TRADE AND TREES

Bård Harstad

February 16, 2021

Abstract

Free trade can often lead to resource depletion, such as deforestation in the tropics. This paper first presents a dynamic model whereby the South (S) depletes to export the extracted units (timber) or the produce (beef) from land available after depletion. Because of the damages, the North benefits from trade liberalization only if the remaining stock is, in any case, diminished. For that reason, S speeds up exploitation. The negative results are reversed if the parties can negotiate a contingent trade agreement, whereby the allocation of gains from trade, and thus the location on the Pareto frontier, is sensitive to the size of the remaining stock. In equilibrium, S conserves to maintain its favorable terms of trade, S conserves more than in autarky, and more when the gains from trade are large. The parties cannot commit to future policies, but they obtain the same outcome as if they could.

Key words: Exhaustible resources, deforestation, international trade, trade agreements, environmental conservation, renegotiation.
JEL: F18, F13, F55, Q56, Q37.

CONTACT: Department of Economics and the Frisch Centre, University of Oslo. Email: bard.harstad@econ.uio.no
ACKNOWLEDGEMENTS: I have benefited from many discussions and suggestions, and I am especially grateful for the detailed comments from Arild Angelsen, Geir Asheim, Torfinn Harding, Henrik Horn, Giovanni Maggi, and Rick van der Ploeg. Kristen Vamsæter and Valer-Olimpiu Suteu provided excellent research assistance and Frank Azevedo helped with the editing. The research received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement 683031).

I. Introduction

There is often tension between proponents of international trade and environmental activists. The last few rounds of multilateral trade negotiations attracted tens ouf thousands of protesters and activists.[1] They requested that trade should be limited rather than liberalized. Trade negotiations have mostly proceeded bilaterally in recent years, but the tension has not weakened.[2]

In June, 2019, Brazil led the Mercosur trade bloc to conclude its largest trade agreement ever with the European Union. Two months later, Mercosur concluded an agreement with EFTA, and it has continued to negotiate with other potential trading partners such as Canada, the US, and Asian countries. The trade agreements will change Brazil's economy. Up to now, Brazil has had a relatively closed economy; 80 percent of the meat it produces, for example, has been consumed domestically.[3]

While the trade negotiations concluded, deforestation rates increased and the forest fires gained international media attention. A year later, deforestation has continued to increase, and it is reported to be even higher than in 2019.[4] Consequently, critics argue that the treaty with the EU should not be ratified in their current form,[5] and it is opposed in countries such as France, Germany, Netherlands, Belgium, Ireland, Austria, and Luxembourg.[6]

The tension between trade and environmental concerns is not unjustified: Trade motivates countries to specialize in their comparative advantages and, for many countries in the South, this specialization leads to resource exploitation and agricultural expansion. Consistent with this logic, several empirical investigations document that trade agreements do cause resource depletion, such as deforestation, and that the damages are immensely costly for the society (Franklin and Pindyck, 2018).[7]

The question is thus not whether there can be a negative relationship between trade and conservation, but whether there must be.

This paper provides two analyses. The first, benchmark case, rests on a simple model of free trade and resource exploitation. The model is stylized but tractable, and it illustrates the negative two-way interaction between trade and conservation. The second analysis employs the very same model to prove

[1] For example, about the 1999 WTO negotiations in Seattle, *The New York Times* wrote (Oct. 13, 1999) that 50,000 demonstrators were expected and, underlying the protests, there "*is a fundamental disagreement about the proper role of the trade organization.*"

[2] When the Transatlantic Trade and Investment Partnership (TTIP) agreement was negotiated, protesters said they expected 250,000 demonstrators to turn out in Germany because "*TTIP threaten environmental and consumer protection*" (*The Guardian*, Sept. 17, 2016).

[3] *The Washington Post,* Aug. 27, 2019.

[4] *The New York Times*, June 6th, 2020.

[5] The trade deal is also critisized by Brazilian scholars. Professor Oliver Stuenkel argue (*Financial Times*, Aug. 10th, 2020): "*Foreign support for Brazil's OECD candidacy, and ratification of the EU-Mercosur trade deal, should be made conditional on Brasília restoring environmental regulations and Ibama's fining capacity... even Mr Bolsonaro understands the value of foreign investment and the worth of a trade deal estimated to add $90bn over the next 15 years to Brazil's sagging economy.*"

[6] On Sept. 6, 2020, *Financial Times* concluded: "*Cynically speaking, the Mercosur association agreement, for example, has no chance of being signed off unless the EU's environmental objectives are given greater prominence.*"

[7] For papers verifying the connection between trade liberalization and deforestation in the tropics, see, for example, the analysis by Barbier (2000), or evidence provided by Faria et al. (2016), Pendrill et al. (2019), and Abman and Lundberg (2020).

how all the negative results are reversed if the parties are allowed to negotiate preliminary "default" tariffs that can depend on measurable aspects, such as the size of the remaining resource stock.

The model is dynamic and it includes the South (S), endowed with a depletable resource, and the North (N), preferring conservation. The resource can be oil or coal but, to fix ideas, I frequently refer to tropical forests. Although forests are in principle renewable resources, for practical purposes tropical forests are depletable because when the forest is logged, the land is converted to agriculture. In the model, S can consume the extracted amount (e.g., the timber) and/or the produce (e.g., the beef) produced on the land that becomes available after exploitation. If instead of consuming its timber and beef domestically S exports to N, S might obtain higher prices. In return, N gains from its access to S's markets.

Trade liberalization increases S's prices for timber and beef, so exploitation increases when a free trade agreement (FTA) is signed. For that reason, an FTA is socially optimal only if the remaining resource stock has already diminished, so that the additional damage to N, when the remaining resource is depleted, is less than the gains from trade. In other words: Trade causes depletion, and depletion leads to an FTA.

Anticipating the FTA that will be signed when the resource stock is diminished, S faces an additional incentive to exploit, even before the FTA is signed. This incentive, in turn, can persuade N to sign the agreement right away. If N could commit, N would commit to autarky when the remaining stock is large. But because S understands that N will always allow for trade later, if the stock is diminished, S is willing to exploit, and therefore, N is willing to trade. The equilibrium might thus be that N and S sign the FTA, and the resource is depleted, even when the damages are larger than the gains from trade.

The second analysis shows how all these negative findings can be reversed by a conservation-sensitive, or contingent trade agreement (CTA). In a CTA, N and S take advantage of the fact that the size of the resource stock is verifiable, and that the gains from trade can be allocated arbitrarily between the parties as a function of the remaining stock. The first challenge is that the more of the forest S has depleted, the larger is the agricultural production and the larger are the total gains available from trade. As long as S can capture a fraction of these additional gains from trade, S has an incentive to extract more than in autarky. The CTA must thus be designed so that the allocation obtained by S decreases when the stock is being depleted. A second challenge, however, is that N and S cannot rely on punishments and plans that are not subgame perfect, or renegotiation proof: For example, if N and S agreed to an FTA that ceased to exist as soon as S had depleted, the two parties would have an incentive to renegotiate later, if the resource has been depleted nevertheless. The main result of this paper is to show how the CTA can still motivate conservation, and more conservation than we would have seen in autarky, despite these two challenges.

Just as with the FTA, the CTA can be implemented by any type of side transfer combined with the trade agreement. The natural type of side transfer is to permit one party, or both, to introduce tariffs in order to improve one's terms of trade. The tariffs will dictate how the gains from trade are allocated between the parties. When this allocation is on the Pareto frontier, it is also renegotiation proof: Any change will then harm and be vetoed by (at least) one of the parties.

The CTA allows the parties to negotiate tariffs, and thus the allocation of gains from trade, that depend on the resource stock. Intuitively, conservation might succeed if S faces no tariff when the stock is large while N faces no tariff when the stock is small. When the tariff functions are renegotiation proof and thus credible, then S is willing to conserve in order to maintain the type of trade agreement that secures S's attractive terms of trade. The larger the total gains from trade are, the bigger is the cake that can be allocated to S as long as the stock is large, and the more can be conserved, I show – if export subsidies are unavailable. If the tariffs can be accompanied by N's export subsidies (as in Grossman and Helpman, 1995), then there is no limit to how large the effective transfer from N to S can be, and then the first best can be implemented.

The optimal CTA is not only feasible, and renegotiation-proof, but it is also the *equilibrium* agreement that N and S will sign as soon as the future (default) tariffs are allowed to depend on the stock size. The contractual environment is only slightly different from that of the situation in which the parties negotiate an FTA. Because this difference is small, it is worth emphasizing that the relationships between trade and exploitation are entirely reversed:

(1) While an FTA causes exploitation, the CTA motivates more conservation than in autarky.

(2) While an FTA is socially valuable if the remaining stock is small, the CTA is more valuable if the stock is large.

(3) While S is tempted to exploit to obtain a (better) FTA, with a CTA, S conserves to avoid a less attractive agreement.

(4) The larger the gains from trade, the more is exploited under the FTA, but the more can be conserved with a CTA.

(5) When side transfers are facilitated by tariff adjustments, then N's equilibrium tariff increases or S's tariff decreases in the stock size with an FTA, but the reverse holds with a CTA.

It is worth repeating that the parties do not tie their hands, and they do not commit to tariffs that are functions of the stock. The parties are free to renegotiate the tariffs at any time, if they both benefit from doing so. In fact, the tariff functions can be interpreted as default tariffs, that is, the agreement N and S revert to unless they both agree on something else. The CTA can even be made robust to unilateral requests to renege on the treaty. Nevertheless, the equilibrium CTA implements the same outcome that N and S would have obtained if had they been able to *commit* to a trade agreement that was conditional on the resource stock. Consequently, in this model, the optimal CTA is not an arbitrary design from which N and S can make further improvements: It implements the first-best outcome if export subsidies are available, and the second-best outcome if they are not.

The analysis of CTAs is explorative rather than aimed at explaining the types of agreements that we have observed so far. Nevertheless, it is important to recognize that the CTA is realistic and feasible.

As discussed in Section VI, forest cover is verifiable, sometimes used to determine fiscal transfers, and influenced by governmental policy (Burgess et al., 2019). Furthermore, the structure of CTAs is related to a 2020 policy initiative from France and Netherlands: In a recent "non-paper," they recommend that the implementation of trade agreements should proceed step-wise and hinge on the gradual implementation of sustainability requirements.[8] The non-paper is brief and not specific on how the implementation can be achieved. The analysis below is thus important because it sheds light on this policy initiative by showing how trade agreements can be designed so that they motivate conservation rather than exploitation.

Literature. My basic model of trade and agreements draws on existing literature (see the surveys by Maggi, 2014; Bagwell and Staiger, 2016). For example, tariff reductions are motivated by the terms-of-trade effects of tariffs (as in Bagwell and Staiger, 2004 and 2011; Ludema and Mayda, 2013; Grossman, 2016), I permit transfers at the negotiation stage (Aghion et al., 2007; Maggi and Ossa, 2020), I consider export subsidies (as in Grossman and Helpman, 1995), and I allow for renegotiation (Ludema, 2001; Maggi and Staiger, 2015).[9]

The literature on how trade and resource extraction are related goes back to Dasgupta et al. (1978), who study depletion rates in open economies.[10] In the traditional literature on trade and the environment, countries may reduce environmental standards to become competitive (Markusen, 1975) or to specialize in their comparative advantages: The South may have comparative advantage in environmentally damaging production because of policies (Pethig, 1976) or because of lower income levels (Copeland and Taylor, 1994).[11] If countries in the South struggle with an open-access problem, and are unable to control extraction rates, then trade can worsen the problem and cause depletion (Chichilnisky, 1994; Brander and Taylor, 1997 and 1998; Karp et al., 2001).

In fear of a negative relationship, the World Bank (2019) and IPBES (2019) have expressed concerns about how trade is liberalized.[12] To mitigate some of the problems, scholars have recommended trade

[8]In the non-paper, they propose: "*Given the lack of progress in compliance with"*... the Trade and Sustainable Development (TSD) Chapters in trade agreements, "*Parties should introduce, where relevant, staged implementation of tariff reduction linked to the effective implementation of TSD provisions and clarify what conditions countries are expected to meet for these reductions, including the possibility of withdrawal of those specific tariff lines in the event of a breach of those provisions.*" The non-paper is available here:
https://nl.ambafrance.org/Non-paper-from-the-Netherlands-and-France-on-trade-social-economic-effects-and

[9]The focus on terms-of-trade is normal in the shallow integration literature, but studies of deep integration also consider behind-the-border policies (Antras and Staiger, 2012), such as domestic regulation and product standards (Grossman et al. 2000), or concentrate on principles such as reciprocity and nondiscrimination (Bagwell and Staiger, 1999) to prevent bilateral opportunism through "concession erosion" (Bagwell and Staiger, 2016). In addition, countries might sign treaties to persuade firms to invest in new technologies (Matsuyama, 1990), to withstand pressure from lobbies (Maggi and Rodriguez-Clare, 1998), to diversify (Caselli et al., 2020), or to win elections (Battaglini and Harstad, 2020).

[10]Relatedly, Hillman and Van Long (1983) studied a country depleting a resource at the same time as it was importing extracted amounts from another country. If there is a (lower) risk of trade disruption, then the country conserves more (less) of its own resource. With a larger number of jurisdictions, depletion can be larger also because prices will be less sensitive to one's own supply (see Markusen, 1981, for the theoretical point, Burgess et al., 2012, for evidence when it comes to deforestation, and Harstad and Mideksa, 2017, for further on the theory).

[11]For this reason, trade can increase global pollution if income differences are large (Copeland and Taylor, 1995). On the other hand, trade can raise income levels, and because of the environmental Kuznets curve, the outcome can be a cleaner environment (Antweiler et al., 2001; Copeland and Taylor, 2004).

[12]The World Bank (2019:8) states: "...*the expansion of livestock production in Brazil could increase deforestation. Only if these adverse impacts are addressed through appropriate spatial and environmental policies will trade integration be a pathway to development.*" IPBES (2019, Ch 6:138) states: "*the potential of WTO and other free trade agreements and WTO regulations to contribute to conservation and sustainability is criticized... While other regional or bilateral free trade agreements such as NAFTA include environmental provisions, these have mostly been implemented in a narrow way and*

sanctions (Barrett, 1997), border tax adjustments (Hoel, 1996; Elliott et al., 2010), and climate clubs (Nordhaus, 2015). The threats to limit trade are not effective (or renegotiation proof) when the resource is non-renewable, however: after the resource is exhausted, it is in everyone's interest to trade.

The first contribution of this paper is to describe the negative relationship between trade and conservation in a dynamic but highly tractable model. The model also uncovers the reverse relationship, from exploitation to trade, and how that relationship, in turn, can motivate further depletion. Most importantly, this paper shows how the countries can reverse the negative relationships by designing a contingent trade agreement. This possibility is highly policy relevant because using explicit compensations in return for conservation is often problematic.[13]

The literature on issue linkages (surveyed by Maggi, 2016) typically considers structurally unrelated issues and that the parties can commit (Abrego et al., 2001; Horstmann et al., 2005). Here, trade and the environment are structurally linked and the parties cannot commit. The two are structurally linked also in the model by Copeland (2000), but he analyzes linkages between a trade agreement and an environmental agreement, while I consider a trade agreement only.[14]

The way in which conservation is implemented in this paper is inspired by how cooperation is implemented in dynamic games when the parties can renegotiate. For example, the standard grim-trigger strategy can motivate cooperation in repeated prisoner dilemma games but, if the parties should happen to end up in a punishment phase, they have an incentive to renegotiate and end the inefficient punishment. To make the punishment credible, one may need to require that the punishment payoffs continue to be on the Pareto frontier, although the payoff must be unattractive for the party that has defected (Mailath and Samuelson, 2006).[15] I combine this logic with the theory of issue linkages because, in my analysis, the Pareto frontier refers to various allocations of the gains from trade, while defection refers to resource depletion. The problem is intricate because the (ex post) Pareto frontier (i.e., the gains from trade) expands if the resource is depleted.

Outline. The next section presents the model of resource depletion and international trade. The model is simple and payoffs are linear functions – not because the simplicity is necessary for the results – but because it is sufficient. Section III derives five propositions: The first three show that trade causes depletion, and depletion causes trade, in every subgame-perfect equilibrium. To obtain sharper

have not resulted in significantly raised levels of environmental protection... At the global level, WTO has started to discuss environmental provisions as part of the Doha negotiations since 2001, but negotiations were not successful and ended in 2016. Since then, bilateral trade agreements have increased in importance, as have the intensification of 'trade wars'."

[13] Explicit compensations for conservation can, in some cases, be highly effective (Souza-Rodrigues, 2019). However, IPBES (2019:54) reports that "*the literature is currently mixed on the success rates of forest carbon projects in general and REDD+ has faced a number of challenges.*" The challenges with this approach include liquidity constraints (Jayachandran, 2013), contractual externalities (Harstad and Mideksa, 2017), and that they lead to embezzlements (Caselli and Michaels, 2013), corruption, and a worse selection of political candidates (Brollo et al., 2013). Furthermore, future payments may not be credible (Harstad, 2016) and they can motivate domestic counter-lobbying (Harstad, 2020).

[14] Horn and Mavroidis (2014) speculate on the reasons for why environmental agreements and trade agreements are seldom linked in reality.

[15] Mailath and Samuelson (2006) provide a textbook treatment of cooperation in dynamic games and of renegotiation proofness (Section 4.6). In plain English, the problem is that "*should [the players] ever find themselves facing an inefficient continuation equilibrium, whether on or off the equilibrium path, they can renegotiate to achieve an efficient equilibrium*" (p. 122). Regarding the solution to this problem "*the key to constructing nontrivial renegotiation-proof equilibria is to select punishments that reward the player doing the punishing*" (p. 135).

predictions, I then characterize the Markov-perfect equilibrium and the equilibrium tariff levels. The same model is employed in Section IV, where Corollary 1 states that all the negative findings are reversed with a CTA. In the robustness section, I show that the CTA can be robust to unilateral requests to renege on the treaty, and that the main results hold if the six linearity assumptions are relaxed. Section VI concludes, while all proofs are in the Appendix.

II. The Model

The South. Let S be a country endowed with a resource stock that can be partly or fully depleted over time. At the beginning of time $t \in \{1, 2, ...\}$, the remaining resource stock is R_t, the part that has been exploited is X_t, and $R_0 = R_t + X_t$ is the original size of the stock. When S exploits $x_t \in [0, R_t]$,

$$R_{t+1} = R_t - x_t \text{ and } X_{t+1} = X_t + x_t.$$

S can benefit from R_t, X_t, and x_t. The resource can represent oil or coal but, to fix ideas, I refer to R_t as the remainder of the rainforest, X_t as the land that has already been logged and converted to agriculture, and the timber currently logged is proportional to x_t. For simplicity, S's agricultural produce (beef) equals its amount of converted land, X_t. In autarky, $\underline{a} \geq 0$ represents the (present-discounted) agricultural value of land, if $(1 - \delta)\,\underline{a}$ measures S's *per-period* utility of the food produced per unit of land and $\delta \in (0, 1)$ is the discount factor. In addition, $\underline{b} \geq 0$ is S's marginal benefit of the extracted resource (timber), while c is the marginal cost of exploitation. The cost c may include the physical as well as the (present-discounted value of the) environmental cost to S when R_t is reduced by a unit. (We will distinguish between the two in Section V.3.)

If S stays in autarky forever, S is a single decision maker maximizing its continuation value:

$$V_S^{AUT}(R_t) \equiv \max_{x_t^{AUT} \in [0, R_t]} (1 - \delta)(X_t + x_t)\,\underline{a} + \underline{b}x_t - cx_t + \delta V_S^{AUT}(R_{t+1}).$$

The linearity in x_t implies that the autarky choices are simple to characterize:

$$x_t^{AUT}(R_t) = \left\{ \begin{array}{l} 0 \text{ if } \underline{a} + \underline{b} \leq c \\ x_t \in [0, R_t] \text{ if } \underline{a} + \underline{b} = c \\ R_t \text{ if } \underline{a} + \underline{b} > c \end{array} \right\}. \tag{1}$$

The North. The North (N) is S's potential trading partner. Just like S, N can experience costs and benefits from the exploitation. In particular, N faces the damage $d > 0$ for each unit that is logged in country S. Equivalently, d represents N's marginal present-discounted value if a unit of the resource is conserved forever.

In addition, N's marginal value from beef is $(1 - \delta)\,\overline{a}$. That is, $\overline{a}$ is N's present-discounted value of consuming a unit of S's agricultural products in every future period. In addition, N's marginal benefit from the extracted resource (timber) is $\overline{b}$.

If $\overline{a} \geq \underline{a}$, it is socially optimal that the beef (X_t) be exported to N, and if $\overline{b} \geq \underline{b}$, it is socially optimal that the timber (x_t) be exported to N. Both inequalities are assumed to hold weakly.[16]

I assume that the seller sets the price.[17] Thus, with a free trade agreement (FTA), S receives $(1-\delta)\,\overline{a}$ for each exported unit of beef in every period, and $\overline{b}$ for each unit of timber. To S, exploitation for trade is strictly beneficial if and only if $\overline{a}+\overline{b}>c$. Exploitation is socially inefficient when

$$\overline{a}+\overline{b}<c+d, \tag{2}$$

which I will assume holds.

N benefits $e>0$ from getting access to S's market. We can endogenize e as follows. Suppose that N has the capacity to export $\psi \geq 0$ units of machines in every period to S. If the citizens of S are willing to pay $\omega \geq 0$, and N's marginal production cost is $\kappa \in [0,\omega]$, then N charges ω and captures the export's entire present-discounted value $(\omega-\kappa)\,\psi/\,(1-\delta)$, henceforth defined as e.

We have a general-equilibrium model, and trade is balanced, if we introduce a numeraire good that can be used as a currency. For example, countries may trade cookies or labor services. All parameters above are then measured relative to the value of the numeraire.

The First Best. In every period t, the gains from trade are given by

$$(1-\delta)\,e+(1-\delta)\,(\overline{a}-\underline{a})\,(X_t+x_t)+\left(\overline{b}-\underline{b}\right)x_t>0.$$

The first-best outcome is simply that the parties trade and that S conserves in every period (i.e., $x_t=0$).

Timing of the Game. In each period, t, the parties first bargain whether to open up for trade, if they haven't opened up already. Second, S decides on $x_t \in [0,R_t]$. S is assumed to conserve whenever indifferent. Finally, trade and consumption take place.

The Bargaining Solution. Both countries must agree to open up for trade, but side transfers can be used if the countries agree on an FTA. Let $\alpha \in [0,1]$ measure S's share of the bargaining surplus while $1-\alpha$ measures N's share. This outcome follows, for instance, if we let the asymmetric Nash bargaining solution represent the outcome and α be S's bargaining strength. Alternatively, this allocation would follow from standard noncooperative bargaining games: Suppose, for example, that the bargaining stage in each period consists of a finite number of offers, with negligible discounting between the offers, and where S makes the final offer with probability α. The outcome of this noncooperative bargaining game is that N and S trade if the surplus is positive and S captures the fraction α of the surplus.

Section III.E explains how transfers can be facilitated by tariffs and subsidies, but also how the results survive without transfers.

Strategies and Equilibrium Concept. S's strategy is mapping from the set of histories to $x_t \in [0,R_t]$. N does not take any action. (As explained, N and S are simply sharing the gains from trade liberalization

[16]This assumption is without loss of generality for the present analysis because if, for example, $\overline{a}<\underline{a}$, S's beef will not be exported and the realized gains from trade will be zero, i.e., the same as when $\overline{a}=\underline{a}$.

[17]Antras and Staiger (2012) show that the traditional justification of "shallow integreation" for trade agreements is unjustified if the prices are, instead, negotiated. As long as the tariffs influence the terms of trade, however, alternative pricing models will not reverse the results emphasized in this paper.

if the gains are positive.) The below inefficiency results (Propositions 1–3) hold for *all* subgame-perfect equilibria (SPEs). The efficiency result in Section IV holds *despite* the restriction to Markov-perfect equilibria (MPEs). It is, at that stage, natural to focus on MPEs given the importance of the state variable R_t.

Linearities and Generalizations. The model is stylized and simple not because a simple model is necessary for the results below, but because it is sufficient. I assume that a trade agreement is binding for the parties, unless both agree on something else, but Section V.1 verifies that the results hold for non-binding treaties, that is, when a country can unilaterally renege on the terms Section V.2 replaces the constants above by concave or convex functions. Depletion rates are then more gradual – and realistic – but the main results emphasized below continue to hold.

III. Free Trade Agreements

The simple model above captures a basic mechanism for how trade and exploitation can be mutually dependent. Before characterizing the equilibrium in detail, it is useful to start with three results that hold for all SPEs.

III.A. Trade Causes Exploitation

With free trade forever, S is, as in autarky, a single decision maker maximizing its continuation value. Now, however, S faces an alternative to consuming the resource domestically. In autarky, S exploits $x_t^{AUT} > 0$ only if $\underline{a} + \underline{b} > c$. With trade, S exploits $x_t^{FTA} > 0$ also if $\overline{a} + \overline{b} > c$.

It follows that S exploits more with trade than in autarky.

Proposition 1. *Free trade causes depletion: In every SPE,*

$$\begin{aligned} x_t^{FTA} &= R_t \geq x_t^{AUT} = 0, \text{ if} \\ c &\in \left[\underline{a} + \underline{b}, \overline{a} + \overline{b}\right). \end{aligned} \tag{3}$$

If $\overline{a} + \overline{b} \leq c$*, then* $x_t^{AUT} = x_t^{FTA} = 0$*. If* $\underline{a} + \underline{b} > c$*, then* $x_t^{AUT} = x_t^{FTA} = R_t$*.*

Proposition 1 implies that no SPE can implement the first best with an FTA and no exploitation.

It is easy to see that the equilibrium survives also if trade liberalization is reversible and must be decided on in every period. As soon as the parties trade in one period, S exploits by choosing $x_t = R_t$. Thereafter, when the resource is depleted, trade is unambiguously efficient in every period. A threat to not trade after depletion is not credible.[18]

III.B. Exploitation Causes Trade

[18] If the transfers/tariffs can depend on the history, there exist SPEs in which N pays S in every period as long as S conserves, if just δ is sufficiently large. Such SPEs cease to exist if the transfer, as in this section, cannot be conditioned on S's action.

With transfers at the negotiation stage, it is the sum of the two continuation values that determines whether an agreement is beneficial. By comparing the autarky payoffs following (1) and the FTA payoffs following Proposition 1, we can conclude that the benefit from an FTA can be negative if R_t is large. When (3) holds, the resource will be depleted with an FTA but not in autarky.

If R_t is already diminished, however, the additional damage is small and outweighed by the gains from trade. Thus, there exists a threshold, R^*, so that the FTA is socially valuable if $R_t \leq R^*$.

PROPOSITION 2. *Suppose trade influences x_t (i.e., (3) holds). The social value of the FTA at time t decreases in R_t and it is positive if the gains from trade are large and R_t is small, i.e., if:*

$$\frac{e + (\overline{a} - \underline{a}) R_0}{R_t} \geq c + d - \underline{a} - \overline{b} \Leftrightarrow \tag{4}$$

$$R_t \leq R^* \equiv \frac{e + (\overline{a} - \underline{a}) R_0}{c + d - \underline{a} - \overline{b}}. \tag{5}$$

The proposition describes a second-best outcome: Given the inefficiency uncovered by Proposition 1, it is socially optimal with trade if and only if the resource has already been exploited so much that the remainder R_t is small and the inequality (5) holds. In this case, the parties strictly benefit from trade, despite the fact that trade will motivate further exploitation.

This result generalizes the claim above that if $R_t = 0$, then trade liberalization is unambiguously beneficial for the countries.

For the inequality (5) to fail, $R_t > R^*$. For such a large R_t, the sum of payoffs is larger in autarky than with trade whenever autarky is necessary for conservation to take place, that is, under (3).

III.C. Exploit to Trade

In reality, the parties cannot commit to stay in autarky forever just because trade can cause depletion. If it should happen that the resource is exploited anyway, so that (5) holds, then N and S will find it optimal to trade.

S anticipates that if it exploits enough so that (5) holds, then it will be able to enjoy the gains from trade. Even if $\underline{a} + \underline{b} < c$, so that S finds it costly to deplete the resource in autarky, this cost is worth paying if R_t is already small or if the gains from trade are large.

PROPOSITION 3. *S is willing to exploit in order to obtain an FTA if the gains from trade are large or R_t is small, i.e., if:*

$$\frac{e + (\overline{a} - \underline{a}) R_0}{R_t} > \frac{c - \underline{b} - \underline{a}}{\delta\alpha} \Leftrightarrow \tag{6}$$
$$R_t < \widehat{R} \equiv \delta\alpha \frac{e + (\overline{a} - \underline{a}) R_0}{c - \underline{b} - \underline{a}} \text{ or } \underline{a} + \underline{b} > c.$$

This proposition implies that the second-best outcome, characterized by Proposition 2, cannot be sustained by any SPE. Even if the FTA is not socially valuable, because $R_t > R^*$, S can always obtain a larger continuation than in autarky if $R_t < \widehat{R}$, simply by first exploiting the resource and then trade.

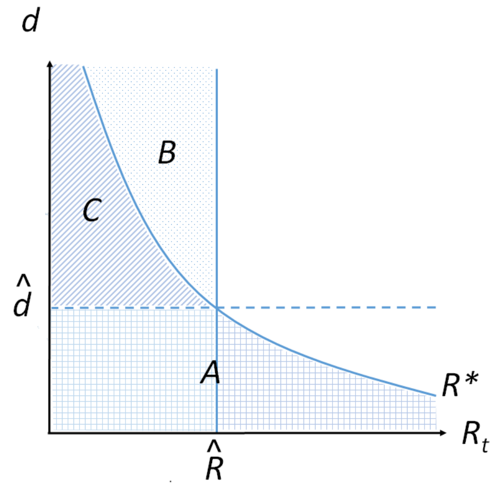

Figure 1: *An FTA is socially optimal if* $R_t < R^*$*, but S is willing to exploit to obtain one if* $R_t < \widehat{R}$.

As illustrated in Figure 1,

$$\widehat{R} > R^* \Leftrightarrow d > \widehat{d} \equiv \left(\frac{1}{\delta \alpha} - 1 \right) (c - \underline{b} - \underline{a}) + \overline{b} - \underline{b}.$$

In region B, where $d > \widehat{d}$ and $R_t \in \left(R^*, \widehat{R}\right)$, trade and exploitation are socially suboptimal, but S is willing to deplete the resource if this is necessary to diminish the stock so much that the FTA becomes acceptable also to N.

Proposition 3 states that S is "willing" to exploit but, on the equilibrium path, N and S find it optimal to sign an FTA as soon as N expects that S will, in any case, exploit the resource.

Remark 1 on Exhaustibility and Irreversability. The negative results from Proposition 2 and 3 follow because the resource is exhaustible and depletion irreversible. If, instead, R_t returned to R_0 after every period, or if the stock was not relevant, then N and S would lose from trade if $\overline{a} + \overline{b} < c + d$, and S would not be able to extract to obtain an FTA.

III.D. Equilibrium

The propositions above hold for every SPE. Proposition 1 states that the first best cannot be supported as an SPE. Proposition 2 describes the second best, given the inefficiency uncovered by Proposition 1, but Proposition 3 states that not even the second best can be supported as an SPE: There is no SPE without trade or exploitation if $R_t \in \left(R^*, \widehat{R}\right)$. There are multiple SPEs, but all of them lead to trade and exploitation.[19]

In games with stocks, it is common to pay special attention to Markov strategies that are functions of the stock. Maskin and Tirole (2001:192–193) write that MPEs "*prescribe the simplest form of behavior*

[19] For example, there are SPEs in which N and S agree to trade every Δ_t period, where Δ_t is so large that S prefers to exploit when there are Δ_t periods left until the next "agreement" period, while S prefers to conserve when there are fewer than Δ_t periods left. These SPEs are not renegotiation proof, however.

that is consistent with rationality" while capturing the fact that "*bygones are bygones more completely than does the concept of subgame-perfect equilibrium.*" So, to offer a sharp characterization of the outcome, I henceforth characterize the MPE in pure and linear strategies. The MPE specifies the default extraction level, $x^D = \phi R_t$, that S will choose at the extraction stage unless the parties have signed an agreement. There are also mixed-strategy MPEs in which $\phi \in [0,1]$ measures the *probability* that S exploits R_t when the parties have not agreed, but the mixed-strategy equilibria are payoff-equivalent to the pure-strategy MPE, thanks to the linear payoffs.[20]

PROPOSITION 4. *In equilibrium, the FTA is signed, and S exploits, if and only if the gains from trade are large or R_t is small. I.e., if*

$$\frac{e + (\overline{a} - \underline{a}) R_0}{R_t} > c + d - \overline{b} - \underline{a}, \qquad \text{and then } x^D = 0, \quad \text{when } d \leq \widehat{d}, \text{ and}$$

$$\frac{e + (\overline{a} - \underline{a}) R_0}{R_t} > \frac{c - \underline{b} - \underline{a}}{\delta \alpha}, \qquad \text{and then } x^D = R_t \cdot \max\{\phi, 1\}, \text{ when } d > \widehat{d}, \text{ where}$$

$$\phi \equiv \frac{d - \widehat{d}}{c + d - \overline{b} - \underline{a} + \frac{\overline{b} - \underline{b}}{1-\delta}} = 1 - \frac{\frac{\overline{b} - \underline{b}}{1-\delta} + \frac{c - \underline{b} - \underline{a}}{\delta \alpha}}{c + d - \overline{b} - \underline{a} + \frac{\overline{b} - \underline{b}}{1-\delta}}.$$

In region A in Figure 1, $d < \widehat{d}$, and S is not willing to exploit in order to obtain an FTA when (4) fails, and S is not willing to exploit in order to obtain a *better* FTA if (4) holds. In this situation, N and S trade if and only if R_t is so small that (4) holds, and the equilibrium threat point is $x^D = 0$. This situation corresponds to the first case in the proposition.

In the second case, and in region B in Figure 1, N and S trade even if the FTA is socially suboptimal, because S will otherwise exploit in order to obtain an FTA later. In equilibrium, S exploits the fraction $\phi > 0$ of the remaining resource if N and S do not sign the FTA. Interestingly, the fraction ϕ is independent of R_t, and it increases in d.

In region C, R_t is even smaller, and both (4) and (6) hold. In this case, S would be willing to exploit R_t in order to obtain an FTA, but doing so is not necessary. After all, N and S jointly benefit from the FTA, even without that threat. In fact, N and S would negotiate an FTA and S would obtain the fraction α of the total surplus even if the "threat point" were that S would not exploit. That surplus, it turns out, can be less than what S can obtain from *first* exploiting and *then* negotiating an FTA. The reason for why it can be less is that when the threat, or "default," outcome is $x^D = 0$, i.e., S does not exploit if N and S fail to sign an FTA at time t, then S must compensate N for the damages N faces given that the FTA, and only the FTA, will cause exploitation. If S exploits first, then the damage is sunk and no such compensation can be requested. The latter option is preferable to S when $d > \widehat{d}$. So, in region C,

[20] The intuition for why there are mixed-strategy MPEs is the following. If S is willing to deplete in order to obtain an FTA, then N and S benefit from signing an FTA right away, but if S can expect an FTA very soon, then S might prefer $x_t = 0$ because S earns more from exporting the timber than from consuming it domestically. This logic shows that, in some cases, there might be no equilibrium in which $x_t = 0$ or $x_t = R_t$ and that instead S might randomize between these two actions. Thanks to the linear payoff functions, a mixed-strategy equilibrium corresponds to a pure-strategy equilibrium in which the default outcome, or threat point, is $x_t^D \in (0, R_t)$. Such an interior pure-strategy exploitation level is payoff equivalent to the mixed-strategy MPE, and it is also more realistic.

S is willing to exploit in order to obtain a *better* FTA. The default is not $x^D = 0$, but $x^D > 0$, and the larger x^D strengthens S's bargaining position.

Because of the linear payoffs, an interior solution, with $\phi \in (0, 1)$, is possible only when S is indifferent between exploitation and conservation. This indifference pins down S's payoff (since it must equal the payoff when S depletes). Since S's payoff equals a fraction α of the total surplus, S's indifference is also pinning down the total gains from signing the agreement and, with that, the equilibrium level of ϕ.

If $\underline{a} + \underline{b} - c > 0$, S benefits from exploiting, even in autarky. For $\underline{a} + \underline{b} - c$ sufficiently large, $\phi \geq 1$, implying that S depletes the resource in the default outcome, that is, if N and S fail to sign the FTA.[21]

III.E. Transfers and Tariffs

The results above continue to hold, qualitatively, even if N and S cannot use transfers at the negotiation stage. For example, Proposition 1 remains unchanged, the inequality in Proposition 2 simplifies to $e/R_t \geq d$, and (6) in Proposition 3 is replaced by $(\overline{a} - \underline{a}) R_0/R_t > c - \underline{a} - \underline{b}$.

But allowing for transfers seems reasonable when it comes to international trade agreements (Grossman and Helpman, 1995; Aghion et al., 2007; Maggi and Ossa, 2020). After all, the transfers do not need to be explicit monetary transfers. They can, alternatively, be facilitated by trading favors, such as when the EU demanded that Russia ratify the Kyoto Protocol in order to obtain the EU's support for Russia's entry into the WTO.[22] International politics are multidimensional and issue linkages like this are quite common.[23]

In addition, the transfers can take the form of tariff adjustments. In practice, there is a large set (i.e., an entire frontier) of Pareto-optimal agreements. After all, trade agreements are not necessarily completely "free" but instead are characterized by reduced tariffs. Reduced tariffs on S's exports benefit S but, because of the terms-of-trade effect, N may be harmed.

This terms-of-trade effect is easily captured in the above model given the inelastic supply that has been assumed. If $\tau_N \leq 1$ measures S's ad valorem import tariff on N's export, then consumers in S are willing to pay only the fraction $(1 - \tau_N)$ for N's goods, relative to how much they would have paid without any tariff. N will find it necessary (and optimal) to reduce the price by this fraction and, therefore, N loses and S gains when τ_N is increased.

A tariff (τ_S) in N on S's beef is similarly benefitting N but harming S. With both tariffs, the payoffs

[21] It is necessary but not sufficient that $\underline{a} + \underline{b} > c$ for $\phi = 1$. Even if S would exploit in autarky, S is still better off by signing an FTA with N, and thus S is willing to conserve so that the extracted resource can be exported in the next period, when an agreement is expected to be signed. If $\underline{a} + \underline{b} - c > 0$ is very large, the harsh threat point $x^D = R_t$ is necessary to ensure that S's share of the total surplus only equals α multiplied by the total surplus (which is larger if x^D is large). This case cannot arise (and we always have $x^D < R_t$) if δ is close to 1. The technicalities and details of the equilibrium are clarified in the Appendix.

[22] https://www.nytimes.com/2004/05/21/international/russia-on-path-to-wto-signs-trade-deal-with-europe-20040521930844453O3.html

[23] Aghion et al. (2007:3) refer to several examples and they conclude: "*We believe...that it is realistic to model trade negotiations as games with transferable utility, because the exchange of concessions on non-trade-related issues often serves the role of transfers that redistribute the gains from trade liberalization.*"

after signing the agreement become:

$$\begin{aligned} V_S^{FTA}(R_t,\overline{\tau}) &= \tau_N e + (1-\tau_S)\,\overline{a}R_0 + \left(\overline{b}-c\right)R_t = V_S^{FTA}(R_t,0) + \overline{\tau}, \text{ and} \\ V_N^{FTA}(R_t,\overline{\tau}) &= (1-\tau_N)\,e + \tau_S\overline{a}R_0 - dR_t = V_N^{FTA}(R_t,0) - \overline{\tau}, \text{ where} \\ \overline{\tau} &\equiv \tau_N e - \tau_S\overline{a}R_0. \end{aligned} \tag{7}$$

Note that $\overline{\tau}$ essentially represents a monetary transfer from N to S relative to what the two countries would have enjoyed if all tariffs were zero. To facilitate this transfer, there is no need for introducing tariffs on timber (x_t) also.

There is no welfare loss associated with the tariffs as long as the tariffs are so small that they do not change the traded quantity. But if there are no export subsidies, the producers in S are willing to export only if the tariff is limited, i.e., if $(1-\tau_S)\,\overline{a} \geq \underline{a}$. (There will be a similar constraint on τ_N.) With more realistic elasticities on the supply and demand functions, the welfare losses from the tariffs would be positive and continuous, but Section V.C explains how two-part tariffs can eliminate the distortions.

In either case, however, the constraints on the tariffs and the welfare losses can be ignored if the tariffs can be accompanied by export subsidies (as in Grossman and Helpman, 1995). If s_S is an ad valorem export subsidy in S, the producers in S are willing to export as long as $(1+s_S)(1-\tau_S)\,\overline{a} \geq \underline{a}$. The export subsidy is just a transfer within the country and it might not influence the countries' payoffs. With this subsidy, and with a similar export subsidy in N, there is no constraint on how large the tariffs or the transfer can be: the producers in country $i \in \{N,S\}$ will continue to export as long as there is no change in $(1+s_i)(1-\tau_i)$.[24]

Alternatively, the transfers can be arbitrarily large if, instead of export subsidies, we permit import subsidies. N subsidizes imports if $\tau_S < 0$ and S subsidizes import if $\tau_N < 0$.

To be able to abstract from the constraint on $\overline{\tau}$, I will in the following allow for subsidies on export or import. The assumption is relaxed in Section IV.C.

In the dynamic game above, the gains from trade include the gains from starting with the FTA at time t instead of at time $t+1$. In the meanwhile, N risks that S exploits x^D. The equilibrium level of x^D is thus going to influence the equilibrium transfer. This is evident in the next proposition, which presents the equilibrium transfer implemented by the tariffs.

Proposition 5. *If N and S sign an FTA at time T, then τ_N is smaller or τ_S is larger if R_T is large:*

$$\overline{\tau}(R_T) = \alpha e - (1-\alpha)\,\Delta_a R_0 - R_T \cdot \left\{ \begin{array}{ll} \alpha d - (1-\alpha)\left(c - \underline{a} - \overline{b}\right) & \text{if} \quad \phi < 0 \\ \overline{b} - \underline{b} + \left[c - \underline{b} - \underline{a}\right](1/\delta - 1) & \text{if} \ \ \phi \in [0,1] \\ \left(\overline{b} - \underline{b}\right)(1-\alpha) & \text{if} \quad \phi > 1 \end{array} \right\}.$$

The comparative statics are interesting but intuitive. When N and S negotiate whether to sign the FTA, the equilibrium transfer from N to S will reflect the bargaining strength (α), the gains from trade, and the payoffs in the outside option (i.e., in autarky).

[24] As Grossman and Helpman (1995:683) write: "*if the home country were to increase its tariff on imports of some good and the foreign country increased its export subsidy by the same percentage amount, then the world price would fall so as to leave the domestic prices in each country unchanged.*"

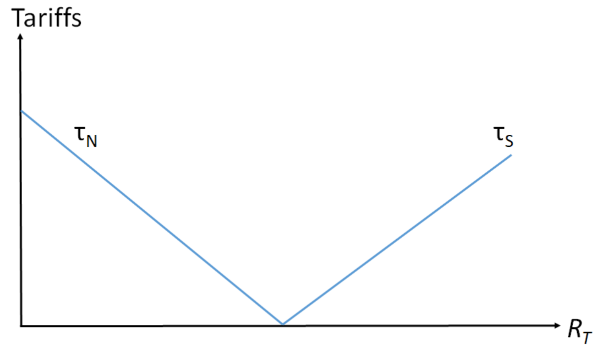

Figure 2: *In the FTA, the terms of trade are favorable to S if R is small.*

The larger R_T is, the larger is S's payoff from the FTA, but the smaller N's payoff is from the FTA when the FTA causes exploitation (i.e., when (3) holds). The transfers or tariffs can ensure that the parties will obtain comparable gains from the bilateral deal. It is thus intuitive that the transfer to S, $\tau_N e - \tau_S \overline{a} R_0$, must be strictly decreasing in R_T. Consequently, τ_N must decrease in R_T, or τ_S must increase in R_T (or both).

For the purpose of reallocating the gains, there is no need for both tariffs to be strictly positive, so we can, without loss of generality, let $\tau_N \tau_S = 0$. The corresponding tariffs are illustrated in Figure 2.

At first, one might think that as soon as the resource has been depleted, and the damage is sunk, then the level of R_T would be irrelevant for the tariffs. This is not true because the FTA that has already been negotiated is renegotiation proof. In particular, N has no interest in renegotiating the agreement in a way that harms N's terms of trade. Any pair of tariffs that ensure that the payoff pair is located on the Pareto frontier is renegotiation proof. The original level of R_T will influence where this allocation is located, even though the historical foundation for this allocation seems irrelevant later, once the resource is depleted.

The fact that all agreements on the Pareto frontier are renegotiation proof is taken advantage of in the next section.

IV. Contingent Trade Agreements

IV.A. Feasibility

With the tariffs introduced in Section III.E, the gains from trade can be distributed in arbitrary ways. Once the parties have agreed to trade, and the gains are allocated according to some pair (τ_N, τ_S), then every such allocation is renegotiation proof in the following sense: any change in $\overline{\tau}$ will harm and thus be vetoed by (at least) one of the parties.

However, we may not want to impose the restriction that the equilibrium allocation of gains, or the pair (τ_N, τ_S), must be constant. This section permits the parties to negotiate tariffs that are functions

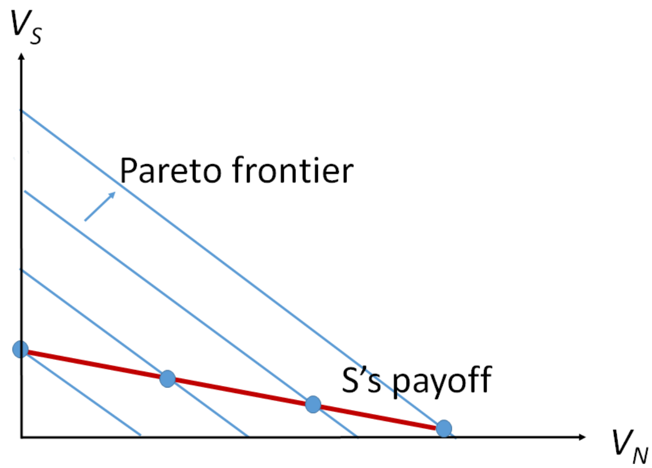

Figure 3: *When the resource is depleted, the Pareto frontier shifts, but S's payoff can decline.*

not only of R_T at the time, T, when they negotiate, but also of every smaller $R_t \in [0, R_T]$ that is imaginable for future dates (even off the equilibrium path).

Definition: *A* CTA, *negotiated at time T, specifies tariffs* $\tau_S(R_t; R_T)$ *and* $\tau_N(R_t; R_T)$, *that can depend on the current* R_t *as well as on* R_T, *unless the parties agree on different tariffs.*

For every R_t, the agreement must give S at last the same payoff as S can obtain in autarky. In addition, the tariff functions must be renegotiation proof. This requirement rules out, for example, a punishment strategy in which S will no longer be able to export if R_t has been reduced. As observed above, if $R_t = 0$, it is always (ex post) better for N and S to trade. The agreement is renegotiation proof if the equilibrium payoff pair is on the Pareto frontier for every R_t that is feasible at $t > T$.

A challenge is that, even if there is no exploitation after the agreement is signed, the sum of the gains from trade is $e + (\overline{a} - \underline{a}) X_t$, increasing in X_t. Thus, the more that has been depleted, the larger are the gains from trade that can be shared. If S receives a constant fraction of this cake, it should not be surprising that S faces a strong incentive to exploit.

But it does not need to be this way. Even if the total gain increases, the gain allocated to S can decrease, as illustrated in Figure 3. This situation is possible if S's tariff is a function that decreases in the stock, while N's tariff is a function that increases in the stock, as in Figure 4. If $\overline{\tau}$ increases sufficiently fast in R_t, then S has an incentive to conserve rather than to exploit.

Lemma 1. *S is willing to conserve and exploit* $x_t = 0 \forall t \geq T, R_t \leq R_T$, *if and only if:*

$$\begin{aligned} \partial \frac{\overline{\tau}(R_t; R_T)}{\partial R_t} &\geq \overline{a} + \underline{b} - c \; \forall R_t \leq R_T, \text{ where} \\ \overline{\tau}(R_t; R_T) &= \tau_N(R_t; R_T)\, e - \tau_S(R_t; R_T)\, \overline{a} X_t. \end{aligned} \tag{8}$$

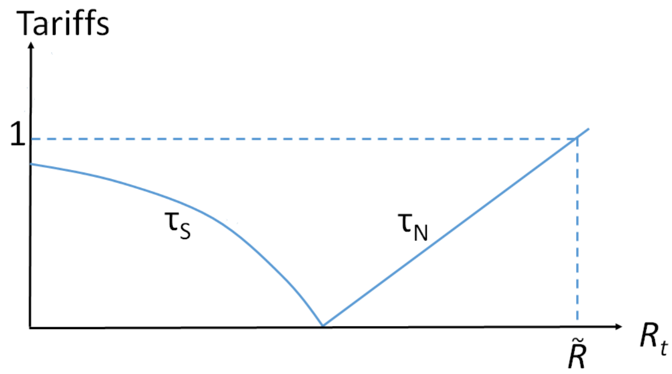

Figure 4: *In the CTA, the terms of trade are favorable to S if R is large.*

In essence, the CTA can specify a rule that allocates the gains from trade to S as long as S conserves, and to N if S exploits its resource. When (8) holds, this reallocation of the gains occurs so fast when R_t is reduced that S is better off conserving than depleting the resource.

IV.B. Equilibrium: Implementing The First Best

The lemma states that it is *possible* to design an agreement that motivates conservation. The next result states that N and S will indeed sign such an agreement in equilibrium, as long as the tariffs are permitted to be a function of the remaining stock, R_t. The intuition for this statement is simply that conservation is socially efficient, and thus both N and S can benefit from an agreement that motivates conservation when the parties can use side transfers (e.g., tariffs). N and S will share the total surplus according to their respective bargaining strengths.

The following proposition describes the tariff levels and x^D in the unique MPE in pure strategies. The tariffs are written as a function of the stock that exists at the time of negotiations, R_T. If future stocks are different, then the tariffs will also change in line with (8). Because (8) can be respected by a continuum of functions, the proposition does not specify exactly how steeply the tariffs will change if (off the equilibrium path) S extracted rather than conserved.

Proposition 6. *Consider a subgame starting at time T without a CTA. In equilibrium, N and S sign a CTA and implement the first-best outcome with $x_t = 0 \forall t \geq T$. The tariffs respect (8) and:*

$$\overline{\tau}\left(R_T; R_T\right) = \left\{ \begin{array}{lll} \alpha e - (1-\alpha)\left(\overline{a} - \underline{a}\right) R_0 + (1-\alpha)\left(\overline{a} - \underline{a}\right) R_T & \text{and } x^D = 0 & \text{if } \varphi < 0 \\ \alpha e - (1-\alpha)\left(\overline{a} - \underline{a}\right) R_0 + \frac{\underline{a} + \underline{b} - c}{\delta} R_T + \left(\overline{a} - \underline{a}\right) R_T & \text{and } x^D = \varphi R_T & \text{if } \varphi \in (0,1) \\ \alpha e - (1-\alpha)\left(\overline{a} - \underline{a}\right) R_0 + \left[(1-\alpha)\left(\overline{a} + \underline{b} - c\right) + \alpha d\right] R_T & \text{and } x^D = R_T & \text{if } \varphi > 1 \end{array} \right\}$$

$$\text{where } \varphi \equiv \frac{\delta\alpha\left(\overline{a} - \underline{a}\right) + \underline{a} + \underline{b} - c}{\delta\alpha\left(\overline{a} - \underline{a}\right) + \delta\alpha\frac{d + c - \overline{a} - \underline{b}}{1-\delta}}.$$

Interestingly, $\overline{\tau}\left(R_T; R_T\right)$ increases in R_T. The intuition is that if R_T is large, then N's benefit from an agreement that leads to conservation is also large. The larger benefit to N implies that, at the bargaining stage, N will accept transfers to S, or will accept to face tariffs when exporting to S. Thus, $\overline{\tau}(\cdot)$ increases in both arguments. The Appendix proves that $\overline{\tau}(R_T; R_T)$ permits (8) to hold for every $R_t \in [0, R_T]$.

It is always first best to conserve the entire resource in the simple model studied here. Full conservation is feasible by letting S obtain a large share of the gains from trade when R_t is large, but a smaller share when R_t is small. If R_T is very large, then the benefit to S might need to be larger than the total gains from trade. Above, this situation is permitted but it requires τ_N to be so large that N must subsidize its export for the producers to be willing to sell, or that $\tau_S < 0$, so that N subsidizes import. When τ_N is accompanied by an export subsidy in country N, or N subsidizes import, then there is no limit to how large $\overline{\tau}$ can be, and there is no limit to how much one can conserve. N agrees to the large $\overline{\tau}$, in equilibrium, because it prevents depletion.

IV.C. Without Subsidies: The Second Best

Export subsidies are rarely used in practice, however, and they are generally prohibited by the WTO. When export and import subsidies cannot be used, then the transfer from N to S is limited by the magnitude of the gains from trade. These gains therefore limit how much S can be persuaded to conserve by simply being allocated the gains from trade. The next result describes the upper boundary for how much it is possible to conserve without export and import subsidies.

PROPOSITION 7. *Consider a subgame at T without a CTA. Suppose subsidies are not available and, for simplicity, that $\alpha = 0$. In equilibrium, N offers a CTA that S immediately accepts:*
(i) The tariffs are in line with Proposition 6 and $x_t = 0$ for every $t \geq T$ if:

$$\frac{e + (\overline{a} - \underline{a}) R_0}{R_T} \geq \overline{a} + \underline{b} - c \Leftrightarrow \tag{9}$$

$$R_T \leq \widetilde{R} \equiv \frac{e + (\overline{a} - \underline{a}) R_0}{\overline{a} + \underline{b} - c} \text{ or } \overline{a} + \underline{b} < c.$$

(ii) When (9) fails, i.e., $R_T > \widetilde{R} > 0$, then, for every $t \geq T$,

$$x_t = \left(R_t - \widetilde{R}\right)\gamma, \text{ where } \gamma \equiv \frac{\overline{a} + \underline{b} - c}{\overline{a} + \underline{b} - c + \frac{\overline{b} - \underline{b}}{1 - \delta}} \in (0, 1), \tag{10}$$

and, on the equilibrium path $\tau_S = 0$ and $\tau_N = 1$.

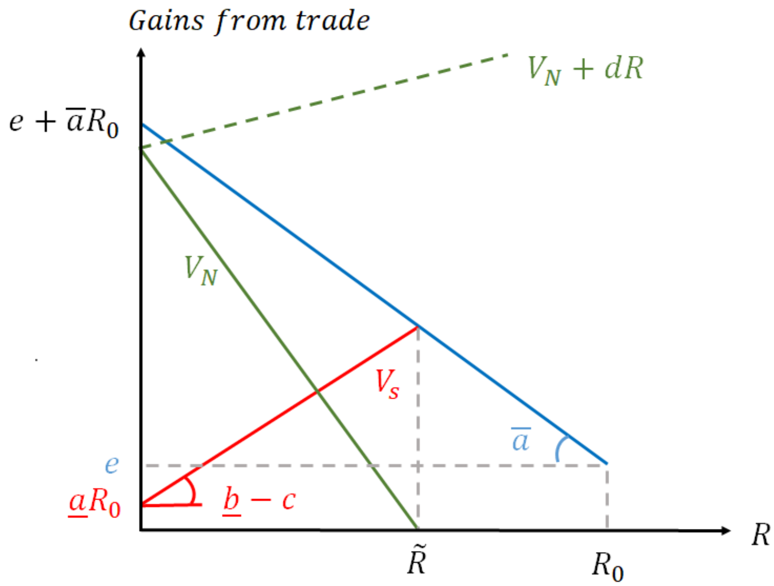

Figure 5: *The gains from trade (blue) decrease in R, but to motivate conservation, V_S (red line) must increase in R. Thus, at most $\widetilde{R}$ can be conserved – when export subsidies cannot be used. Although N's gains from trade (green) must decrease in R, also N is better off when R is large when the conservation value is taken into account.*

Eq. (9) can be rewritten as:

$$\frac{e + (\overline{a} - \underline{a})(R_0 - R_T)}{R_T} \geq \underline{a} + \underline{b} - c.$$

With this reformulation, part (i) in Proposition 7 states that if $\underline{a} + \underline{b} < c$, so that S does not want to reduce R_t in autarky, then the CTA can always ensure that there is no further exploitation, even if export subsidies cannot be used. If $\underline{a} + \underline{b} > c$, so that S would exploit in autarky, then the CTA can still motivate conservation, so that $R_t \in (0, R_0) \, \forall t$, but the amount of resource conservation, $\widetilde{R}$, is limited by the value of trade (Figure 5) because of the tariff ceiling (Figure 4). The larger the gains from trade are, the more it is possible to conserve in this way (i.e., the larger $\widetilde{R}$ is).

Part (ii) states that if $R_T > \widetilde{R} > 0$, the gains from trade are insufficient to motivate full conservation. The CTA can nevertheless be used to motivate a slower extraction rate. The speed at which S extracts can be reduced by allocating most of the gains from trade to S as long as S sticks to (10). If S exploits more, S will face higher tariffs. The larger are the gains from trade, $\overline{b} - \underline{b}$, the more it is possible to persuade S to conserve in each period.

A similar result holds if we face the restriction $\tau_N = 0$, so that the best N can offer S is free trade, and so that tariffs will not be used on the equilibrium path. The constraint $\tau_N = 0$ limits the cake available for S and thus the amount of the resource that can be conserved: (9) holds if just e is replaced by zero.

Remark on α. If $\alpha > 0$, the equilibrium x_t, when $R_T > \widetilde{R}$, would be larger than the x_t described by (10). When $R_T > \widetilde{R}$, the level of α is thus not only affecting the distribution of surplus, but also efficiency: a larger α is less efficient because S, then, requests a CTA that tolerates faster extraction.

IV.D. Comparison to the FTA

It is interesting to note that the mechanisms of the CTA are fundamentally different from those associated with the FTA. While the FTA is associated with more resource exploitation, the CTA is associated with conservation. The main result of this paper is that the CTA overturns *all* the results established for FTAs in Section III.

COROLLARY 1. *With a CTA, Propositions 1–5 are overturned:*
(1) The CTA leads to more conservation than in autarky.
(2) The CTA is more valuable when R_T is large.
(3) S conserves to maintain the most attractive CTA.
(4) In equilibrium, S conserves when the gains from trade are large and R_t is small.
(5) τ_N strictly increases or τ_S strictly decreases (or both) in R_T as well as in R_t.

Part (1) follows because if there is any conservation in autarky (i.e., if $\underline{a} + \underline{b} < c$), then (9) always holds and the CTA ensures full conservation, even when export subsidies are not available. Part (2) holds because in addition to exploiting the gains from trade, the CTA is beneficial because it ensures that R_T is conserved. The larger this stock is, the larger is the benefit of the CTA. Part (3), which reverses Proposition 3, is explained above. Part (4), reversing Proposition 4, follows when export or import subsidies cannot be used: Proposition 7 shows that S will exploit when R_t is large, but not when it is small, and the larger the gains from trade are, the more can be conserved. Part 5 follows when we compare Figures 2 and 4.

Remark 1 is also reversed.

REMARK 2 ON EXHAUSTIBILITY AND IRREVERSABILITY. The CTA can secure conservation *because* the resource is exhaustible. If, instead, R_t returned to R_0 in every period, or if R_t were not relevant, then it would not be credible that $\overline{\tau}$ would decrease if S extracted. If the anticipation of such a decrease could motivate S to conserve, then N would prefer to "restart the clock" after S had extracted. For this reason, the CTA would not be renegotiation proof if the resource were renewable.

Consequently, while the exhaustibility feature intensifies the conflict between trade and conservation under the FTA, it is this feature that makes the CTA effective and credible in motivating conservation.

IV.E. Commitment

Proposition 6 shows that the CTA can implement the first-best outcome. When subsidies cannot be used, Proposition 7 describes the amount that the CTA will conserve, in equilibrium. In this situation, one might wonder if there can be other designs that can be even better and that can motivate more conservation.

The answer to this question is no. Even if the parties could *commit* to policies that were conditioned on the resource stock and extraction levels, they would not be able to obtain higher payoffs than from the CTA, as long as export subsidies cannot be used. To prove this claim, it is sufficient to maximize the amount of conservation subject to the harshest punishment on S if S deviates from the plan. The harshest punishment is autarky. The autarky payoff is also what S obtains if S decides to *fully* deplete under the CTA, and full depletion is indeed a best response for S as long as *marginal* depletion (i.e., x_t marginally larger than (10)) is a best response. The last statement follows because S's payoff is linear in the size of the stock.

PROPOSITION 8. *The CTAs described by Propositions 6 and 7 implement the same outcome, and secure the same payoffs, as N and S would have achieved if they could commit to future policies as a function of the history.*

This result is important because it suggests that the CTA is not simply a design that improves marginally on the FTA, and from which N and S might be able to make further improvements. Instead, the CTA often implements the best N and S can hope for, even if they could have committed, although the CTA does not require them to be endowed with an ability to commit.

V. ROBUSTNESS

The model is simple for pedagogical reasons, but the main results are robust and can be derived in more general models. While the previous subsection showed that the CTA implements the outcome under commitment, Section V.1 shows how the CTA can be robust to unilateral requests to renege on the treaty. Section V.2 relaxes all linearity assumptions and prove that the main insights emphasized above continue to hold.

V.1. Binding vs. Non-binding Agreements – and Implementation

So far, the CTA has been praised as renegotiation proof because it distributes all gains from trade and, therefore, no other agreement is weakly better for both parties and strictly better for one. Renegotiation proofness is a natural requirement for a treaty that is binding, that is, if the agreement binds each party unless both countries agree to renegotiate the terms. If the agreement is non-binding, however, an individual country is free to tear it apart. If a country does so, the parties will find it in their interests to agree on another Pareto optimal allocation where S captures the fraction α of the time t surplus relative to autarky. I will say that an agreement is "renege proof" if at no $t \geq T$ or $R_t \in [0, R_T]$, no party can strictly benefit from leaving the agreement (e.g., in order to negotiate a new one).[25]

The CTA, described above, can be renege proof as well as renegotiation proof. As mentioned after Proposition 6, $\overline{\tau}(R_T; R_T)$ increases in R_T. When the CTA leads to conservation, then N will pay S, in

[25] In principle, it is not clear whether this threat of sticking with autarky should be taken seriously by the opponent. After all, the country reneging harms itself unless it soon wins the war of attrition it has just initiated. The credibility of this threat will depend on the details of the bargaining structure. This ambiguity has motivated a variety of definitions of renegotiation proofness that I do not intend to survey here. The above notion of renegotiation proofness is referred to as "the standard one" by Abreu et al. (1993) and Bergin and MacLeod (1993), and these authors propose concepts that are related to renege proofness. Mailath and Samuelson (2006) review the early literature on this topic.

equilibrium, for the conservation benefits that N enjoys from the CTA. If, instead, S has already depleted the resource, then S will obtain less favorable terms of trade because N benefits less from the CTA. This fact is often sufficient to motivate S to conserve the resource.

PROPOSITION 9. *Suppose N and S are free to renege on the CTA at any point in time.*
(i) Suppose export subsidies are available. If $c \notin (0, \alpha(\overline{a} - \underline{a}))$*, the equilibrium CTA is* $\overline{\tau}(R_t; R_T) = \overline{\tau}(R_t; R_t)$*, where* $\overline{\tau}(R_t; R_t)$ *is given by Proposition 6 if just* R_T *is replaced by the current* $R_t \leq R_T$*. If* $c \in (0, \alpha(\overline{a} - \underline{a}))$*, the equilibrium CTA is, instead:*

$$\overline{\tau}(R_t; R_t) = \alpha e - (1 - \alpha)(\overline{a} - \underline{a}) R_0 + (\overline{a} + \underline{b} - c) R_t. \tag{11}$$

In either case, the CTA implements the first best.
(ii) If export subsidies cannot be used, and $\alpha = 1$, *the equilibrium CTA is given by Proposition 7.*

When $\alpha \to 0$, it is always true that $c - \underline{a} - \underline{b} \notin (0, \alpha(\overline{a} - \underline{a}))$, and, thus, that the CTA characterized by Proposition 6 is renege proof. When N has all the bargaining power, it is intuitive that S cannot benefit from reneging. The CTA described by Proposition 7, where $\alpha = 0$ was assumed, is thus always renege proof.

Implementation. When the CTA is renege proof, it is straightforward to implement. It is sufficient to let parties negotiate $\overline{\tau}(R_t; R_T)$ in period T, and allow either party renege on the agreement in any subsequent period, and at any point in time during that period. Therefore, it is sufficient that $\overline{\tau}(R_t; R_T)$ holds through period T, and that it is sensitive to the stock that is relevant at the consumption stage at time T, that is, $R_{T+1} = R_T + x_T$. The $\overline{\tau}$ for subsequent periods can be negotiated later.

V.3. Non-Linear Payoff Functions

The most striking simplification has been that the payoffs have been linear in all variables, but all these linearity assumptions can be relaxed. If $\underline{a}(X^S)$ and $\overline{a}(X^N)$ are concave functions of the beef consumed in S and N, respectively, then decisions and outcomes can be an interior rather than corner solutions. Similarly, let $\underline{b}(x^S)$ and $\overline{b}(x^N)$ be concave functions of the quantity of timber that is consumed in S and N. With an extraction cost $\underline{c}(x)$, that is convex in x, depletion is likely to be gradual rather than instantaneous.

N's damage, $d(X)$, is now a convex function of X, and S may also face a convex environmental harm function, $h(X)$. These damages measure the present-discounted harm when the accumulated extracted quantity is, and forever will be, X. Note that $h(X)$ and $\underline{c}(x)$ together replace the linear cx in the model above. There, it was immaterial whether c reflected the current physical extraction cost or the environmental damage (or the sum of them), but here it is sensible to distinguish between the two.

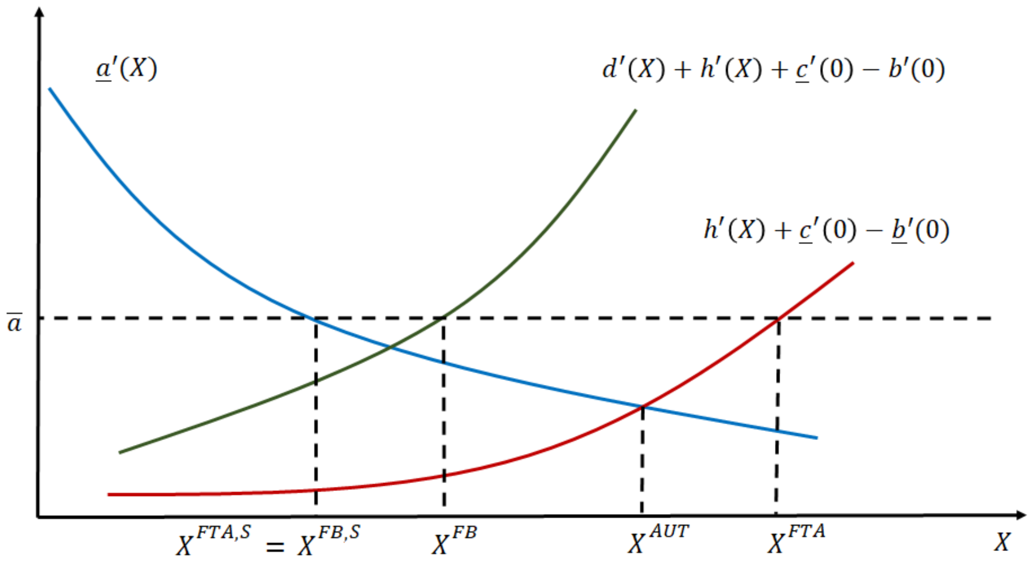

Figure 6: *The model can be generalized to non-linear functions and interior solutions.*

Figure 6 illustrates the steady-state levels of X in the first best, in autarky, and with free trade. For simplicity, the figure takes $\overline{a}$ as a constant, and applies the definition $b'(0) \equiv \max\left\{\underline{b}'(0), \overline{b}'(0)\right\}$.

The Appendix proves that the main result continues to hold with non-linear payoffs.[26] Deadweight losses and distortions because of tariffs can, once again, be avoided if the tariffs are accompanied with export subsidies. Without export subsidies, the distortions can be avoided if the tariff applies to the first $\chi_i > 0$ units of i's exported goods, but not for additional traded units, if just χ_i is set below the quantities that will be traded in equilibrium. With such two-stage tariff schedules, the tariffs will not influence the marginal decisions. A crucial difference between the non-linear case and the linear model of Section II is that the non-linear model might lead to interior solutions for the variables. To focus on this difference, I henceforth presume that all solutions are interior.

Proposition 10. *Suppose $\underline{a}(\cdot)$, $\overline{a}(\cdot)$, $\underline{b}(\cdot)$, and $\overline{b}(\cdot)$ are concave functions, while $d(\cdot)$, $\underline{c}(\cdot)$ and $h(\cdot)$ are convex functions, as explained above. Resource depletion is gradual, but the steady states are characterized as follows.*

(i) The resource stock is smaller with the free trade than in autarky:

$$R^{FTA} < R^{AUT}.$$

(ii) The resource is larger with a CTA than with an FTA:

$$R^{FTA} < R^{CTA}.$$

[26] Propositions 2 and 3, however, might not necessarily hold. If the marginal environmental value of the stock increases dramatically when the stock is depleted, then Proposition 2, above, might be reversed: An FTA can then be less attractive when the stock is threatened.

(iii) With export subsidies, $R^{CTA} = R^{FB}$*. Without export subsidies,* $R^{CTA} \in \left(R^{AUT}, R^{FB}\right]$ *is conserved when* $R^{AUT} < R^{FB}$*. If* $R^{CTA} < R^{FB}$*,* R^{CTA} *increases with the gains from trade:*

$$\frac{\partial R^{CTA}}{\partial e} > 0. \tag{12}$$

To complement (12), the Appendix shows that R^{CTA} also increases in the gains from trading beef. If N's marginal value is constant, $\overline{a}$, then $\partial R^{CTA}/\partial \overline{a} > 0$.

In general, R^{AUT} and R^{FB} cannot be ranked because N's damage decreases but N's demand decreases the first-best X.[27] The intuition for the other statements is similar to the intuition in the basic model, explained above. Conservation is motivated by the CTA because gains from trade are allocated to S when R_t is large. The larger these gains are, the more the CTA can conserve.

In practice, we rarely see export subsidies or the two-step tariff schedule suggested above. Tariffs can then lead to inefficiency losses, but for small tariffs, the losses are smaller than the conservation benefit. Furthermore, a large τ_S is used only off the equilibrium path. Under that threat, even free trade ($\tau_N (R_T; R_T) = \tau_S (R_T; R_T) = 0$) can motivate S to conserve $R_T > R^{FTA}$. If $R^{FB} > R^{AUT}$, then $R^{CTA} > R^{AUT}$. When $R^{CTA} < R^{FB}$, it holds, as before, that the larger are the gainst from trade, the more S is willing to conserve under the CTA.

VI. Conclusion

This paper starts by presenting a simple model that illustrates the negative two-way interaction between trade and environmental conservation. When the extracted resource can be exported as well as consumed domestically, extraction increases. Anticipating this situation, trade is worthwhile only if the remaining stock is, in any case, small. That fact, in turn, can motivate resource owners to reduce the stock to the point at which trade is acceptable for everyone. In equilibrium, the parties trade and the resource is exploited even when the status quo would be more efficient.

A contingent trade agreement reverses the negative results. Because every allocation on the Pareto frontier is renegotiation proof, conservation succeeds when exploitation worsens one's terms of trade. With the CTA, the South faces a low tariff when the stock is high and the North faces a low tariff low when the stock is small. Thus, the South conserves more than without any agreement, and more if the gains from trade are large.

These results are important because they show that although trade is often associated with resource depletion, such as deforestation, it *must* not be so. Clever agreements exploit the gains from trade and use the gains to motivate conservation rather than exploitation. This possibility should be kept in mind by scholars studying trade and environmental problems, but also by policymakers, public officials, and activists who struggle with how to balance trade and conservation. In fact, the analysis can shed light on

[27] From the Appendix, one can show that $X^{FB} < X^{AUT}$ if $\overline{a}'\left(X^{FB,N}\right) - \underline{a}'\left(X^{AUT}\right) < d'\left(X^{FB}\right)$

how one might implement the proposal in the recent non-paper by France and Netherland, mentioned in the Introduction. The analysis has uncovered that more can be conserved if the WTO allows for export subsidies, for instance. And CTAs are viable: verifiable measures of forest cover are available, thanks to satellite monitoring.[28] In India, the regional forest cover has, since 2015, been part of the central government's allocation of tax revenue to its 29 states (Busch and Mukherjee, 2018). Angelsen et al. (2018:51) elaborate on this policy and conclude that: "*This represents the first large-scale ecological fiscal transfers for forest cover, and could serve as a model for other countries.*"

The model above is tractable and future reseach can generalize it in several directions. In particular, the model has abstracted from politics and political constraints that must be respected for an agreement to be politically feasible. With lobbying, rent-seeking, elections, and legislative bargaining, the tariffs and subsidies will be influenced by many factors, beyond those studied here. These factors will be necessary to incorporate into the analysis in order to uncover how the CTA can be implemented effectively. With new research along these lines, we will continue to learn how trade can be exploited so as to motivate – rather than to discourage – the highest degree of conservation.

[28] As IPBES (2019, Ch. 6:56) states: "*The monitoring systems have been improved to the point of offering daily real-time data, constituting one of the most important tools for the fight against deforestation in Brazil.*"

Appendix

Proof of Proposition 1.

When $R_t = 0$, it is a best response for both parties to trade. When S obtains the fraction α of the total gains from trade in addition to S's autarky payoff, then:

$$\begin{aligned} V_S^{FTA}(0) &= V_S^{AUT}(0) + \alpha\left[e + (\overline{a} - \underline{a}) R_0\right] = \underline{a}R_0 + \alpha e + \alpha(\overline{a} - \underline{a}) R_0, \text{ because} \\ V_S^{AUT}(0) &= \underline{a}R_0. \end{aligned}$$

When $R_t > 0$ and the parties trade, then S solves:

$$V_S^{FTA}(R_t) \equiv \max_{x_t^{FTA} \in [0, R_t]} (1-\delta)(X_t + x_t)\overline{a} + \overline{b}x_t - cx_t + \delta V_S(R_{t+1}),$$

which implies, regardless of whether $V_S(R_{t+1}) = V_S^{FTA}(R_{t+1})$ or $V_S(R_{t+1}) = V_S^{AUT}(R_{t+1})$, that

$$x_t^{FTA}(R_t) = \left\{ \begin{array}{c} 0 \text{ if } \overline{a} + \overline{b} \leq c \\ x_t \in [0, R_t] \text{ if } \overline{a} + \overline{b} = c \\ R_t \text{ if } \overline{a} + \overline{b} > c \end{array} \right\}.$$

QED

Proof of Proposition 2.

When (3) holds, autarky leads to the total payoff $V^{AUT}(R_t) \equiv V_S^{AUT}(R_t) + V_N^{AUT}(R_t) = (R_0 - R_t)\underline{a}$, while the FTA leads to depletion and the total payoff

$$V^{FTA}(R_t) \equiv V_S^{FTA}(R_t) + V_N^{FTA}(R_t) = e + \overline{a}R_0 + \left(\overline{b} - c - d\right) R_t, \tag{13}$$

which is larger if:

$$e + \overline{a}R_0 + \left(\overline{b} - c - d\right) R_t > (R_0 - R_t)\underline{a} \Leftrightarrow (4).$$

QED

Proof of Proposition 3.

Even if there is no trade at time t, S is strictly better off with $x_t = R_t$ than with $x_t = 0$ and autarky forever if:

$$(1-\delta) R_0 \underline{a} + (\underline{b} - c) R_t + \delta V_S^{FTA}(0) > \underline{a}(R_0 - R_t) \Leftrightarrow (6).$$

QED

Proof of Proposition 4.

Trade equilibria. Consider, first, the situation in which N and S trade at every $R \leq R_t$.

The bargaining surplus. Let $x_t^D = \eta R_t$, with $\eta \in [0,1]$, measure S's extraction after disagreement. (In principle, η can be a function of R_t.) The proof proceeds by deriving the fixed point where S's best response, given η, coincides with η.

Given η, if the parties have disagreed at t, but expect to agree at $t+1$, the sum of disagreement payoffs is:

$$\begin{aligned} V^{DIS}(R_t) &= (1-\delta)(X_t + \eta R_t)\underline{a} - (c + d - \underline{b})\eta R_t + \delta V^{FTA}((1-\eta) R_t) \\ &= (1-\delta)(R_0 - R_t)\underline{a} - \left[(1-\delta)\left(c + d - \underline{a} - \overline{b}\right) + \left(\overline{b} - \underline{b}\right)\right]\eta R_t + \delta\left[e + \overline{a}R_0 + \left(\overline{b} - c - d\right) R_t\right]. \end{aligned}$$

The total gains from agreeing at t, (13), minus the above disagreement payoff, $V^{DIS}(R_t)$, is:

$$\begin{aligned} \Delta_R &= (1-\delta)\left[e + \overline{a}R_0 + \left(\overline{b} - c - d\right) R_t\right] - (1-\delta) X_t \underline{a} + \left[(1-\delta)\left(c + d - \underline{a} - \overline{b}\right) + \left(\overline{b} - \underline{b}\right)\right]\eta R_t \\ &= (1-\delta)\left[e + (\overline{a} - \underline{a}) R_0 + \left(\underline{a} + \overline{b} - c - d\right) R_t\right] + \left[(1-\delta)\left(c + d - \underline{a} - \overline{b}\right) + \left(\overline{b} - \underline{b}\right)\right]\eta R_t. \end{aligned} \tag{14}$$

S's best response. Consider, first, the case in which $x_t^D = 0$ is among S's best responses. Even after disagreement at t, N and S will agree at $t+1$ and, then, S can expect S's default payoff plus the fraction α of (14):

$$V_S^{FTA}(R_t) = (1-\delta)\underline{a}X_t + \delta V_S^{FTA}(R_t) + \alpha\Delta_R = \underline{a}X_t + \alpha\frac{\Delta_R}{1-\delta}. \quad (15)$$

Thus, at the disagreement stage at time t, S's payoff is (when S conserves):

$$(1-\delta)\underline{a}X_t + \delta V_S^{FTA}(R_t) = \underline{a}X_t + \delta\alpha\frac{\Delta_R}{1-\delta}.$$

If, instead, S depletes, then S obtains:

$$\underline{a}X_t + (\underline{a}+\underline{b}-c)R_t + \delta\alpha\left[e + (\overline{a}-\underline{a})R_0\right].$$

By comparison, S is better off conserving if:

$$\delta\alpha\frac{\Delta_R}{1-\delta} \geq (\underline{a}+\underline{b}-c)R_t + \delta\alpha\left[e+(\overline{a}-\underline{a})R_0\right] \Leftrightarrow$$

$$\delta\alpha\left(\underline{a}+\overline{b}-c-d\right)R_t + \delta\alpha\left[\left(c+d-\underline{a}-\overline{b}\right)+\left(\frac{\overline{b}-\underline{b}}{1-\delta}\right)\right]\eta R_t \geq (\underline{a}+\underline{b}-c)R_t \Leftrightarrow$$

$$\eta \geq \phi, \text{ where } \phi \equiv \frac{\underline{a}+\underline{b}-c-\delta\alpha\left(\underline{a}+\overline{b}-c-d\right)}{\delta\alpha\left[\left(c+d-\underline{a}-\overline{b}\right)+\left(\frac{\overline{b}-\underline{b}}{1-\delta}\right)\right]} = \frac{c+d-\underline{a}-\overline{b}-\frac{c-\underline{a}-\underline{b}}{\delta\alpha}}{c+d-\underline{a}-\overline{b}+\frac{\overline{b}-\underline{b}}{1-\delta}}.$$

Consequently, at the disagreement stage, S's best response is to conserve if η is large, and to exploit if η is small. S's best response is a decreasing (step-)function of η and there is a unique fixed point.

The fixed point. If $\phi > 1$, it is never a best response for S to conserve because $\eta > \phi$ contradicts $\eta \in [0,1]$. In equilibrium, then, $\eta = 1$. If $\phi \leq 0$, it is always a best response to conserve because $\eta \geq \phi$ always holds: thus, $\eta = 0$. If $\phi \in (0,1]$, the fixed point is $\eta = \phi$.

Non-trade equilibria. Note that if $x^D > 0$, then S strictly benefits from agreeing at t, instead of disagreeing. When $\alpha \in (0,1)$, this requires that N, too, strictly benefits from agreeing at t, instead of agreeing at $t+1$. Consequently, there cannot be an equilibrium where N and S do not agree at t, if $x^D > 0$. Therefore, to end by considering equilibria in which N and S do not trade, it must be that $x^D = 0$. For this to be an equilibrium, it must be that (4) fails, so that N and S do not benefit from trading, and that (6) fails, so that S does not want to exploit in order to obtain an FTA. *QED*

Proof of Proposition 5.

When S is willing to conserve after a disagreement, S's payoff is given by (15). At the same time, S's payoff is also given by (7). The two are equal if:

$$\overline{\tau} + \overline{a}R_0 + \left(\overline{b}-c\right)R_t = \underline{a}X_t + \alpha\frac{\Delta_R}{1-\delta} \Leftrightarrow$$

$$\begin{aligned}
\overline{\tau} &= \left(c-\underline{a}-\overline{b}\right)R_t - (\overline{a}-\underline{a})R_0 \\
&\quad +\alpha\left[e+(\overline{a}-\underline{a})R_0+\left(\underline{a}+\overline{b}-c-d\right)R_t+\left[c+d-\underline{a}-\overline{b}+\frac{\overline{b}-\underline{b}}{1-\delta}\right]\eta R_t\right] \\
&= \alpha e-(1-\alpha)(\overline{a}-\underline{a})R_0 - R_t\left[\alpha d-(1-\alpha)\left(c-\underline{a}-\overline{b}\right)-\alpha\eta\left[c+d-\underline{a}-\overline{b}+\frac{\overline{b}-\underline{b}}{1-\delta}\right]\right].
\end{aligned}$$

We have three cases to consider. If $\phi \leq 0$, $\eta = 0$, so:

$$\overline{\tau} = \alpha e - (1-\alpha)(\overline{a}-\underline{a})R_0 - R_t\left[\alpha d - (1-\alpha)\left(c-\underline{a}-\overline{b}\right)\right].$$

If $\phi = \frac{d-\widehat{d}}{c+d-\overline{b}-\underline{a}+\frac{\overline{b}-\underline{b}}{1-\delta}} \in (0,1]$, $\eta = \phi$, and then:

$$\begin{aligned}
\overline{\tau} &= \alpha e-(1-\alpha)(\overline{a}-\underline{a})R_0 - R_t\left[\alpha d-(1-\alpha)\left(c-\underline{a}-\overline{b}\right)-\alpha\left(d-\widehat{d}\right)\right] \\
&= \alpha e-(1-\alpha)(\overline{a}-\underline{a})R_0 - R_t\left[\frac{1-\delta}{\delta}(c-\underline{b}-\underline{a})+\left(\overline{b}-\underline{b}\right)\right].
\end{aligned}$$

If $\phi > 1$, S's best response is $x^D = R_t$, so:

$$V_S^{FTA}(R_t) = \underline{a}R_0 + (\underline{b} - c) R_t + \alpha \left[e + (\overline{a} - \underline{a}) R_0 + \overline{b} - \underline{b} \right].$$

This payoff equals (7) if:

$$\begin{aligned}
\overline{\tau} + \overline{a}R_0 + (\overline{b} - c) R_t &= \underline{a}R_0 + (\underline{b} - c) R_t + \alpha \left[e + (\overline{a} - \underline{a}) R_0 + \overline{b} - \underline{b} \right] \Leftrightarrow \\
\overline{\tau} &= \alpha e - (1 - \alpha)(\overline{a} - \underline{a}) R_0 - (1 - \alpha)(\overline{b} - \underline{b}) R_t.
\end{aligned}$$

QED

Proof of Lemma 1.

If N and S have signed the agreement, and S conserves at R_t, then S obtains:

$$V_S^{CTA}(R_t) = \overline{a}X_t + \overline{\tau}(R_t; R_T).$$

Alternatively, if S extracts x_t to enjoy $V_S^{CTA}(R_t - x_t)$, in addition to $(\underline{b} - c) x_t$, where $\overline{\tau}(R_t - x_t; R_T)$ depends on the new level of converted land, then S obtains:

$$V_S^{CTA}(R_t - x_t) + (\underline{b} - c) x_t,$$

which is worse if

$$V_S^{CTA}(R_t) - V_S^{CTA}(R_t - x_t) \geq (\underline{b} - c) x_t,$$

which always holds if and only if:

$$\partial \frac{V_S^{CTA}(R_t)}{\partial R_t} \geq \underline{b} - c, \tag{16}$$

and, note that with $\partial \frac{V_S^{CTA}(R_t)}{\partial R_t} = \partial \frac{\overline{\tau}(R_t; R_T)}{\partial R_t} - \overline{a}$, we obtain (8). Since $\overline{\tau}(0; R_T) \geq -(\overline{a} - \underline{a}) R_0$, to respect that S benefits from trade, we can integrate over (8) to get:

$$\overline{\tau}(R_t; R_T) \geq (\overline{a} + \underline{b} - c) R_t - (\overline{a} - \underline{a}) R_0. \tag{17}$$

QED

Proof of Proposition 6.

The bargaining surplus. Again, let $x_t^D = \eta R_t$, with $\eta \in [0, 1]$, measure S's extraction after disagreement. (In principle, η can be a function of R_t.) The proof proceeds by deriving the fixed point where S's best response, given η, coincides with η.

If the parties have disagreed at T, but expect to agree at $T + 1$, the sum of disagreement payoffs is:

$$\begin{aligned}
V^{DIS}(R_T) &= (1 - \delta)(X_T + \eta R_T) \underline{a} - (c + d - \underline{b}) \eta R_T + \delta V^{CTA}((1 - \eta) R_T) \\
&= (1 - \delta)(X_T + \eta R_T) \underline{a} - (c + d - \underline{b}) \eta R_T + \delta \left[e + \overline{a}(X_T + \eta R_T) \right].
\end{aligned}$$

The total gains from agreeing at T, $e + \overline{a}X_T$, minus the above disagreement payoff, $V^{DIS}(R_T)$, is:

$$\begin{aligned}
\Delta_R &= (1 - \delta) \left[e + (\overline{a} - \underline{a}) X_T \right] - (1 - \delta) \eta R_T \underline{a} + (c + d - \underline{b}) \eta R_T - \delta \overline{a} \eta R_T \\
&= (1 - \delta) \left[e + (\overline{a} - \underline{a}) X_T \right] + (c + d - \underline{a} - \underline{b} - \delta (\overline{a} - \underline{a})) \eta R_T.
\end{aligned} \tag{18}$$

S's best response. Consider, first, the case in which $x_t^D = 0$ is among S's best responses. Even after disagreement at T, N and S will agree at $T + 1$ and, then, S can expect S's default payoff plus the fraction α of (18):

$$V_S^{CTA}(R_T) = (1 - \delta) \underline{a}X_T + \delta V_S^{CTA}(R_T) + \alpha \Delta_R = \underline{a}X_T + \alpha \frac{\Delta_R}{1 - \delta}. \tag{19}$$

Thus, after disagreeing at time T, S's payoff is (when S conserves):

$$(1 - \delta) \underline{a}X_T + \delta V_S^{CTA}(R_T) = \underline{a}X_T + \delta \alpha \frac{\Delta_R}{1 - \delta}.$$

If, instead, S depletes, then S obtains:

$$\underline{a}X_T + (\underline{a} + \underline{b} - c) R_T + \delta\alpha [e + (\overline{a} - \underline{a}) R_0].$$

By comparison, S is better off conserving if:

$$\begin{aligned} \delta\alpha\tfrac{\Delta_R}{1-\delta} \geq (\underline{a} + \underline{b} - c) R_T + \delta\alpha [e + (\overline{a} - \underline{a}) R_0] \Leftrightarrow \\ -\delta\alpha (\overline{a} - \underline{a}) R_T + \delta\alpha \tfrac{(c+d-\underline{a}-\underline{b}-\delta(\overline{a}-\underline{a}))}{1-\delta}\eta R_T \geq \quad (\underline{a} + \underline{b} - c) R_T \Leftrightarrow \\ \eta \geq \varphi, \text{ where } \varphi \equiv \tfrac{1-\delta}{\delta\alpha} \tfrac{\underline{a}+\underline{b}-c+\delta\alpha(\overline{a}-\underline{a})}{c+d-\underline{a}-\underline{b}-\delta(\overline{a}-\underline{a})}. \end{aligned}$$

Consequently, at the disagreement stage, S's best response is to conserve if η is large, and to exploit if η is small. S's best response is a decreasing (step-)function of η and there is a unique fixed point.

The fixed point. If $\varphi > 1$, it is never a best response for S to conserve because $\eta > \varphi$ would contradict $\eta \in [0, 1]$. In equilibrium, then, $\eta = 1$. If $\varphi \leq 0$, it is always a best response to conserve because $\eta \geq \varphi$ always holds: thus, $\eta = 0$. If $\varphi \in (0, 1]$, the fixed point is $\eta = \varphi$.

Tariffs. When S is willing to conserve after a disagreement, S's payoff is given by (19). At the same time, because the CTA motivates conservation, this payoff is also given by $\overline{a}X_T + \overline{\tau}$. The two are equal if:

$$\overline{a}X_T + \overline{\tau} = \underline{a}X_T + \alpha\frac{\Delta_R}{1-\delta} \Leftrightarrow \overline{\tau} = \alpha e - (1-\alpha)(\overline{a} - \underline{a}) X_T + \alpha \left(\frac{c + d - \underline{a} - \underline{b} - \delta (\overline{a} - \underline{a})}{1-\delta}\right) \eta R_T.$$

We have three cases to consider. If $\varphi \leq 0$, $\eta = 0$, so:

$$\overline{\tau} = \alpha e - (1-\alpha)(\overline{a} - \underline{a}) X_T.$$

If $\varphi \equiv \frac{1-\delta}{\delta\alpha} \frac{\underline{a}+\underline{b}-c+\delta\alpha(\overline{a}-\underline{a})}{c+d-\underline{a}-\underline{b}-\delta(\overline{a}-\underline{a})} \in (0, 1]$, $\eta = \varphi$, and then:

$$\overline{\tau} = \alpha e - (1-\alpha)(\overline{a} - \underline{a}) R_0 + \left[\frac{\underline{a} + \underline{b} - c}{\delta} + \overline{a} - \underline{a}\right] R_T.$$

If $\varphi > 1$, S's best response after disagreement is $x^D = R_t$, so:

$$V_S^{CTA}(R_T) = \underline{a}R_0 + (\underline{b} - c) R_T + \alpha [e + (\overline{a} - \underline{a}) X_T + (c + d - \underline{a} - \underline{b}) R_T].$$

This payoff equals $\overline{a}X_T + \overline{\tau}$ if:

$$\overline{\tau} = \alpha e - (1-\alpha)(\overline{a} - \underline{a}) R_0 + [\alpha d + (1-\alpha)(\overline{a} + \underline{b} - c)] R_T.$$

It is easy to check that, in all three cases, $\overline{\tau}$ satisfies (17) when $R_t = R_T$. This implies that with the above equilibrium $\overline{\tau}$, it is possible to find a function $\overline{\tau}(R_t; R_T)$ that satisfies (8) for every possible future $R_t \in [0, R_T]$. *QED*

Proof of Proposition 7.

(i) If $\alpha = 0$, N will minimize $V_S^{CTA}(R_T)$ s.t. the conditions that S conserves: i.e., (16), and that S accepts to trade, which requires $V_S^{CTA}(0) \geq \underline{a}R_0$. When these two bind:

$$V_S^{CTA}(R_T) = \underline{a}R_0 + (\underline{b} - c) R_T.$$

With the largest tariffs on e, and on x_t, and the lowest on S's beef, $V_S^{CTA}(R_T) \leq e + \overline{a}(R_0 - R_T)$, so:

$$\begin{aligned} \underline{a}R_0 + (\underline{b} - c) R_T &\leq e + \overline{a}(R_0 - R_T) \\ R_T &\leq \widetilde{R} \equiv \frac{e + (\overline{a} - \underline{a}) R_0}{\overline{a} + \underline{b} - c}. \end{aligned}$$

(ii) This part takes advantage of the fact that if the tariffs can depend on the current stock, they can depend on the previous stock, and thus also on the difference in the stocks (i.e., x_t). When $x(R_t)$ is tolerated by the agreement, then S's continuation value on the equilibrium path is:

$$V_S^{CTA}(R_t) = x(R_t)\left(\overline{b}(1-\tau_b)-c\right)+(1-\delta)(1-\tau_S)(X_t+x(R_t))\overline{a}+(1-\delta)\tau_N e+\delta V_S^{CTA}(R_t-x(R_t)),$$

where τ_b is the tariff on timber. When $\alpha = 0$, N ensures that any deviation leaves S with its autarky payoff, $\underline{a}R_0 + (\underline{b}-c)R_t$. N's problem is to minimize $x(R_t)$, subject to $V_S^{CTA}(R_t) \geq \underline{a}R_0 + (\underline{b}-c)R_t$. Thus, N sets $\tau_b = \tau_S = 0$ and $\tau_N = 1$, and $V_S^{CTA}(R) = V_S^{AUT}(R)$. With this, $x(R_t)$ is given by:

$$\underline{a}R_0 + (\underline{b}-c)R_t = x(R_t)\left(\overline{b}-c\right) + (1-\delta)(X_t+x(R_t))\overline{a} + (1-\delta)e + \delta\left(\underline{a}R_0 + (\underline{b}-c)(R_t - x(R_t))\right) \Leftrightarrow$$

$$\begin{aligned} x(R_t) &= \frac{\underline{a}R_0 + (\underline{b}-c)R_t - (1-\delta)X_t\overline{a} - (1-\delta)e - \delta\underline{a}R_0 - \delta(\underline{b}-c)R_t}{\overline{b}-c+(1-\delta)\overline{a}-\delta(\underline{b}-c)} \\ &= \frac{(\overline{a}+\underline{b}-c)R_t - (\overline{a}-\underline{a})R_0 - e}{\overline{a}+\underline{b}-c+\frac{\overline{b}-\underline{b}}{1-\delta}} = \left(R_t - \widetilde{R}\right)\frac{\overline{a}+\underline{b}-c}{\overline{a}+\underline{b}-c+\frac{\overline{b}-\underline{b}}{1-\delta}}. \end{aligned}$$

Note that, when the inequality binds, $x(R_t)$ approaches zero when $R_t \downarrow \widetilde{R}$. *QED*

Proof of Proposition 8.

The proof follows the reasoning in the text. *QED*

Proof of Proposition 9.

(i) If either party can walk away from the CTA and negotiate a new CTA, under the threat of autarky, then the equilibrium will be characterized by $\overline{\tau}(R_t;R_T) = \overline{\tau}(R_t;R_t)$, where $\overline{\tau}(R_t;R_t)$ is given by Proposition 6 (where R_T is replaced by the current $R_t \leq R_T$), if the CTA leads to conservation. This $\overline{\tau}(R_t;R_t)$ increases in R_t, and it might satisfy Lemma 1. If we compare the derivatives $\partial\overline{\tau}(R_T;R_T)/\partial R_T$ for the three cases in Proposition 6 with the requirement (8), it is easy to verify that (8) is satisfied whenever $c - \underline{a} - \underline{b} \notin (0, \alpha(\overline{a}-\underline{a}))$. In this case, therefore, the CTA $\overline{\tau}(R_t;R_T) = \overline{\tau}(R_t;R_t)$, where $\overline{\tau}(R_t;R_t)$ is given by Proposition 6, is both renegotiation proof and renege proof, and it is the equilibrium treaty when the parties negotiate. No party will ever want to renege on this CTA, not even off the equilibrium path. If $c - \underline{a} - \underline{b} \in (0, \alpha(\overline{a}-\underline{a}))$, then $\overline{\tau}(R_T;R_T)$, in Proposition 6, does not increase sufficiently fast in R_T to motivate conservation. When S can renege on the CTA, the CTA will be renege proof and it will motivate S to conserve only if $\overline{\tau}$ satisfies (8) for every $R_t \in [0, R_T]$. By integrating (8) from $R_t = 0$ to $R_t = R_T$, we can see that N must agree on the following $\overline{\tau}$:

$$\overline{\tau}(R_t;R_t) = \overline{\tau}(0;0) + (\overline{a}+\underline{b}-c)R_t.$$

Further, for $\overline{\tau}(0;0)$ to be renege proof, it must be given by Proposition 6 when $R_T = 0$. When we combine the two terms, we get (11).This CTA is renege proof, it implements the first best, and it is larger than the one in Proposition 6 if and only if $c - \underline{a} - \underline{b} \in (0, \alpha(\overline{a}-\underline{a}))$.

(ii) The CTA described by Proposition 7 is renege proof by construction because N has all bargaining power and must respect S's participation constraint at every $R_t \in [0, R_T]$. *QED*

Proof of Proposition 10.

In line with the text in Section V.3, $\underline{a}\left(X_t^S\right)$ and $\overline{a}\left(X_t^N\right)$ are S's and N's present-discounted value from forever consuming the produce X_t^S and X_t^N, respectively, where $X_t \equiv X_t^S + X_t^N$. Similarly, $\underline{b}\left(x_t^S\right)$ and $\overline{b}\left(x_t^N\right)$ are values of consuming the extracted quantities, where $x_t \equiv x_t^N + x_t^S$. In contrast to the linear model, it is now necessary to distinguish between S's extraction cost, $\underline{c}(x_t)$, and S's environmental harm, $h(X_{t+1})$. The damage for N is $d(X_{t+1})$.

(i) *Autarky.* In autarky, the first-order condition for the steady state is:

$$\underline{a}'\left(X^{AUT}\right) + \underline{b}'(0) = \underline{c}'(0) + h'\left(X^{AUT}\right), \tag{20}$$

and, because the second-order condition holds given the assumptions,

$$\underline{a}'(X) + \underline{b}'(0) < \underline{c}'(0) + h'(X) \text{ if } X > X^{AUT}. \tag{21}$$

S's steady-state payoff is $\underline{a}\left(R_0 - R^{AUT}\right) - h\left(R_0 - R^{AUT}\right)$.

FTA. With an FTA, the steady state is:

$$\begin{aligned}\underline{a}'\left(X^{FTA,S}\right) + \overline{a}'\left(X^{FTA,N}\right) + b'(0) &= \underline{c}'(0) + h'\left(X^{FTA}\right), \text{ where} \\ b'(0) &\equiv \max\left\{\underline{b}(0), \overline{b}'(0)\right\},\end{aligned}$$

and where $\underline{a}'\left(X^{FTA,S}\right) = \overline{a}'\left(X^{FTA,N}\right)$, in equilibrium. It follows that $X^{FTA} > X^{AUT}$ as long as $\overline{a}'(0) > \underline{a}'\left(X^{FTA,S}\right)$. The latter inequality must hold if, in equilibrium, $X^{FTA,N} > 0$.

(ii) *First Best.* Note that when the first best is interior, then it requires:

$$\underline{a}'\left(X^{FB,S}\right) + \overline{a}'\left(X^{FB,N}\right) + b'(0) = \underline{c}'(0) + h'\left(X^{FB}\right) + d'\left(X^{FB}\right),$$

and $\underline{a}'\left(X^{FB,S}\right) = \overline{a}'\left(X^{FB,N}\right)$ when both $X^{FB,S} > 0$ and $X^{FB,N} > 0$. When we compare to free trade, we can see that $X^{FB} < X^{FTA}$ when $d'\left(X^{FB}\right) > 0$.

CTA. To simplify, the remainder of this proof limits attention to the case in which N has all the bargaining power: $\alpha = 0$.

Let $V_S^{CTA}(R_t)$ be S's continuation payoff with full conservation of R_t under the CTA. N's willingness to pay (i.e., the price in terms of the numeraire good) for beef is $\overline{a}'\left(X^{CTA,N}\right)$, and thus we can write

$$\begin{aligned}V_S^{CTA}(R_t) &= \underline{a}\left(X^{CTA,S}\right) + (1-\tau_S)\overline{a}'\left(X^{CTA,N}\right)X^{CTA,N} + \tau_N e - h\left(X^{CTA}\right), \\ &= \underline{a}\left(X^{CTA,S}\right) + \overline{a}'\left(X^{CTA,N}\right)X^{CTA,N} - h\left(X^{CTA}\right) + \overline{\tau}, \text{ where} \\ \overline{\tau} &= \tau_N e - \tau_S \overline{a}'\left(X^{CTA,N}\right)X^{CTA,N} \text{ and } X^{CTA,N} = R_0 - R_t - X^{CTA,S}.\end{aligned}$$

Extracting $x_t > 0$ is not beneficial to S if:

$$\begin{aligned}\arg\max_{x_t \in [0,R_t]} V_S^{CTA}(R_t - x_t) + \underline{b}(x_t) - \underline{c}(x_t) &= 0 \Rightarrow \\ \partial\frac{V_S^{CTA}(R_t)}{\partial R_t} \geq \underline{b}'(0) - \underline{c}'(0).\end{aligned} \tag{22}$$

When N offers a CTA that guarantees conservation of R^{CTA}, and $\alpha = 0$, then, for any R^{CTA}, N minimizes $\overline{\tau}$ or, equivalently, $V_S^{CTA}(R_t)$, subject to the incentive constraint, (22), and subject to S's participation constraint, $V_S^{CTA}(R_t) \geq V_S^{AUT}(R_t)$.[29]

When $R_t \leq R^{AUT}$, then S's participation constraint can bind, i.e., $V_S^{CTA}(R_t) = V_S^{AUT}(R_t)$, without violating the incentive constraint, (22): it follows from (21) that $\partial V_S^{AUT}(R_t)/\partial R_t = -\underline{a}'\left(X^{AUT}\right) + h'\left(X^{AUT}\right) > \underline{b}'(0) - \underline{c}'(0)$ when $R^t < R^{AUT}$. It also follows that R^{AUT} can be conserved by the CTA, if $X^{FB} < X^{AUT}$.

When $R_t > R^{AUT}$, however, (22) requires that $V_S^{CTA}(R_t) > V_S^{AUT}(R_t)$ when $R_t > R^{CTA}$. In particular:

$$\begin{aligned}V_S^{CTA}(R_t) &= V_S^{CTA}\left(R_t^{AUT}\right) + \int_{R^{AUT}}^{R_t} \partial\frac{V_S^{CTA}(R)}{\partial R}dR \geq V_S^{AUT}\left(R_t^{AUT}\right) + \int_{R^{AUT}}^{R_t} \underline{b}'(0) - \underline{c}'(0)\, dR \\ &= \underline{a}\left(X^{AUT}\right) + \left[\underline{b}'(0) - \underline{c}'(0)\right]\left(R_t - R^{AUT}\right).\end{aligned}$$

When $\alpha = 0$, the inequality will bind, and

$$V_S^{CTA}(R_t) = \underline{a}\left(X^{AUT}\right) - h\left(X^{AUT}\right) + \left[\underline{b}'(0) - \underline{c}'(0)\right]\left(R_t - R^{AUT}\right).$$

With this, any R_t can be conserved by the CTA if just $\overline{\tau}$ is sufficiently large. With (such) transfers at the bargaining stage, N and S will thus agree on a CTA in which $R^{CTA} = R^{FB}$, since the sum of payoffs is largest in the first-best outcome.

The required transfer is the following. The steady state payoff for S can be written as:

$$V_S^{CTA}(R_t) = \underline{a}\left(X^{CTA,S}\right) + \overline{a}'\left(X_N^{CTA}\right)X_N^{CTA} - h\left(X^{CTA}\right) + \overline{\tau}.$$

[29] Here, $V_S^{AUT}(R_t)$ reflects S's payoff in autarky and not simply in the autarky's steady state.

The two expressions for $V_S^{CTA}(R_t)$ are equal if and only if:

$$\begin{aligned}\overline{\tau} &= \underline{a}\left(X^{AUT}\right)+h\left(X^{CTA}\right)-h\left(X^{AUT}\right)+\left[\underline{b}'(0)-\underline{c}'(0)\right]\left(R_t-R^{AUT}\right) \\ &\quad -\underline{a}\left(R_0-R_t-X^{CTA,N}\right)-\overline{a}\left(X^{CTA,N}\right) \qquad (23)\\ &= \underline{a}\left(X^{AUT}\right)+h\left(X^{CTA}\right)-h\left(X^{AUT}\right)+\left[\underline{b}'(0)-\underline{c}'(0)\right]\left(R_t-R^{AUT}\right) \\ &\quad -\underline{a}\left(X^{CTA,S}\right)-\overline{a}\left(R_0-R_t-X^{CTA,S}\right). \qquad (24)\end{aligned}$$

(iii) From (23),

$$\begin{aligned}\frac{\partial\overline{\tau}}{\partial R^{CTA}} &= \underline{a}'\left(R_0-R_t-X^{CTA,N}\right)+\underline{b}'(0)-\underline{c}'(0)-h'\left(X^{CTA}\right) \\ &= \underline{a}'\left(X^{CTA,S}\right)+\underline{b}'(0)-\underline{c}'(0)-h'\left(X^{CTA}\right)>0 \text{ when } X^{CTA,S}<X^{AUT}. \qquad (25)\end{aligned}$$

The last claim follows from (20) and (21). When export subsidies cannot be used, $\overline{\tau}\leq e$, and the largest R_t that can be conserved is given by $\overline{\tau}=e$. Combined with (25), we get:

$$\frac{\partial R^{CTA}}{\partial e}=\frac{1}{\underline{a}'\left(X^{CTA,S}\right)+\underline{b}'(0)-\underline{c}'(0)-h'\left(X^{CTA}\right)}>0.$$

In addition, when $\overline{a}$ is a constant, then we can differentiate (24) w.r.t. dR^{CTA} and $d\overline{a}$ to get:

$$\begin{aligned}-h'\left(X^{CTA}\right)dR^{CTA}+\left[\underline{b}'(0)-\underline{c}'(0)+\overline{a}\right]dR^{CTA}-\left(R_0-R^{CTA}-X^{CTA,S}\right)d\overline{a}=0\Leftrightarrow \\ \frac{dR^{CTA}}{d\overline{a}}=\frac{R_0-R^{CTA}-X^{CTA,S}}{\overline{a}+\underline{b}'(0)-\underline{c}'(0)-h'\left(X^{CTA}\right)}=\frac{X^{CTA,N}}{\overline{a}+\underline{b}'(0)-\underline{c}'(0)-h'\left(X^{CTA}\right)}>0.\end{aligned}$$

QED

References

Abman, Ryan, and Lundberg, Clark (2020): "Does Free Trade Increase Deforestation? The Effects of Regional Trade Agreements," Journal of the Association of Environmental and Resource Economists 7(1): 35-72.

Abrego, Lisandro; Perroni, Carlo; Whalley, John, and Wigle, Randall M. (2001): "Trade and Environment: Bargaining Outcomes from Linked Negotiations," Review of International Economics 9(3): 414-28.

Abreu, Dilip; Pearce, David, and Stacchetti, Ennio (1993): "Renegotiation and symmetry in repeated games," Journal of Economic Theory 60(2): 217-40.

Aghion, Philippe; Antràs, Pol, and Helpman, Elhanan (2007): "Negotiating free trade," Journal of International Economics 73(1): 1-30.

Angelsen, Arild; Hermansen, Erlend A. T.; Rajajo, Raoni and van der Hoff, Richard (2018): "Results-based payment: Who should be paid, and for what?" Ch. 4 in A. Angelsen, C. Martius, V. de Sy, A.E. Duchelle, A.M. Larson, and Pham T.T. (Eds.), Transforming REDD+: Lessons and new directions. Center for International Forestry Research (CIFOR).

Antràs, Pol, and Staiger, Robert W. (2012): "Trade Agreements and the Nature of Price Determination," American Economic Review 102(3): 470-76.

Antweiler, Werner; Copeland, Brian R., and Taylor, M. Scott (2001): "Is Free Trade Good for the Environment?" American Economic Review 91 (4): 877-908.

Bagwell, Kyle, and Staiger, Robert W. (1999): "An economic theory of GATT," American Economic Review 89(1): 215-48.

Bagwell, Kyle, and Staiger, Robert W. (2004): The economics of the world trading system. Cambridge: MIT press.

Bagwell, Kyle, and Staiger, Robert W. (2011): "What do trade negotiators negotiate about? Empirical evidence from the World Trade Organization," American Economic Review 101(4): 1238-73.

Bagwell, Kyle, and Staiger, Robert W. (2016): "The design of trade agreements," in K. Bagwell and R. Staiger (Eds.), Handbook of Commercial Policy 1: 435-529.

Barbier, Edward B. (2000): "Links between economic liberalization and rural resource degradation in the developing regions," Agricultural Economics 23(3): 299-310.

Barrett, Scott (1997): "The strategy of trade sanctions in international environmental agreements," Resource and Energy Economics 19(4): 345-61.

Battaglini, Marco, and Harstad, Bård (2020): "The Political Economy of Weak Treaties," Journal of Political Economy 128(7): 2653-89.

Bergin, James, and MacLeod, Bentley (1993): "Efficiency and renegotiation in repeated games," Journal of Economic Theory 61(1): 42-73.

Brander, James A., and Taylor, Scott M. (1997): "International trade and open access renewable resources: the small open economy case," Canadian Journal of Economics 30(3): 526-52.

Brander, James A., and Taylor, Scott M. (1998): "Open access renewable resources: Trade and trade policy in a two-country model," Journal of International Economics 44(2): 181-209.

Brollo, Fernanda; Nannicini, Tommaso; Perotti, Roberto, and Tabellini, Guido (2013): "The Political Resource Curse," American Economic Review 103(5): 1759-96.

Burgess, Robin; Costa, Francisco; and Olken, Ben (2019): "The Brazilian Amazon's Double Reversal of Fortune." Mimeo, MIT.

Burgess, Robin, Hansen, Matthew, Olken, Ben, Potapov, Peter and Sieber, Stefanie (2012): "The Political Economy of Deforestation in the Tropics," Quarterly Journal of Economics 127(4): 1707-54.

Busch, Jonah, and Mukherjee, Anit (2018): "Encouraging State Governments to protect and restore forests using ecological fiscal transfers: India's tax revenue distribution reform," Conservation Letters 11(2): 11(2): 1–10..

Bustos, Paula; Garber, Gabriel, and Ponticelli, Jacopo (2020): "Capital accumulation and structural transformation," Quarterly Journal of Economics 135(2): 1037-94.

Caselli, Francesco; Koren, Miklos; Lisicky, Milan and Tenreyro Silvana (2020): "Diversification Through Trade," Quarterly Journal of Economics 135(1): 449-502.

Caselli, Francesco, and Michaels, Guy (2013): "Do Oil Windfalls Improve Living Standards? Evidence from Brazil " American Economic Journal: Applied Economics 5(1): 208-38.

Chichilnisky, Graciela (1994): "North-south trade and the global environment," American Economic Review 84(4): 851-74.
Copeland, Brian R. (2000): "Trade and environment: policy linkages," Environment and Development Economics 5(4): 405-32.
Copeland, Brian R., and Taylor, M. Scott (1994): "North-South Trade and the Environment," Quarterly Journal of Economics 109(3): 755-87.
Copeland, Brian R. and Taylor, M. Scott (1995): "Trade and Transboundary Pollution," American Economic Review 85 (4): 716-37.
Copeland, Brian R., and Taylor, M. Scott (2004): "Trade, growth and the environment," Journal of Economic Literature 42(1): 7-71.
Dasgupta, Partha; Eastwood, Robert, and Heal, Geoffrey (1978): "Resource management in a trading economy," Quarterly Journal of Economics 92(2): 297-306.
Elliott, Joshua; Foster, Ian; Kortum, Samuel; Munson, Todd; Perez Cervantes, Fernando; and Weisbach, David (2010): "Trade and Carbon Taxes," American Economic Review: Papers & Proceedings 100(2): 465-69.
Faria, Weslem Rodrigues, and Almeida, Alexandre Nunes (2016): "Relationship between openness to trade and deforestation: Empirical evidence from the Brazilian Amazon," Ecological Economics 121: 85-97.
Franklin Jr, Sergio L., and Pindyck, Robert S. (2018): "Tropical Forests, Tipping Points, and the Social Cost of Deforestation," Ecological Economics 153: 161-71.
Grossman, Gene M. (2016): "The purpose of trade agreements," in K. Bagwell and R. Staiger (Eds.), Handbook of Commercial Policy 1: 379-434.
Grossman, Gene M., and Helpman, Elhanan (1995): "Trade wars and trade talks," Journal of Political Economy 103(4): 675-08.
Grossman, Gene M.; McCalman, Phillip, and Staiger, Robert W.(2020): "The" new" economics of trade agreements: From trade liberalization to regulatory convergence?" Econometrica, forthcoming.
Harstad, Bård (2016): "The market for conservation and other hostages," Journal of Economic Theory 166: 124-51.
Harstad, Bård (2020): "The Conservation Multiplier," Mimeo.
Harstad, Bård, and Mideksa, Torben (2017): "Conservation Contracts and Political Regimes," Review of Economic Studies 84(4): 1708-34.
Hillman, Arye L., and Van Long, Ngo (1983): "Pricing and depletion of an exhaustible resource when there is anticipation of trade disruption," Quarterly Journal of Economics 98(2): 215-33.
Hoel, Michael (1996): "Optimal international trade agreements and dispute settlement procedures," Journal of Public Economics 59(1): 17-32.
Horn, Henrik, and Mavroidis, Petros C. (2014): "Multilateral environmental agreements in the WTO: Silence speaks volumes," International Journal of Economic Theory 10(1): 147-66.
Horstmann, Ignatius J.; Markusen, James R., and Robles, Jack (2005): "Issue Linking in Trade Negotiations: Ricardo Revisited or No Pain No Gain," Review of International Economics 13(2): 185-204.
IPBES (2019): Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. E. S. Brondizio, J. Settele, S. Díaz, and H. T. Ngo (editors). IPBES secretariat, Bonn, Germany.
Jayachandran, Seema (2013): "Liquidity Constraints and Deforestation: The Limitations of Payments for Ecosystem Services," American Economic Review, Papers and Proceedings 103(3): 309-13.
Karp, Larry; Sacheti, Sandeep, and Zhao, Jinhua (2001): "Common Ground Between Free-Traders and Environmentalists," International Economic Review 42(3): 617-48.
Ludema, Rodney D. (2001): "Optimal international trade agreements and dispute settlement procedures," European Journal of Political Economy 17(2): 355-76.
Ludema, Rodney D., and Mayda, Anna Maria (2013): "Do terms-of-trade effects matter for trade agreements? Theory and evidence from WTO Countries," Quarterly Journal of Economics 128(4): 1837-93.
Maggi, Giovanni (2014): "International trade agreements," in G. Gopinath, E. Helpman, and K. Rogoff (Eds.), Handbook of International Economics 4: 317-90. Elsevier.
Maggi, Giovanni (2016): "Issue linkage," in K. Bagwell and R. Staiger (Eds.), Handbook of Commercial Policy 1: 513-64.

Maggi, Giovanni, and Ossa, Ralph (2020): "Are Trade Agreements Good For You?" NBER Working Paper No. 27252.
Maggi, Giovanni, and Rodriguez-Clare, Andres (1998): "The value of trade agreements in the presence of political pressures," Journal of Political Economy 106(3): 574-601.
Maggi, Giovanni, and Staiger, Robert W. (2015): "Optimal design of trade agreements in the presence of renegotiation," American Economic Journal: Microeconomics 7(1): 109-43.
Mailath, George J., and Samuelson, Larry (2006): Repeated games and reputations: long-run relationships. Oxford University Press.
Markusen, James R. (1975): "International externalities and optimal tax structures," Journal of International Economics 5(1): 15-29.
Markusen, James R. (1981): "Trade and the gains from trade with imperfect competition," Journal of International Economics 11(4): 531-51.
Maskin, Eric, and Tirole, Jean (2001): "Markov Perfect Equilibrium: I. Observable Actions," Journal of Economic Theory 100(2): 191-219.
Matsuyama, Kiminori (1990): "Perfect Equilibria in a Trade Liberalization Game," American Economic Review 80(3): 480-92.
Nordhaus, William (2015): "Climate clubs: Overcoming free-riding in international climate policy," American Economic Review 105(4): 1339-70.
Pendrill, Florence; Persson, U. Martin; Godar, Javier; Kastner, Thomas; Moran, Daniel; Schmidt, Sarah, and Wood, Richard (2019): "Agricultural and forestry trade drives large share of tropical deforestation emissions," Global Environmental Change 56: 1-10.
Pethig, Rüdiger (1976): "Pollution, welfare, and environmental policy in the theory of comparative advantage," Journal of Environmental Economics and Management 2(3): 160-69.
Souza-Rodrigues, Eduardo (2019): "Deforestation in the Amazon: A unified framework for estimation and policy analysis," Review of Economic Studies 86(6): 2713-44.
World Bank (2019): "Trade integration as a pathway to development?" Semiannual Report of the Latin America and Caribbean Region.