Online Appendix A. Comparison to Commitment

Above, the parties are not endowed with the ability to commit: The CTA tariffs have been required to be renegotiation proof. Nevertheless, 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 (11)) is a best response.[9]

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 here 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.

Online Appendix B. 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).[10]

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 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:*

$$\overline{\tau}(R_t; R_t) = \alpha e - (1 - \alpha)(\overline{a} - \underline{a}) R_0 + (\overline{a} + \underline{b} - c) R_t. \tag{20}$$

In either case, the CTA implements the first best.

[9] The last statement follows because S's payoff is linear in the size of the stock.

[10] 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.

(ii) If export subsidies cannot be used, and $\alpha = 1$, the equilibrium CTA is given by Proposition 7.

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 (20).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*

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.

Online Appendix C. Remarks 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.

The CTA, in contrast, 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.

References for Online Appendix

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

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

Mailath, George J., and Samuelson, Larry (2006): Repeated games and reputations: long-run relationships. Oxford University Press.