis for full border adjustments). Jakob et al. (2013) go on to show that full border adjustment can actually exacerbate carbon leakage, and that this is empirically relevant in the case of the EU imposing a set of import and export adjustments on China. In our simulations we find that, while applying carbon based export subsidies reduces the gap between the domestic Pigouvian tax and the trade adjustment, it does not eliminate the gap. Consistent with Jakob et al. (2013), full border adjustment based on the domestic carbon price is not optimal as a unilateral policy, and we extend this to demonstrate that it is not optimal even in the case of cooperative trade.[7]

# 3 Theory

In this section we first present the Markusen (1975) theory, indicating the additive nature of the environmental and strategic (rent seeking) components of a country's optimal tariff. We then introduce a constraint representing the GATT commitment, which effectively eliminates

[7]Apart from the discussion in the economic literature, full border adjustment could face international legal problems. The carbon rebate on exports could be viewed as a *per se* violation of GATT rules on export subsidies. Cosbey et al. (2012) argue that export adjustments are not recommended because they clash with trade laws and their administration is otherwise problematic.

the strategic term. This framework allows us to analyze national incentives to distort trade for purely environmental objectives. We derive a simple closed-form relationship between the optimal environmental tariff and the optimal domestic (Pigouvian) production tax. We build directly on Markusen (1975), but our theory would extend to the multi-commodity environments examined in Hoel (1996) and Keen and Kotsogiannis (2014). The key insights provided here include a theoretic foundation for the separability of the environmental and strategic components of commercial policy and the divergence of optimal domestic environmental taxes and the optimal border adjustment.

Consider a simple two-good two-country (North-South) trade model. Both countries, country $N$ and country $S$, produce and trade the goods $X$ and $Y$, and pollution is a function of the domestic and foreign production of good $X$. The pollution level, $Z$, is represented as follows:

$$Z = Z(X_N, X_S). \tag{1}$$

The efficient transformation function that determines a country's output of $X$ and $Y$ is given by:

$$F_r(X_r, Y_r) = 0 \qquad \text{or} \qquad Y_r = L_r(X_r), \qquad r \in \{N, S\}, \tag{2}$$

where $L_r(X_r)$ maps out the efficient frontier (PPF) in terms of $Y_r$ as a function of $X_r$. Letting $C_{iN}$ represent the consumption of good $i$ in country $N$, the welfare of the North is

$$U_N = U_N(C_{XN}, C_{YN}, Z). \tag{3}$$

We use $Y$ as a numeraire so that all prices are ratios in terms of $Y$. Let $q$, $p$, and $p^*$ denote the price ratio faced by consumers in the North, the price ratio faced by producers in the North, and world price ratio faced by consumers and producers in the South. The policy instruments considered are $\tau$, a tariff rate set by the North, and $t_X$, as the production tax rate in the North. Assuming no other distortions, the price relationships are

$$q = p(1 + t_X) = p^*(1 + \tau). \tag{4}$$

Pollution is not priced in the market equilibrium, but let us denote the marginal rate of substitution between pollution and good $Y$ as $q_Z = \frac{\partial U_N/\partial Z}{\partial U_N/\partial C_{YN}}$, where $q_Z$ is negative reflecting the negative impact of pollution on welfare.

Let $m_i$ indicate the North's net imports of good $i$. Thus the balance-of-payments constraint is given by

$$p^* m_X + m_Y = 0, \qquad m_X = C_{XN} - X_N, \qquad m_Y = C_{YN} - Y_N. \tag{5}$$

We are primarily interested in the case where the North imports the polluting good ($m_X > 0$) to inform current climate policy debates. The theory, however, generalizes to either trade pattern.[8]

As in Markusen (1975), we first consider the case where both environmental policy and trade policy are noncooperative. Given the trade equilibrium with the cross-border externality, and no other distortions, we can derive formulas for the North's optimal unilateral tariff $\tau$ and production tax $t_X$.

**Theorem 1.** *(Markusen, 1975). The optimal unilateral tariff and production tax in a noncooperative trade setting are given by:*

$$\begin{aligned} \tau &= \frac{m_X}{p^*}\frac{dp^*}{dm_X} - \frac{q_Z}{p^*}\frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*}\frac{dp^*}{dm_X}, \\ t_X &= -\frac{q_Z}{p}\frac{\partial Z}{\partial X_N}. \end{aligned} \tag{6}$$

[8] In the case that $m_X < 0$, where the North exports the polluting good, $\tau$ is interpreted as the North's export subsidy (or equivalently $-\tau$ is the export tax). Thus, the general pricing equation (4) is preserved in any case.

*Proof.* See Appendix A □

The optimal import tariff consists of the (non-environmental) strategic component as the first term and the environmental component as the second term. The optimal production tax is the Pigouvian rate because setting the production tax at $-\frac{q_Z}{p}\frac{\partial Z}{\partial X_N} = -\frac{1}{p}\frac{\partial U_N/\partial Z}{\partial U_N/\partial C_{YN}}\frac{\partial Z}{\partial X_N}$ represents pricing $X$ inclusive of the marginal environmental damage resulting from a unit of the North's production of $X$. Notice that in the absence of the environmental externality (where $\frac{\partial Z}{\partial X_N} = \frac{\partial Z}{\partial X_S} = 0$) the standard neo-classical trade result is obtained, where the domestic production tax is zero and the trade distortion is purely a strategic optimal tariff, $\tau = \frac{m_X}{p^*}\frac{dp^*}{dm_X}$ (which is the inverse of the North's import-supply elasticity).

While the above constitutes optimal policy in a noncooperative trade setting, the first component of the optimal tariff (the strategic term) is inconsistent and works against the principals of the GATT, as it exploits leverage over the terms-of-trade to extract rents from the South.[9] As such, we next determine the optimal policy in a cooperative trade setting, where such beggar-thy-neighbor strategic tariffs are not allowed. We thus modify the Markusen model by adding an endogenous lump-sum transfer that eliminates this strategic incentive to distort trade, per Böhringer et al. (2014).[10] The transfer payment $T$ is determined such that the South is not made worse off by trade policy implemented in the North. Let $\bar{U}_S$ be the measure of welfare in the South in the absence of tariffs and let $U_S = U_S(C_{XS}, C_{YS})$ equal the South's realized welfare.[11] A complementary slack condition

[9] Technically, countries only commit to not exceed their tariff bindings under the GATT (not necessarily a commitment to fully cooperative trade). The point here is that, even in the absence of environmental motives, any attempt by a country to exploit market power by exceeding tariff bindings would draw attention, and potential compensatory judgment, by the international courts.

[10] There are alternative ways to represent the constraints imposed by cooperative trade agreements, such as the potential for retaliatory tariffs. Our formulation of the endogenous lump-sum transfer, however, captures the purest (transparent) instrument which perfectly neutralizes the strategic trade incentives. Distortional retaliation available under WTO rules would have additional general equilibrium effects and therefore are not considered.

[11] Note that we only include private consumption in the South's utility function. This should not be read as an argument that the South does not value the environment. It is simply an assumption that the WTO-consistent compensatory action is restricted to lost private consumption.

is indicated that ensures GATT consistency of added trade distortions; where $U_S - \bar{U}_S \geq 0$ and $T \geq 0$, and $T(U_S - \bar{U}_S) = 0$. Under a set of border adjustments imposed by the North there is downward pressure on $U_S$ and we can be sure that the following holds:

$$U_S = \bar{U}_S; \qquad T > 0. \tag{7}$$

The balance-of-payments equation, (5), is modified as follows when we include the transfer, $T$:

$$p^* m_X + m_Y + T = 0. \tag{8}$$

We now consider the optimal policy as chosen in the North when environmental policy is noncooperative, but trade policy is subject to cooperative trade agreements. Given these modifications, the North sets its tariff $\tau$ and production tax $t_X$ unilaterally as before, but accounting for the fact that losses in the South's welfare require compensation via the endogenous transfer:

**Proposition 1.** *The optimal unilateral tariff and production tax in a cooperative trade setting are given by:*

$$\begin{aligned} \tau &= -\frac{q_Z}{p^*}\frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*}\frac{dp^*}{dm_X}, \\ t_X &= -\frac{q_Z}{p}\frac{\partial Z}{\partial X_N}. \end{aligned} \tag{9}$$

*Proof.* See Appendix B □

Comparing Proposition 1 with Theorem 1, the addition of the transfer has effectively eliminated the strategic component in the optimal tariff in a cooperative trade setting. While this isolates the environmental component of the optimal tariff, nonetheless it is clear that the optimal tariff is not simply equal to the Pigouvian rate and critically depends on the

North's ability to affect international prices with its tariff. That is, if $dp^*/dm_X = 0$ the optimal environmental tariff is zero. A *small* country cannot send a price signal to foreign agents through a tariff and optimally chooses free trade.

We next consider how the optimal tariff derived above compares with the production tax rate, which is optimally set at the Pigouvian rate. Let $\theta_N \equiv -\frac{q_Z}{p}\frac{\partial Z}{\partial X_N}$ represent the Pigouvian rate (marginal external damage) for production in the North, and $\theta_S \equiv -\frac{q_Z}{p^*}\frac{\partial Z}{\partial X_S}$ represent the Pigouvian rate for production in the South. If it were allowed, $\theta_S$ is the rate at which the North would like to directly regulate production in the South.

**Proposition 2.** *In a cooperative trade setting, the optimal tariff* $\tau$ *is:*

*i) less than the Pigouvian rate for production in the South, such that* $\tau < \theta_S$*;*

*ii) less than the production tax rate in the North,* $\tau < t_X$*, if emissions per unit of output are the same,* $\frac{\partial Z}{\partial X_N} = \frac{\partial Z}{\partial X_S}$*;*

*iii) greater than the production tax rate in the North,* $\tau > t_X$*, if and only if* $\frac{\partial Z/\partial X_N}{\partial Z/\partial X_S} < \frac{dX_S}{dp^*}\frac{dp^*}{dm_X}$.

*Proof. i)* From (9) the optimal tariff is

$$\tau = \theta_S \frac{dX_S}{dp^*}\frac{dp^*}{dm_X}, \tag{10}$$

where $m_X$ is the North imports of good $X$. In order to prove that the optimal tariff is less than the Pigouvian rate, we derive the following equation from the supply and demand relationship (analogous to (5)) in the South ($X_S = C_{SX} + m_X$):

$$\frac{dX_S}{dp^*}\frac{dp^*}{dm_X} = \frac{dC_{SX}}{dp^*}\frac{dp^*}{dm_X} + \frac{dm_X}{dp^*}\frac{dp^*}{dm_X}. \tag{11}$$

The left-hand term is positive, given convexity of the production set and the fact that $\frac{dp^*}{dm_X}$

is positive.[12] The last term on the right-hand side is equal to unity, and the term $\frac{dC_{SX}}{dp^*}$ must be negative under (7) as consumers in the South will substitute away from the more expensive good, noting that under (7) we only have a substitution effect for the South.[13] Taken together signing the elements of (11) gives

$$\frac{dX_S}{dp^*}\frac{dp^*}{dm_X} < 1. \tag{12}$$

Thus, $\tau < \theta_S$.

*ii)* From (9), $t_X = \theta_N$. If the environmental damage associated with producing the good $X$ is the same in the North and the South ($\partial Z/\partial X_N = \partial Z/\partial X_S$) local to the optimal, then $\theta_S = \theta_N \frac{p}{p^*} = \theta_N \frac{1+\tau}{1+t_X}$ and thus $\frac{t_X}{1+t_X} = \frac{\theta_S}{1+\tau}$. Given $\tau < \theta_S$ from *i)* it follows that $\frac{t_X}{1+t_X} = \frac{\theta_S}{1+\tau} > \frac{\tau}{1+\tau}$. Therefore we must have $t_X > \tau$ given that $1 + t_X$ and $1 + \tau$ are non-negative price wedges.

*iii)* First, we show that if $\tau > t_X$, then $\frac{\partial Z/\partial X_N}{\partial Z/\partial X_S} < \frac{dX_S}{dp^*}\frac{dp^*}{dm_X}$. Note (9) can be rewritten as:

$$\begin{aligned} \frac{\tau}{1+\tau} &= -\frac{q_Z}{q}\frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*}\frac{dp^*}{dm_X}, \\ \frac{t_X}{1+t_X} &= -\frac{q_Z}{q}\frac{\partial Z}{\partial X_N}. \end{aligned} \tag{13}$$

Solving the above for $\tau$ and $t_X$ respectively, and assuming $\tau > t_X$ yields $\frac{\partial Z/\partial X_N}{\partial Z/\partial X_S} < \frac{dX_S}{dp^*}\frac{dp^*}{dm_X}$.

Next, we show that if $\frac{\partial Z/\partial X_N}{\partial Z/\partial X_S} < \frac{dX_S}{dp^*}\frac{dp^*}{dm_X}$, then $\tau > t_X$. First, if $\frac{\partial Z/\partial X_N}{\partial Z/\partial X_S} < \frac{dX_S}{dp^*}\frac{dp^*}{dm_X}$, then $-q_Z\frac{\partial Z}{\partial X_N} < -q_Z\frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*}\frac{dp^*}{dm_X}$, which from (9) implies $p^*\tau > pt_X$. This gives $\tau > \frac{p}{p^*}t_X = \frac{1+\tau}{1+t_X}t_X$, which simplifies to $\tau > t_X$. □

To understand the first result, note that although the Pigouvian rate $\theta_S$ reflects the marginal environmental damage of production in the South, it is adjusted in the optimal

[12] An increase in North imports ($m_X$) drives up the international price ($p^*$).

[13] If the South were not compensated the sign of $\frac{dC_{SX}}{dp^*}$ is ambiguous, given the possibility of being on a *backward-bending* portion of the offer curve.

tariff by two terms: (1) the ability of the North to influence prices in the South through changing import volumes ($\frac{dp^*}{dm_X}$), and (2) the impact of that price change on production in the South ($\frac{dX_S}{dp^*}$). The tariff decreases the price faced by producers in the South, and production of $X$ is discouraged in the South. The lower price also encourages consumption of $X$ in the South. Thus, the decrease in environmental damage from decreased imports is partially offset by the increase in consumption in the South. Intuitively, $\tau$ is an imperfect instrument for influencing production in the South because the price change is limited by the negative $\frac{dC_{SX}}{dp^*}$ term. This is the unintended consumption effect of the environmental tariff. Consumption of the polluting good is encouraged in the South making the optimal tariff less than the Pigouvian rate that the North would like to impose on production in the South. Notice that, in our model, to arrive at the restrictive case highlighted by Keen and Kotsogiannis (2014), where the optimal environmental tariff has the simple structure envisaged in the policy debate ($\tau = \theta_S$) one would need consuming agents in the South to be completely unresponsive to price changes ($\frac{dC_{SX}}{dp^*}$=0). This restriction is not easily defended, and as such we maintain our assumption of strictly negative substitution effects throughout the analysis in this paper.

The second result shows that if the environmental damage associated with producing good $X$ is the same between the North and the South ($\partial Z/\partial X_N = \partial Z/\partial X_S$), then the optimal tariff is always less than the optimal production tax. It is possible, however, that the marginal environmental damage per unit of production may be higher in the South. Nonetheless, the third result shows that in order for the tariff to exceed the production tax rate, a large difference in marginal damages from production is required to offset the general equilibrium effect on the South's consumption response. For example, if $\frac{dX_S}{dp^*}\frac{dp^*}{dm_X} = 0.5$, such that a one unit decrease in imports leads to a 0.5 unit decrease in production in the South, then the marginal damage from production in the South would need to be more than double that in the North for $\tau > t_X$. Taken together, the above results indicate that it is unlikely that following the typical advice to set $\tau = t_X$ is optimal.

Empirically, equation (10) provides some insight into determining which commodities potentially have large differences between optimal and Pigouvian tariff rates. If $\frac{dp^*}{dm_X}$ is small, the optimal tariff becomes small and the gap with the Pigouvian rate becomes large. In other words, if changes in imports do not affect world prices significantly the price signal to foreign agents is weak, and the optimal tariff is close to zero. The amount of imports relative to world production (import share) can indicate whether $\frac{dp^*}{dm_X}$ is small or large. For example, if the imports are a small share of the world market, it is likely that changing the import amount will not substantially affect world prices.

Also from (10) we see that if $\frac{dX_S}{dp^*}$ is small the optimal tariff becomes small. In this case, if the world price change does not affect production in non-regulated regions significantly, the optimal tariff is close to zero. The key responses come from both the consumption and the production sides of the foreign economy. In the case that consumers in the South are very responsive to price (high elasticities of substitution) the more negative is $\frac{dC_{SX}}{dp^*}$, the smaller is the optimal tariff. On the production side, if production is relatively insensitive to the price changes (low elasticities of transformation) the smaller is the optimal tariff.

In our final extension of the Markusen (1975) theory we consider taxes on pollution and embodied pollution tariffs.[14] In the previous discussion, optimal policies are derived in terms of an ad valorem tax or tariff on *production* ($X$), while carbon policies under consideration

[14]Copeland (1996) makes a similar extension to the theory to look at strategic motives to extract international rents through environmental policy. The pollution-content tariff introduced by Copeland (1996), however, is slightly different in that it allows for a direct identification of the exporting firm's emissions on the units exported. The tariff varies with the amount of pollution during the production of the *traded* output. This sets up an incentive for firms to use different processes for domestic versus export markets, and gives Copeland a relatively *sharp* policy instrument to target the crossborder externality. In contrast, we assume the tariff is based on the average emissions rate for the foreign industry as a whole, which is probably more realistic from an administrative perspective. Even industry-wide measures are ambitious in the context of carbon emissions. With carbon, indirect emissions associated with intermediate non-fossil inputs—like electricity—are important. See Cosbey et al. (2012) for a discussion of the practical challenges of setting up embodied carbon tariffs, and Böhringer et al. (2013a) for technical details on how one might use (imperfect) input-output techniques for calculating the full carbon content by good and country.

are typically framed as a specific unit tax or tariff on *pollution* ($Z$). To explore this nuance, suppose the tax and tariff is levied on pollution, such that the price relationships in equation (4) are now:

$$q = p + \tilde{t}_X \frac{\partial Z}{\partial X_N} = p^* + \tilde{\tau} \frac{\partial Z}{\partial X_S}. \tag{14}$$

Notice that the instruments are now specific taxes (i.e., dollars per ton) on the marginal pollution content of the transaction.

**Proposition 3.** *When the tax and tariff are levied on pollution $Z$,*

*i) The optimal unilateral tariff and production tax in a noncooperative trade setting are given by:*

$$\begin{aligned} \tilde{\tau} &= \frac{m_X}{\partial Z / \partial X_S} \frac{dp^*}{dm_X} - q_Z \frac{dX_S}{dp^*} \frac{dp^*}{dm_X}, \\ \tilde{t}_X &= -q_Z. \end{aligned} \tag{15}$$

*ii) The optimal unilateral tariff and production tax in a cooperative trade setting are given by:*

$$\begin{aligned} \tilde{\tau} &= -q_Z \frac{dX_S}{dp^*} \frac{dp^*}{dm_X}, \\ \tilde{t}_X &= -q_Z. \end{aligned} \tag{16}$$

*iii) The optimal tariff on embodied emissions in a cooperative setting is strictly less than the Pigouvian tax, $\tilde{\tau} < \tilde{t}_X$.*

*Proof.* The proofs of *i)* and *ii)* follow from the new price wedges implied by (14): $(q - p) =$

$\tilde{t}_X \frac{\partial Z}{\partial X_N}$ and $(q - p^*) = \tilde{\tau} \frac{\partial Z}{\partial X_S}$. Inserting these wedges, equation (26) in Appendix A becomes

$$\begin{aligned} \frac{dU_N}{\partial U_N / \partial C_{YN}} &= \left[ \tilde{\tau} \frac{\partial Z}{\partial X_S} - m_X \frac{dp^*}{dm_X} + q_Z \frac{\partial Z}{\partial X_S} \frac{dX_S}{dp^*} \frac{dp^*}{dm_X} \right] dm_X \\ &+ \left[ \tilde{t}_X \frac{\partial Z}{\partial X_N} + q_Z \frac{\partial Z}{\partial X_N} \right] dX_N, \end{aligned} \tag{17}$$

which yields the formulas in *i)* at the optimal. By the same substitution, under the GATT constraint equation (35) in Appendix B becomes

$$\begin{aligned} \frac{dU_N}{\partial U_N / \partial C_{YN}} &= \left[ \tilde{\tau} \frac{\partial Z}{\partial X_S} + q_Z \frac{\partial Z}{\partial X_S} \frac{X_S}{dp^*} \frac{dp^*}{m_X} \right] dm_X \\ &+ \left[ \tilde{t}_X \frac{\partial Z}{\partial X_N} + q_Z \frac{\partial Z}{\partial X_N} \right] dX_N, \end{aligned} \tag{18}$$

and we have *ii)*. The proof of *iii)* then follows from the fact that $\tilde{\tau} = \tilde{t}_X \frac{dX_S}{dp^*} \frac{dp^*}{dm_X}$ under cooperative trade, and the proof of *i)* in Proposition 2 (where we show that $\frac{dX_S}{dp^*} \frac{dp^*}{dm_X} < 1$). □

The first result shows that although pollution is not explicitly traded, nonetheless the North's optimal tariff contains a strategic component in a noncooperative setting. Turning to the cooperative trade setting, with the transfer in place, the strategic component is once again eliminated in the second result. When the tariff and tax are placed on units of pollution, the tariff simplifies to $\tilde{\tau} = -q_Z \frac{dX_S}{dp^*} \frac{dp^*}{m_X}$ and the production tax is simply the specific Pigouvian tax on units of pollution $\tilde{t}_X = -q_Z$. Because the tariff again lowers the world price of $X$ and encourages consumption in the South, the optimal tariff rate levied on pollution is strictly less than the Pigouvian rate. As before, if changes in imports do not affect world prices significantly, or if the South's production of the polluting good is relatively unresponsive to changes in world price, the optimal tariff may be quite small. Taken together, these indicators of smaller optimal tariffs imply a larger gap between the optimal domestic carbon price and the optimal trade adjustments. In the following section we explore the size of this gap, and illustrate its significance in a model calibrated to data.

# 4 Optimal border adjustments on Nonferrous Metals

In this section, we use a specific, data driven, illustration of the potential difference between the optimal domestic carbon price and the trade adjustment. The context for the illustration is Annex-I subglobal carbon abatement, where there is an option to impose border adjustments on trade in aluminum and other nonferrous metals. Nonferrous metals are a good choice for the empirical experiment because of their energy and trade intensity.[15] These characteristics make nonferrous metals a likely target of border carbon adjustments. Focusing on nonferrous metals also provides a relatively clean experimental setting for our illustration. As a sensitivity case we include all energy intensive goods (iron, steel, chemicals, rubber, plastic, and other nonmetallic mineral products) in the coverage of border adjustments. In this case, our conclusion that the optimal environmental border adjustment is well below the Pigouvian rate is maintained.[16]

We first describe the model and calibration. Next, we calculate and compare the optimal tariff and domestic price in noncooperative and cooperative (GATT consistent) trade settings. We also consider the proposed so-called *full border adjustment*, where an export rebate is placed on exported embodied carbon in addition to the import tariff placed on imported embodied carbon. We conclude with sensitivity analysis that links the simulations back to the basic lessons from the theory.

[15] As noted in Cosbey et al. (2012), primary aluminum is identified as an energy-intensive, trade-exposed industry. A set of full results focused exclusively on aluminum as a subcategory appear in Yonezawa's thesis [Yonezawa (2012), Chapter 4].

[16] We also explored experiments with a broader coverage on non-energy intensive goods. In these cases the optimal environmental border adjustment was zero for most parameter settings. While in the simple theory presented above we can be sure that the marginal environmental benefit of a *small* tariff exceeds the international compensation costs (at the reference case of a Pigouvian domestic policy), this will not necessarily be the case in the data-driven simulation model.

## 4.1 Model and calibration

Our numeric model is a multi-commodity multi-region static general-equilibrium representation of the global economy with detailed carbon accounting.[17] We adopt the structure employed by Rutherford (2010) in his examination of carbon tariffs. We also follow Rutherford (2010) and Böhringer et al. (2013a) in calculating carbon embodied in trade using the multi-region input-output (MRIO) technique. For every trade flow, a carbon coefficient is calculated that includes the direct and indirect carbon content, as well as the carbon associated with transport.[18]

We augment the Rutherford (2010) model to include an explicit representation of environmental valuation. We include a preference for the environment (disutility from global emissions) in the Annex-I expenditure system. We use a simple formulation that assumes environmental quality is separable from consumption with a constant elasticity of substitution between environmental quality and private consumption of 0.5.[19] We calibrate Annex-I environmental preference to be roughly consistent with contemporary proposals on climate policy. The model is used to compute a carbon cap that yields a carbon price of $35 per ton of $CO_2$ in the Annex-I region (approximately an 80% cap relative to business as usual). With this reference equilibrium established we recalibrate the Annex-I expenditure function such that this is the money-metric marginal utility of (separable) emissions abatement. Therefore, in the calibrated reference case, the Annex-I region is pursuing optimal unilateral abatement with $35 per ton emissions pricing, conditional on no border adjustments. With targeted

[17] There is an extensive literature utilizing similar numeric simulation models to analyze border carbon adjustments and climate policy more generally. A recent special issue of *Energy Economics* was specifically focused on border carbon adjustments. This issue included 12 papers from different teams studying different aspects of border adjustments. An overview of the special issue and a set of model comparison exercises is provided by Böhringer et al. (2012).

[18] When calculating the carbon content of Annex-I exports for the case of full border adjustments below, we do not include the carbon associated with transport. It is the carbon content at the border that is of interest. Embodied imported carbon is gross of transport carbon, whereas embodied export carbon is net.

[19] Non-separabilities could be important in the context of climate change as emphasized by Carbone and Smith (2013), but this consideration is beyond the theory we illustrate.

Table 1: Scope of the Empirical Model

| Regions: | | Goods: | | Factors: | |
|---|---|---|---|---|---|
| Annex-I | Annex I (except Russia) | OIL | Refined oil products | LAB | Labor |
| MIC | Middle-High Income, n.e.c. | GAS | Natural Gas | CAP | Capital |
| LIC | Low Income Countries, n.e.c. | ELE | Electricity | RES | Natural Resources |
| | | COL | Coal | | |
| | | CRU | Crude Oil | | |
| | | ALU | Aluminum | | |
| | | NFM | Other Nonferrous metals | | |
| | | EIT | Energy Intensive, n.e.c. | | |
| | | TRN | Transportation | | |
| | | AOG | All other goods | | |

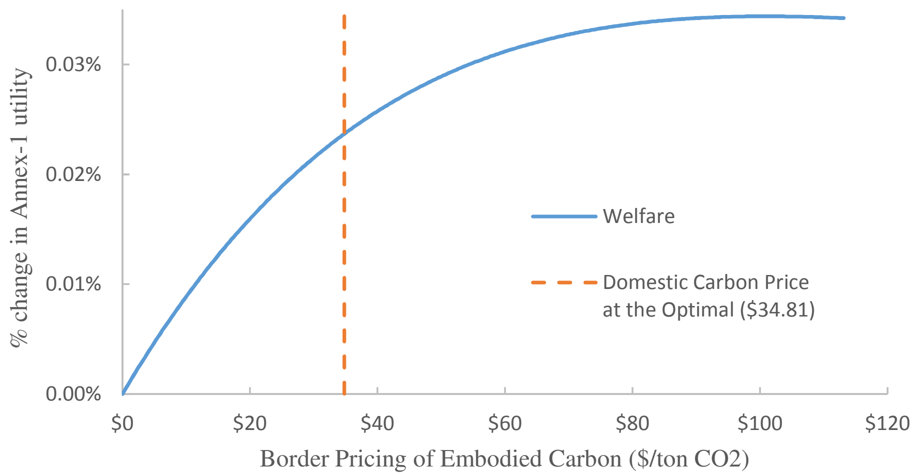
border adjustments, Annex-I can improve its welfare, because, on the margin, emissions reductions achieved through border adjustments on nonferrous metals are less costly than domestic abatement.

We also modify the Rutherford (2010) model to include the Böhringer et al. (2014) complementary slack condition, which under border adjustments is given by equation (7). This eliminates the strategic incentive for the Annex-I coalition to extract rents from other regions. In this context carbon-based border adjustments are only used to achieve the environmental objective, per the preceding theory.

To calibrate the model we use GTAP 7.1 data (Narayanan and Walmsley, 2008), which represents global production and trade with 113 countries/regions, 57 commodities, and five factors of production. For our purpose, we aggregate the data into three regions, nine commodities (one of which is nonferrous metals), and three factors of production. To explore targeted border adjustments on aluminum we split out the primary and secondary aluminum industry from the nonferrous metals accounts using data from Allen (2010) and the United States Geological Survey report on aluminum (Bray, 2010).[20] Table 1 summarizes the aggregate regions, commodities, and factors of production represented in the model. Annex-I

[20] A full description of the augmentation to the GTAP data to include aluminum (and the computer code used) is offered in Yonezawa (2012).

Figure 1: Welfare Responses to Border Adjustments with No GATT Constraint

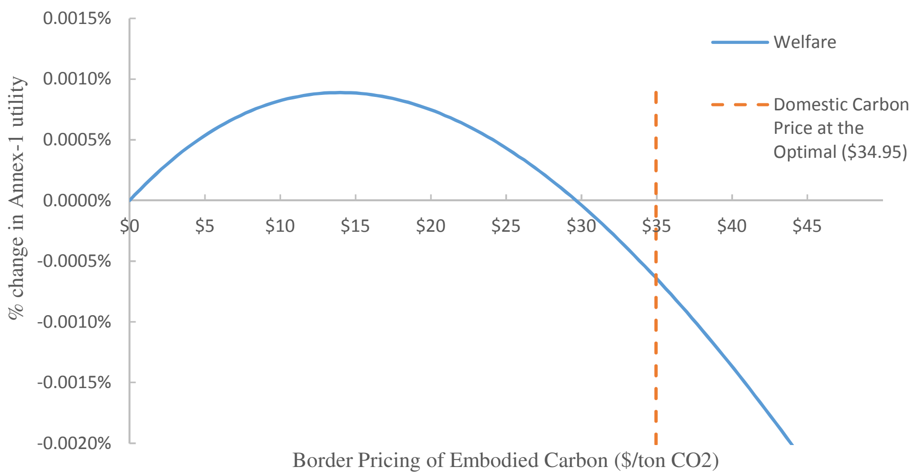



parties to the United Nations Framework Convention on Climate Change (UNFCCC) except Russia are aggregated as carbon-regulated regions. The rest of the world is divided into two aggregate regions according to World Bank income classifications.

## 4.2 Optimal carbon tariffs

We begin by first considering the optimal border adjustment in a noncooperative trade setting, which shows that the Annex-I coalition has a relatively large incentive to impose tariffs on aluminum and nonferrous metal imports. In this noncooperative setting, Annex-I countries are motivated by both strategic and environmental objectives, and the optimal pricing of embodied carbon associated with imports is $101 per ton $CO_2$ as illustrated in Figure 1. This is nearly three times the domestic carbon price. Translating the $101 per ton embodied carbon price into an ad valorem tariff equivalent results in a 31% tariff on `MIC` aluminum imports and a 44% tariff on `LIC` aluminum imports. The ad valorem rates are lower on other nonferrous metals (23% for `MIC` imports and 33% for `LIC` imports). The differences in these rates across products and trade partners reflect different carbon intensities.

Figure 2: Welfare Responses to Border Adjustments with GATT Constraint

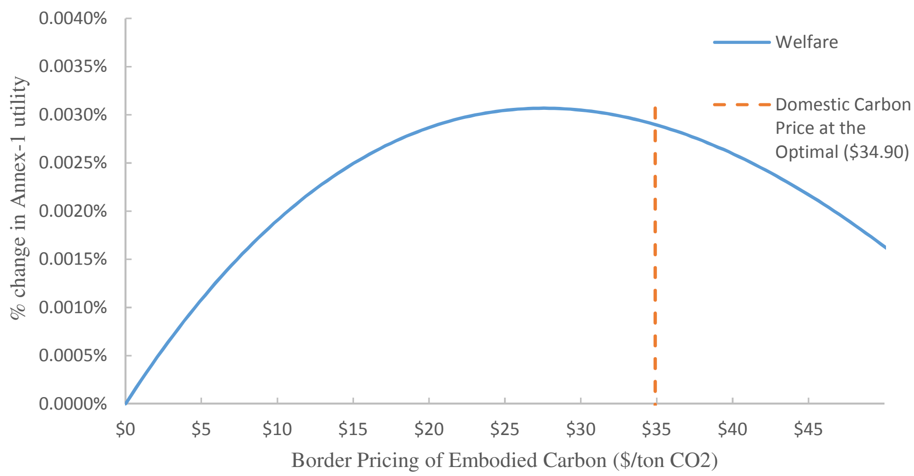



With the optimal unconstrained policy established, we now consider a comparison of embodied-carbon pricing and the domestic carbon price when the border objective is purely environmental. With the GATT constraint imposed, Figure 2 shows that the optimal trade distortion drops dramatically to $14 per ton. This is less than half of the domestic carbon price at the optimal. As such, following the standard prescription of imposing the domestic carbon price on embodied carbon imports indicates that over half of the trade distortion is a hidden beggar-thy-neighbour policy. At $14 per ton of $CO_2$, the ad valorem equivalents are modest: 4% on aluminum from `MIC`, 6% on aluminum from `LIC`, 3% on other nonferrous metals from MIC, and 5% on other nonferrous metals from `LIC`. Thus, in these relatively transparent numeric simulations, we find substantially lower optimal border adjustments, on the order of 60% lower than the domestic price.

## 4.3 Full border adjustment

We now consider the proposal of *full border adjustments*. In Figure 3 we plot Annex-I welfare as a function of the carbon price imposed on imports, as well as exports, of aluminum and

Figure 3: Full Border Carbon Adjustment with GATT Constraint




other nonferrous metals (full border adjustment). Two results are of note. First, optimal carbon pricing of trade is much closer to the domestic carbon price. The optimal pricing on embodied carbon in trade is $28 per ton, which is about 80% of the domestic carbon price. As highlighted by Yonezawa et al. (2012), a version of Lerner's symmetry (Lerner, 1936) applies, in that import tariffs are offset by export subsidies. In this sense, a higher overall pricing of carbon on imports is optimal as long as there is a counteracting export subsidy. Second, comparing Figure 2 with Figure 3, optimal welfare in Annex-I is higher under full border adjustments relative to an import-only policy. This reflects the cost savings due to driving world nonferrous metal consumption toward relatively low emissions intensive sources.[21]

The above simulations reinforce the findings of our theoretical analysis that the optimal border adjustment on carbon is less than the domestic carbon price under a GATT constraint. Furthermore, the simulations show that this difference may be of first-order importance, such

[21] Aluminum and other nonferrous metals produced in Annex-I countries have a relatively lower carbon intensity (reflected in the embodied carbon coefficients calculated using the MRIO method), and thus Annex-I can improve welfare through export subsidies which displace high carbon intensive aluminum in other countries.

Table 2: Optimal Ad Valorem Tariffs and Subsidies on Aluminum and Nonferrous Metals

| Trade Partner | Import Tariff | Export Subsidy | Embodied $CO_2$ Price ($\tilde{\tau}$) | Domestic $CO_2$ Price ($\tilde{t}_X$) | **Ratio:** $\tilde{\tau}/\tilde{t}_X$ |
|---|---|---|---|---|---|
| | **GATT Constrained** | | | | |
| ALU: Aluminum | | | | | |
| MIC | 4.3% | | \$13.98 | \$34.95 | **0.40** |
| LIC | 6.0% | | \$13.98 | \$34.95 | **0.40** |
| NFM: Other Nonferrous Metals | | | | | |
| MIC | 3.2% | | \$13.98 | \$34.95 | **0.40** |
| LIC | 4.6% | | \$13.98 | \$34.95 | **0.40** |
| | **Not GATT Constrained** | | | | |
| ALU: Aluminum | | | | | |
| MIC | 31.2% | | \$100.95 | \$34.81 | **2.90** |
| LIC | 43.5% | | \$100.95 | \$34.81 | **2.90** |
| NFM: Other Nonferrous Metals | | | | | |
| MIC | 23.4% | | \$100.95 | \$34.81 | **2.90** |
| LIC | 33.4% | | \$100.95 | \$34.81 | **2.90** |
| | **GATT Constrained: Full Border Adjustment** | | | | |
| ALU: Aluminum | | | | | |
| MIC | 8.5% | 4.2% | \$27.57 | \$34.90 | **0.79** |
| LIC | 11.9% | 4.2% | \$27.57 | \$34.90 | **0.79** |
| NFM: Other Nonferrous Metals | | | | | |
| MIC | 6.4% | 3.2% | \$27.57 | \$34.90 | **0.79** |
| LIC | 9.2% | 3.2% | \$27.57 | \$34.90 | **0.79** |

that border adjustments set at the domestic price may be substantially excessive relative to the optimal. Table 2 summarizes the above results for the three scenarios considered. The final column reports the ratio of the optimal embodied $CO_2$ price relative to the domestic carbon price at the optimal. An alternative, but equivalent, interpretation of our analysis is that it would be optimal to reduce the *amount* of embodied carbon on each trade flow according to the ratio in the final column of Table 2 if the embodied carbon price were equal to the domestic price. That is, the specific tariff is simply the product of the applied carbon price and the carbon coefficient so there are any number of combinations that can result in the optimal. Our point is that the optimal specific tariff is substantially below the standard advice to apply the full carbon price on measured embodied carbon.

### 4.4 Sensitivity analysis

We conclude our numeric simulations with a set of model runs that draw the applied model back to the theory. We focus on piecemeal parametric changes that impact the important determinants of the optimal tariff in the formulas derived in Section 3. First, Propositions 2 and 3 show that the optimal tariff is increasing in market power. We adjust the trade elasticities in the model to illustrate this effect. Second, the optimal tariff is decreasing in the foreign consumption response. We alter the elasticity of substitution between the focus goods (aluminum and other nonferrous metals) and other goods to illustrate this effect. Third, the optimal tariff is increasing in the foreign production response. We alter the elasticity of substitution between energy and other inputs, and the elasticity of substitution between sector-specific energy resources and other inputs, to illustrate this effect. Finally, we change the coverage of the tariffs relative to our central case. We decrease the coverage to only include aluminum, and increase the coverage to include all energy intensive imports. Table 3 shows the impact on the ratio of the optimal embodied carbon tariff and optimal domestic carbon pricing across these sensitivity runs.

Table 3: Sensitivity Analysis on Optimal Border Carbon Pricing Relative to Domestic

| | **Settings** | | | **Ratio:** $\tilde{\tau}/\tilde{t}_X$ | | |
|---|---|---|---|---|---|---|
| | low | central | high | low | central | high |
| Armington Substitution Multiplier | 0.5 | 1.0 | 2.0 | 0.55 | 0.40 | 0.23 |
| Materials Substitution Elasticity | 0.0 | 0.5 | 1.0 | 0.44 | 0.40 | 0.36 |
| Energy Substitution Elasticity | 0.05 | 0.5 | 5.0 | 0.39 | 0.40 | 0.43 |
| Resource Substitution Multiplier | 0.5 | 1.0 | 2.0 | 0.37 | 0.40 | 0.43 |
| Import Coverage | `ALU` | `ALU+NFM` | `ALU+NFM+EIT` | 0.55 | 0.40 | 0.59 |

The trade structure in our model is based on the standard formulation of differentiated regional goods (the Armington assumption). Under this structure each region's absorption is in a nested constant-elasticity-of-substitution composite of imported and domestically produced output. The trade responses are controlled through the assumed elasticities. In the central cases we use the elasticities as provided by GTAP, and their weighted averages for aggregates. In the first row of Table 3 we scale all of these elasticities for the non-regulated regions down by 50% (low) and then up by 100% (high). As these trade elasticities are scaled down, the Annex-I region gains market power, because the other regions are not as easily able to substitute out of Annex-I exports. As expected, the optimal environmental tariff falls with higher elasticities. When the elasticities are doubled, the ratio of the optimal embodied-carbon tariff drops to 23% of domestic carbon pricing.

In the second row of Table 3 we change the demand response in the middle income and low income countries by increasing the elasticity of substitution between intermediate materials. In the production functions, adopted from Rutherford (2010), the composite of non-energy and non-value-added inputs substitute at the top level for *materials*. In our case, materials include aluminum (`ALU`), other nonferrous metals (`NFM`), other energy intensive goods (`EIT`), and all other goods (`AOG`). The central elasticity of substitution between materials and the composite of energy and value-added inputs is 0.5. To explore the model's sensitivity to this parameter we scale it down to Leontief (0.0) and up to Cobb-Douglas (1.0) in the non-regulated regions. As predicted by the theory, the more responsive is the foreign demand, the lower is the optimal environmental tariff. This is the key general equilibrium effect that we highlight in this paper. Environmental tariffs, while discouraging foreign production of the dirty good, inevitably encourage foreign consumption of the dirty good. In the numeric simulations, agents in the middle and low income countries react to the tariffs by intensifying their own use of aluminum and other nonferrous metals. As we increase the elasticity of substitution for materials, this reaction is larger and the resulting optimal Annex-I environmental tariff is smaller.

In the third and fourth rows of Table 3 we consider the foreign production response. We expect higher optimal Annex-I tariffs the easier it is for non-regulated regions to substitute out of energy intensive production. We manipulate two different elasticities to capture this response. First, we scale the elasticity of substitution between energy and value-added inputs (row 3 of Table 3). We show that higher elasticities indicate higher Annex-I optimal environmental tariffs, but noticeable responses require large changes in this elasticity, likely due to the fact that this is an indirect method of manipulating the production response. In the central case the energy elasticity is 0.5, and we consider a low value of 0.05 and a high value of 5.0. Even at an elasticity of 5.0 (making energy a close substitute for value-added in the non-regulated regions) the optimal environmental tariff only rises to 43% of the domestic tax relative to 40% in the central case. For nonferrous metals, changing the energy substitution elasticity often has to work through primary fuels used in electricity generation and then downstream to electricity used in smelting (the most energy intensive stage of production). This reinforces a robust finding in the literature (see Böhringer et al. (2012)) that carbon tariffs are a blunt instrument for affecting foreign energy intensity.

To explore the foreign response of energy-intensive production from a different angle, in row 4 of Table 3, we manipulate the elasticity of substitution between the sector-specific resource in primary energy (`COL`, `GAS`, and `CRU`) and other inputs. In our model, following

Rutherford (2010), this elasticity of substitution is calibrated to yield specific, local, supply-elasticity targets in the central case ($\eta_{COL} = 1$, $\eta_{GAS} = 0.5$ and $\eta_{CRU} = 0.5$, where $\eta_i$ is the local price elasticity of supply). We scale the elasticity of substitution down by 50% and up by 100%. This has a direct impact on quantity responses for fuel production in non-regulated regions. As the theory predicts, greater response indicates higher optimal environmental tariffs.

In our final set of model runs we consider decreasing the embodied tariff coverage to only aluminum, and then increasing the coverage to include all energy intensive sectors (`ALU`, `NFM`, and `EIT`). In the case of just aluminum, the ratio of the optimal carbon tariff to the domestic tax rises to 55%, and when broadening the coverage to all energy intensive goods the ratio rises even further to 59%. Given that these sectors have a number of data-driven differences in the simulation model, it is difficult to obtain a clear prediction from the theory. The Annex-I global share of consumption is increasing as we increase the coverage, indicating higher optimal environmental tariffs, but aluminum production and consumption is more concentrated in the `LIC` region, also indicating more effective environmental tariffs. Overall, the results are consistent with our central argument that the GATT consistent environmental border adjustment is below the domestic carbon price across a broad range of energy-intensive products.

# 5 Conclusion

In this paper, we consider optimal border adjustments in a setting with noncooperative environmental policy, but cooperative trade policy. Following Markusen (1975) we establish optimal border policy in the presence of cross-border environmental damage. Because the optimal border policy includes a strategic component that is inconsistent with legal commitments under the General Agreement on Tariffs and Trade (GATT), we reevaluate optimal border policy that incorporates a GATT restriction reflecting cooperative trade. We show in this setting that the optimal border adjustment taxes the carbon content of trade below the domestic Pigouvian carbon price.

This finding is of first-order importance for policymakers, as it stands in contrast to the standard advice to impose the domestic carbon price on the carbon content of trade. The wedge between the domestic carbon price and the optimal border adjustment arises in general equilibrium because border adjustments inadvertently drive up consumption of emissions-intensive goods in unregulated regions. The magnitude of this wedge depends on the ability of the country imposing the tariffs to affect world prices and ultimately production of the polluting good in unregulated countries. If world prices are unaffected by policy in the regulating country, or if production in unregulated countries is unresponsive to changes in world price, the optimal border adjustment tends toward zero.

Our numerical simulations of Annex-I carbon policy illustrate that this is not simply a theoretical concern. We find an optimal import tariff on the carbon content of aluminum and nonferrous metals that is on the order of 40% of the domestic carbon price. The numeric simulations support the theoretic findings that optimal environmental tariffs are sensitive to the regulated region's international market power and the unregulated region's consumption and production responses. We caution that optimal border carbon adjustments are below the domestic carbon price under cooperative trade. Countries that impose border carbon adjustments at the domestic carbon price will be extracting rents from unregulated regions at the expense of efficient environmental policy and consistency with international law.

# Appendix A: Proof of Theorem 1 (Markusen 1975)

We derive one equation from (5) and two equations from (2), and we substitute those equations into the welfare change equations in the following pages. First, if a unique domestic import quantity is associated with every world price ratio, from the balance-of-payments constraint (5), we can specify the world price ratio as a function of the import quantity as follows:

$$p^* = g(m_X); \qquad dp^* = g' dm_X, \qquad g' > 0. \tag{19}$$

Second, as Vandendorpe (1972) derives from (2), the supply relationships are

$$\frac{dX_r}{dp_r} = R_{Xr}, \qquad where \quad R_{Xr} = \left(-\frac{\partial^2 L_r}{\partial (X_r)^2}\right)^{-1}, \qquad r \in \{N, S\}. \tag{20}$$

Third, totally differentiating (2) and dividing by $\frac{\partial F_r}{\partial Y_r}$ yields

$$\frac{\partial F_r/\partial X_r}{\partial F_r/\partial Y_r} dX_r + dY_r = p_r dX_r + dY_r = 0, \qquad r \in \{N, S\}, \tag{21}$$

and at an equilibrium, $\frac{\partial F_r/\partial X_r}{\partial F_r/\partial Y_r}$ equals $p_r$, where $p_N = p$ and $p_S = p^*$.

Totally differentiating (3) and dividing by $\frac{\partial U_N}{\partial C_{YN}}$ yields the change in the North welfare in terms of consumption good $Y$, $\frac{dU_N}{\partial U_N/\partial C_{YN}}$. Since the welfare in $N$ is maximized when $\frac{dU_N}{\partial U_N/\partial C_{YN}} = 0$, we find the conditions to make this true. The welfare change is as follows:

$$\frac{dU_N}{\partial U_N/\partial C_{YN}} = \frac{\partial U_N/\partial C_{XN}}{\partial U_N/\partial C_{YN}} dC_{XN} + dC_{YN} + \frac{\partial U_N/\partial Z}{\partial U_N/\partial C_{YN}} dZ = q dC_{XN} + dC_{YN} + q_Z dZ, \tag{22}$$

where, $q = \frac{\partial U_N/\partial C_{XN}}{\partial U_N/\partial C_{YN}}$ is the marginal rate of substitution between goods $X$ and $Y$, and $q_Z = \frac{\partial U_N/\partial Z}{\partial U_N/\partial C_{YN}}$ is the marginal rate of substitution between pollution $Z$ and good $Y$. Again note that $q_Z$ is negative because the pollution level $Z$ has a negative impact on the welfare ($\partial U_N/\partial Z$ is negative). We make several substitutions to derive the optimal policy conditions. First, using $dC_{iN} = di_N + dm_i$ from (5) yields

$$\frac{dU_N}{\partial U_N/\partial C_{YN}} = dY_N + dm_Y + q dX_N + q dm_X + q_Z dZ. \tag{23}$$

Second, using $dm_Y = -m_X dp^* - p^* dm_X$ from (5) and $dY_N = -pdX_N$ from (21) yields

$$\frac{dU_N}{\partial U_N/\partial C_{YN}} = (q-p)dX_N + (q-p^*)dm_X - m_X dp^* + q_Z dZ. \tag{24}$$

Differentiating (1), and noting that the supply response in $S$ [see (20)] is driven by a change in the international price ($p^*$), yields

$$dZ = \frac{\partial Z}{\partial X_N} dX_N + \frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*} dp^*. \tag{25}$$

Finally, by using $q - p^* = p^*\tau$ and $q - p = pt_X$ from (4) and replacing $dZ$ from (25) and $dp^*$ from (19), (24) becomes

$$\frac{dU_N}{\partial U_N/\partial C_{YN}} = \left[p^*\tau - m_X \frac{dp^*}{dm_X} + q_Z \frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*}\frac{dp^*}{dm_X}\right] dm_X + \left[pt_X + q_Z \frac{\partial Z}{\partial X_N}\right] dX_N. \tag{26}$$

Since the welfare change (or (26)) is zero at optimal, the optimal tariff and production tax are thus given by

$$\begin{aligned} \tau &= \frac{m_X}{p^*}\frac{dp^*}{dm_X} - \frac{q_Z}{p^*}\frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*}\frac{dp^*}{dm_X}, \\ t_X &= -\frac{q_Z}{p}\frac{\partial Z}{\partial X_N}. \end{aligned} \tag{27}$$

# Appendix B: Proof of Proposition 1

We now modify the previous model to incorporate the transfer from the North to the South such that $U_S = \bar{U}_S$. Correspondingly, we modify (19) as follows:

$$p^* = G(m_X, T), \qquad dp^* = G_{m_X} dm_X + G_T dT. \tag{28}$$

We use the same procedure as we used to derive (26), but now considering (28) and $dm_Y = -dT - m_X dp^* - p^* dm_X$ [derived from (8)]. We arrive at the following:

$$\begin{aligned}\frac{dU_N}{\partial U_N/\partial C_{YN}} &= \left[p^*\tau - m_X G_{m_X} + q_Z \frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*} G_{m_X}\right] dm_X \\ &+ \left[pt_X + q_Z \frac{\partial Z}{\partial X_N}\right] dX_N \\ &+ \left[-1 - m_X G_T + q_Z \frac{\partial Z}{\partial X_S}\frac{dX_S}{dp^*} G_T\right] dT. \end{aligned} \tag{29}$$

Thus, we need to determine $dT$, or the change in the transfer required to hold the South's welfare constant. Let $E_S(p^*, \bar{U}_S)$ indicate the expenditure function of the representative agent in the South. At the solution this equals income, which is the value of production at world prices plus the transfer. Thus we have the following:

$$E_S(p^*, \bar{U}_S) = p^* X_S + Y_S + T, \tag{30}$$

and solving for $T$ we have

$$T = E_S(p^*, \bar{U}_S) - p^* X_S - Y_S. \tag{31}$$

Differentiating (31) and noting that $p^* dX_S + dY_S = 0$ from (21) gives

$$dT = \left(\frac{\partial E(p^*, \bar{U}_S)}{\partial p^*} - X^S\right) dp^*. \tag{32}$$

Applying Shephard's lemma yields

$$dT = -m_X dp^*. \tag{33}$$

Replacing $dp^*$ by using (28) gives us

$$dT = -\frac{m_X G_{m_X}}{1 + m_X G_T} dm_X. \tag{34}$$

Now substituting (34) into (29) yields

$$\frac{dU_N}{\partial U_N / \partial C_{YN}} = \left[ p^* \tau + q_Z \frac{\partial Z}{\partial X_S} \frac{dX_S}{dp^*} \frac{G_{m_X}}{1 + m_X G_T} \right] dm_X + \left[ p t_X + q_Z \frac{\partial Z}{\partial X_N} \right] dX_N. \tag{35}$$

Furthermore, we substitute $\frac{G_{m_X}}{1+m_X G_T}$ out as follows. From (28) we have

$$\frac{dp^*}{dm_X} = G_{m_X} + G_T \frac{dT}{dm_X}. \tag{36}$$

Now from (34), (36) becomes

$$\frac{dp^*}{dm_X} = \frac{G_{m_X}}{1 + m_X G_T}. \tag{37}$$

Thus, (35) becomes

$$\frac{dU_N}{\partial U_N / \partial C_{YN}} = \left[ p^* \tau + q_Z \frac{\partial Z}{\partial X_S} \frac{dX_S}{dp^*} \frac{dp^*}{dm_X} \right] dm_X + \left[ p t_X + q_Z \frac{\partial Z}{\partial X_N} \right] dX_N. \tag{38}$$

Since the welfare change is zero at optimal, the optimal tariff and production tax are

$$\begin{aligned} \tau &= -\frac{q_Z}{p^*} \frac{\partial Z}{\partial X_S} \frac{dX_S}{dp^*} \frac{dp^*}{dm_X}, \\ t_X &= -\frac{q_Z}{p} \frac{\partial Z}{\partial X_N}. \end{aligned} \tag{39}$$

# References

Aldy, Joseph E., and Robert N. Stavins (2008) 'Designing the Post-Kyoto climate regime: Lessons from the Harvard project on international climate agreements.' Report, Harvard Kennedy School, Belfer Center for Science and International Affairs, November

Aldy, Joseph E., Peter R. Orszag, and Joseph E. Stiglitz (2001) 'Climate change: An agenda for global collective action.' Conference Paper, Pew Center on Global Climate Change, October

Allen, Derry (2010) 'OECD global forum on environment focusing on sustainable materials management: Materials case study 2: Aluminum.' OECD Working Document, OECD Environment Directorate

Bagwell, Kyle, and Robert W Staiger (2011) 'What do trade negotiators negotiate about? Empirical evidence from the World Trade Organization.' *American Economic Review* 101(4), 1238–1273

Barrett, Scott, and Robert Stavins (2003) 'Increasing participation and compliance in international climate change agreements.' *International Environmental Agreements* 3(4), 349–376

Böhringer, Christoph, Andreas Lange, and Thomas F. Rutherford (2014) 'Optimal emission pricing in the presence of international spillovers: Decomposing leakage and terms-of-trade motives.' *Journal of Public Economics* 110(1), 101 – 111

Böhringer, Christoph, Edward J. Balistreri, and Thomas F. Rutherford (2012) 'The role of border carbon adjustment in unilateral climate policy: Overview of an Energy Modeling Forum study (EMF 29).' *Energy Economics* 34, Supplement 2(0), S97 – S110

Böhringer, Christoph, Jared C. Carbone, and Thomas F. Rutherford (2013a) 'Embodied carbon tariffs.' Working Papers, University of Calgary, August

_ (2013b) 'The strategic value of carbon tariffs.' Working Papers 4482, Center for Economic Studies and the Ifo Institute (CESifo), November

Bray, E. Lee (2010) '2008 minerals yearbook: Aluminum.' U.S. Geological Survey annual publication, U.S. Geological Survey

Broda, Christian, Nuno Limao, and David E Weinstein (2008) 'Optimal tariffs and market power: The evidence.' *American Economic Review* 98(5), 2032–2065

Carbone, Jared C., and V. Kerry Smith (2013) 'Valuing nature in a general equilibrium.' *Journal of Environmental Economics and Management* 66(1), 72 – 89

Copeland, Brian R. (1996) 'Pollution content tariffs, environmental rent shifting, and the control of cross-border pollution.' *Journal of International Economics* 40(3-4), 459 – 476

Cosbey, Aaron, Susanne Droege, Carolyn Fischer, Julia Reinaud, John Stephenson, Lutz Weischer, and Peter Wooders (2012) 'A guide for the concerned: Guidance on the elaboration and implementation of border carbon adjustments.' Entwined Policy Report 03, Entwined

de Cendra, Javier (2006) 'Can emissions trading schemes be coupled with border tax adjustments? An analysis vis-à-vis WTO law.' *Review of European Community & International Environmental Law* 15(2), 131–145

Elliott, Joshua, Ian Foster, Samuel Kortum, Todd Munson, Fernando Pérez Cervantes, and David Weisbach (2010) 'Trade and carbon taxes.' *American Economic Review* 100(2), 465 – 469

Gros, Daniel (2009) 'Global welfare implications of carbon border taxes.' *CEPS Working Document No. 315*

Hoel, Michael (1996) 'Should a carbon tax be differentiated across sectors?' *Journal of Public Economics* 59(1), 17 – 32

Horn, Henrik, and Petros C. Mavroidis (2011) 'To B(TA) or Not to B(TA)? On the legality and desirability of border tax adjustments from a trade perspective.' *The World Economy* 34(11), 1911–1937

Jakob, Michael, Robert Marschinski, and Michael Hübler (2013) 'Between a rock and a hard place: A trade-theory analysis of leakage under production- and consumption-based policies.' *Environmental and Resource Economics* 56(1), 47–72

Keen, Michael, and Christos Kotsogiannis (2014) 'Coordinating climate and trade policies: Pareto efficiency and the role of border tax adjustments.' *Journal of International Economics* 94(1), 119 – 128

Krutilla, Kerry (1991) 'Environmental regulation in an open economy.' *Journal of Environmental Economics and Management* 20(2), 127–142

Lerner, Abba P. (1936) 'The symmetry between import and export taxes.' *Economica* 3(11), 306–313

Ludema, Rodney D, and Anna Maria Mayda (2013) 'Do terms-of-trade effects matter for trade agreements? Theory and evidence from WTO Countries.' *Quarterly Journal of Economics* 128(4), 1837–1893

Ludema, Rodney D., and Ian Wooton (1994) 'Cross-border externalities and trade liberalization: The strategic control of pollution.' *Canadian Journal of Economics* 27(4), 950–966

Markusen, James R. (1975) 'International externalities and optimal tax structures.' *Journal of International Economics* 5(1), 15 – 29

Narayanan, G. Badri, and Terrie L. Walmsley (2008) *Global Trade, Assistance, and Production: The GTAP 7 Data Base* (Center for Global Trade Analysis: Purdue University)

Nordhaus, William (2015) 'Climate clubs: Overcoming free-riding in international climate policy.' *American Economic Review* 105(4), 1339–70

Pauwelyn, Joost (2013) 'Carbon leakage measures and border tax adjustments under WTO law.' In *Research Handbook on Environment, Health and the WTO,* ed. Geert Van Calster and Denise Prévost Research Handbooks on the WTO series (Cheltenham, UK: Edward Elgar Publishing, Inc.) chapter 15, pp. 448–506

Rutherford, Thomas F. (2010) 'Climate-linked tariffs: Practical issues.' In 'Thinking Ahead on International Trade (TAIT) Second Conference Climate Change, Trade and Competitiveness: Issues for the WTO'

Stiglitz, Joseph E. (2013) 'Sharing the burden of saving the planet: Global social justice for sustainable development.' In *The Quest for Security: Protection without Protectionism and the Challenge of Global Governance,* ed. Mary Kaldor and Joseph E. Stiglitz (New York: Columbia University Press) pp. 161–190

Tamiotti, Ludivine (2011) 'The legal interface between carbon border measures and trade rules.' *Climate Policy* 11(5), 1202–1211

van Asselt, Harro, Thomas L. Brewer, and Michael A. Mehling (2009) 'Addressing leakage and competitiveness in US climate policy: Issues concerning border adjustment measures.' Working Papers, Climate Strategies

Vandendorpe, Adolf L. (1972) 'Optimal tax structures in a model with traded and non-traded goods.' *Journal of International Economics* 2(3), 235–256

Yonezawa, Hidemichi (2012) *Theoretic and Empirical Issues Related to Border Carbon Adjustments* (Ph.D. Thesis: Colorado School of Mines)

Yonezawa, Hidemichi, Edward J. Balistreri, and Daniel T. Kaffine (2012) 'The suboptimal nature of applying Pigouvian rates as border adjustments.' Working Paper 2012-02, Colorado School of Mines: Division of Economics and Business