Multilateral Tariff Cooperation under Reciprocity

Costas Hadjiyiannis[*], Doruk Iris[†], and Chrysostomos Tabakis[‡]

Abstract

This paper explores trade policy and trade agreements between governments with reciprocal preferences. Such governments respond kindly to actions that are perceived to be kinder than expected and they retaliate when others are perceived to engage in unkind behavior. We compare the results with the standard case where governments maximize national welfare and we find that reciprocity has a significant impact on trade agreements. We show that if the "fair" tariff is higher than a critical level, governments with reciprocal preferences can achieve lower cooperative tariffs in a dynamic infinitely repeated tariff game. However, if governments believe that "fair" tariffs are too low then it is possible that they end up with higher cooperative tariffs. The intuition gained from these results could provide insights for the failure of the Doha round.

Keywords: Multilateral Tariff Cooperation, Reciprocity, Trade Agreements

JEL classification: F13, F15, D03

[*]University of Cyprus
[†]Universidade Nova de Lisboa
[‡]Universidade Nova de Lisboa

1. Introduction

There is plenty of evidence that individuals very often do not simply maximize their own self-interested utility. There is a growing body of literature in behavioral and experimental economics showing that in many cases individuals care to some degree about fairness and equity incorporating these considerations into their preferences. One of the most famous cases we often teach undergraduate students in game theory classes is the "split the dollar" game. In this game a player is asked to propose a way to split one dollar with another player. If the second player agrees, then they both get what they agreed upon. Otherwise they get nothing. If players were simply maximizing their own welfare, the first player should get everything and the second agree to getting next to nothing. However, this is very rarely the outcome of these experiments with individuals opting for a more "equitable" or "fair" split.

This and many other experiments confirm that individuals very often exhibit reciprocal preferences, that is, they respond kindly to actions that are perceived to be kinder than expected and they retaliate when others are perceived to engage in unkind behavior. If individuals behave this way, it is possible that governments have similar preferences for a number of reasons. From the literature on political economy, we know that governments maximize a weighted average of the welfare of individuals. If individuals have reciprocal preferences then it is possible that governments also have reciprocal preferences. In simpler terms, governments want to be popular to get re-elected and to do so they have to satisfy the median voter. If the median voter has reciprocal preferences then the government is going to mirror these preferences. One could, also, make the argument that governments are run by individuals and they themselves might have reciprocal preferences.

In this paper we set out to explore the implications of reciprocal preferences on trade policy and trade agreements without making any claims about how common or widespread these preferences are. That is a question for empirical work and is outside the scope of this paper. To this end, we develop a dynamic game in which countries with reciprocal preferences attempt to maintain multilateral tariff cooperation. In our context, a reciprocal country places a positive (negative) weight on the welfare of a trading partner if it expects the latter to apply an import tariff that is smaller (larger) than the one it perceives as "fair." Our modeling of fairness and reciprocity follows Segal and Sobel (2007). Moreover, we maintain the usual assumption in the trade-agreement literature that binding commitments cannot be made at the international level and countries are therefore limited to cooperative agreements that are self-enforcing. In such a setting, a country will choose to adhere to the cooperative tariff path as long as the onetime gains it could achieve by unilaterally deviating from the agreed-upon trade policy do not outweigh the future welfare losses due to the trade war a defection would ignite.

We show that provided the tariffs that are perceived as "fair" by different countries are not too low, reciprocity facilitates multilateral tariff cooperation. In particular, we demonstrate that as compared with the standard dynamic game with self-interested countries, in our game with reciprocal countries (i) any given cooperative tariff can be more easily sustained (i.e., the critical (minimum) discount factor required is smaller); and (ii) the most cooperative equilibrium tariff is also lower. However, if the tariffs that are perceived as "fair" are below a critical threshold, the effect of reciprocity on multilateral trade cooperation is ambiguous. In other words, for such low "fair" tariffs, there are cases where self-interested countries can support lower cooperative tariffs in equilibrium than reciprocal ones. We show that these results are robust to alternative punishment schemes and apply to very general social-welfare functions. Our findings could therefore shed some light on the failure of the Doha round. The success of the previous trade rounds and the overall economic environment at the time may have raised expectations too high (i.e., lowered the "fair" tariffs substantially), hindering the efforts for further trade liberalization and more multilateral trade cooperation.

2. The Model

We assume the world consists of two countries, A and B, that trade two goods, a and b. In order to keep our analysis as simple as possible, we do not rigorously examine the production process in the two countries. Instead, we simply assume that country J is endowed with 1 unit of good $-j$ and zero units of good j, where $J \in \{A, B\}$ and $j \in \{a, b\}$. Therefore, good j is potentially exported from country $-J$ to country J. On the consumption side, we maintain the assumption that demand functions are symmetric across countries and goods, and that the demand for good j is independent of the price of good $-j$. More specifically, the demand for good j in country J is given by $D\left(P_j^J\right)$, where P_j^J is good j's price in country J. We make the standard assumptions that $D\left(P_j^J\right)$ is strictly positive on some bounded interval $\left[0, \overline{P}_j^J\right)$ with $D\left(P_j^J\right) = 0$ for $P_j^J \geq \overline{P}_j^J$, and that $D'\left(P_j^J\right) < 0$

for $P_j^J \in \left[0, \overline{P}_j^J\right)$.

In each period, the countries simultaneously select specific import tariffs so as to maximize their individual *reciprocal* welfare. The tariffs are picked with perfect information as to all past tariff choices. Let τ^J be country J's import tariff (on good j). It follows that:

$$P_j^J = P_j^{-J} + \tau^J.$$

Moreover, the export supply functions are given by:

$$X_j^{-J}\left(P_j^{-J}\right) = 1 - D\left(P_j^{-J}\right).$$

Market clearing then requires:

$$D\left(P_j^J\right) = X_j^{-J}\left(P_j^{-J}\right).$$

2.1. Static Game

We now compare the static Nash equilibrium of our model with the one that would emerge in a game with self-interested countries. Let the self-interested static game be denoted by $\Gamma^S(SW)$, where SW is the welfare function of a self-interested country. Furthermore, let

$\Gamma^R(RW, w, \tau_f)$ denote our reciprocal static game, where RW is the welfare function of a reciprocal country, w is the weight function determining the weight a country places on its trading partner's self-interested welfare SW, and $\tau_f \equiv \left(\tau_f^J, \tau_f^{-J}\right)$ with τ_f^J being the τ^J country $-J$ deems as "fair."

The welfare of reciprocal country J is given by:

$$RW^J = \int_{P_j^J}^{\frac{\alpha}{\beta}} D\left(P\right) dP + \int_{P_{-j}^J}^{\frac{\alpha}{\beta}} D\left(P\right) dP + P_{-j}^J + \tau^J X_j^{-J} + \gamma w^J(\tau^{-J}, \tau_f^{-J}) SW^{-J}$$

where $w^J(\tau^{-J}, \tau_f^{-J})$ is the weight country J places on the self-interested welfare of country $-J$. The latter is given by:

$$SW^{-J} = \int_{P_{-j}^{-J}}^{\frac{\alpha}{\beta}} D\left(P\right) dP + \int_{P_j^{-J}}^{\frac{\alpha}{\beta}} D\left(P\right) dP + P_j^{-J} + \tau^{-J} X_{-j}^J.$$

The preferences for reciprocity are given by:

$$w^J(\tau^{-J}, \tau_f^{-J}) \begin{cases} > 0 \text{ if } \tau_f^{-J} > \tau^{-J} \\ = 0 \text{ if } \tau^{-J} = \tau_f^{-J} \\ < 0 \text{ otherwise} \end{cases}, \tag{1}$$

that is, country J places a positive weight on the trading partner's self-interested welfare when the latter's tariff is less than τ_f^{-J}, country J places zero weight on the partner's self-interested welfare when the latter's tariff is exactly equal to τ_f^{-J}, and finally, country J places a negative weight on its partner's self-interested welfare when the latter's tariff exceeds τ_f^{-J}.

These conditions capture the fact that a country with reciprocal welfare cares about the intentions of its trading partner. In particular, the first condition expresses positive reciprocity. If country J expects the tariff of its partner to be smaller than its own perception of how high the fair tariff is, then country J is willing to sacrifice some of its self-interested welfare to increase the partner's self-interested welfare. On the other hand, the third condition expresses negative reciprocity. When country J expects the tariff of its trading partner to be higher than country J's perception of what the fair tariff is, then country J is willing to sacrifice some of its own self-interested welfare to reduce the partner's self-interested welfare.

We assume that the weight function is twice differentiable in both arguments, and is decreasing with the other country's tariff (i.e., $\frac{\partial w^J(\tau^{-J},\tau_f^{-J})}{\partial \tau^{-J}} < 0$) and increasing with the fair-tariff perception (i.e., $\frac{\partial w^J(\tau^{-J},\tau_f^{-J})}{\partial \tau_f^{-J}} > 0$). We now have that $\frac{\partial RW^J}{\partial \tau^J} = -D\left(P_j^J\right)\frac{\partial P_j^J}{\partial \tau^J} + X_j^{-J} + \tau^J \frac{\partial X_j^{-J}}{\partial \tau^J} + \gamma w^J(\tau^{-J}, \tau_f^{-J})\frac{\partial SW^{-J}}{\partial \tau^J}$. Noting that $\frac{\partial SW^{-J}}{\partial \tau^J} = -D\left(P_j^{-J}\right)\frac{\partial P_j^{-J}}{\partial \tau^J} + \frac{\partial P_j^{-J}}{\partial \tau^J} = \left(1 - D\left(P_j^{-J}\right)\right)\frac{\partial P_j^{-J}}{\partial \tau^J}$, we then have that $\frac{\partial RW^J}{\partial \tau^J} = -D\left(P_j^J\right)\frac{\partial P_j^J}{\partial \tau^J} + X_j^{-J} + \tau^J \frac{\partial X_j^{-J}}{\partial \tau^J} + \gamma w^J(\tau^{-J}, \tau_f^{-J})\left(1 - D\left(P_j^{-J}\right)\right)\frac{\partial P_j^{-J}}{\partial \tau^J}$.

The cross-partial derivative of the welfare function of reciprocal country J with respect to its own tariff and country $-J$'s tariff is positive since:

$$\frac{\partial RW^J}{\partial\tau^J\partial\tau^{-J}} = \underbrace{-\underbrace{\frac{\partial D\left(P_j^J\right)}{\partial\tau^{-J}}}_{=0}\underbrace{\frac{\partial P_j^J}{\partial\tau^J}}_{\geq 0} - \underbrace{D\left(P_j^J\right)}_{\geq 0}\underbrace{\frac{\partial^2 P_j^J}{\partial\tau^J\partial\tau^{-J}}}_{=0} + \underbrace{\frac{\partial X_j^{-J}}{\partial\tau^{-J}}}_{=0} + \tau^J\underbrace{\frac{\partial^2 X_j^{-J}}{\partial\tau^J\partial\tau^{-J}}}_{=0}}_{=0}$$

$$+\gamma\left(\begin{array}{c} \underbrace{\underbrace{\frac{\partial w^J(\tau^{-J},\tau_f^{-J})}{\partial\tau^{-J}}}_{\leq 0}\underbrace{\left(1-D\left(P_j^{-J}\right)\right)}_{\geq 0}\underbrace{\frac{\partial P_j^{-J}}{\partial\tau^J}}_{\leq 0}}_{\geq 0} \\ \underbrace{+\,w^J(\tau^{-J},\tau_f^{-J})\left(1-D\left(P_j^{-J}\right)\right)\underbrace{\frac{\partial^2 P_j^{-J}}{\partial\tau^J\partial\tau^{-J}}}_{=0} - w^J(\tau^{-J},\tau_f^{-J})\frac{\partial P_j^{-J}}{\partial\tau^J}\underbrace{\frac{\partial D\left(P_j^{-J}\right)}{\partial\tau^{-J}}}_{=0}}_{=0} \end{array}\right),$$

meaning that in our model, the choice variables are strategic complements, that is, country J's incremental returns from increasing its own tariff are increasing in its partner's tariff.

Furthermore, the cross-partial derivative of the welfare function of reciprocal country J with respect to its own tariff and its perception of the fair tariff of country $-J$ is negative since:

$$\frac{\partial RW^J}{\partial\tau^J\partial\tau_f^{-J}} = \underbrace{-\underbrace{\frac{\partial D\left(P_j^J\right)}{\partial\tau_f^{-J}}}_{=0}\underbrace{\frac{\partial P_j^J}{\partial\tau^J}}_{\geq 0} - \underbrace{D\left(P_j^J\right)}_{\geq 0}\underbrace{\frac{\partial^2 P_j^J}{\partial\tau^J\partial\tau_f^{-J}}}_{=0} + \underbrace{\frac{\partial X_j^{-J}}{\partial\tau_f^{-J}}}_{=0} + \tau^J\underbrace{\frac{\partial^2 X_j^{-J}}{\partial\tau^J\partial\tau_f^{-J}}}_{=0}}_{=0}$$

$$+\gamma\left(\begin{array}{c}\underbrace{\underbrace{\frac{\partial w^J(\tau^{-J},\tau_f^{-J})}{\partial\tau_f^{-J}}}_{\geq 0}\underbrace{\left(1-D\left(P_j^{-J}\right)\right)}_{\geq 0}\underbrace{\frac{\partial P_j^{-J}}{\partial\tau^J}}_{\leq 0}}_{\leq 0}\\ \underbrace{+w^J(\tau^{-J},\tau_f^{-J})\left(1-D\left(P_j^{-J}\right)\right)\underbrace{\frac{\partial^2 P_j^{-J}}{\partial\tau^J\partial\tau_f^{-J}}}_{=0}-w^J(\tau^{-J},\tau_f^{-J})\frac{\partial P_j^{-J}}{\partial\tau^J}\underbrace{\frac{\partial D\left(P_j^{-J}\right)}{\partial\tau_f^{-J}}}_{=0}}_{=0}\end{array}\right).$$

We assume that $\tau^J \in \Theta^J \subset \mathcal{R}_+$ for any country J, where Θ^J is a compact interval.

Lemma 1: *If for any* J *(i)* Θ^J *is a compact interval in* $\mathcal{R}_+$, *(ii)* RW^J is twice continuously differentiable on Θ^J, *and (iii)* $\frac{\partial RW^J}{\partial\tau^J\partial\tau^{-J}} \geq 0$, *then* $\Gamma^R(RW, w, \tau_f)$ *is a supermodular game.*

Proof: All conditions of Theorem 4 in Milgrom and Roberts (1990) are satisfied in our model.

From Milgrom and Roberts (1990) we know that if $\Gamma^R(RW, w, \tau_f)$ is a supermodular game, then there exist largest and smallest serially undominated strategies for each player J, $\overline{\tau}^J$ and $\underline{\tau}^J$. Moreover, the strategy profiles $\overline{\tau} \equiv \left(\overline{\tau}^J, \overline{\tau}^{-J}\right)$ and $\underline{\tau} \equiv \left(\underline{\tau}^J, \underline{\tau}^{-J}\right)$ are pure-strategy Nash equilibrium profiles. Thus, the existence of a Nash equilibrium for our stage game is guaranteed. Our next result shows how countries' perceptions of the fair tariffs of their trading partners influence the extremal equilibrium tariffs of this static game.

Proposition 1: If $n = 2$, $\Gamma^R(RW, w, \tau_f)$ is a supermodular game, and RW^J has decreasing differences in (τ^J, τ_f^{-J}) for any J, then the largest and the smallest Nash equilibria of $\Gamma^R(RW, w, \tau_f)$, i.e., $\overline{\tau}_{NR} \equiv \left(\overline{\tau}_{NR}^J, \overline{\tau}_{NR}^{-J}\right)$ and $\underline{\tau}_{NR} \equiv \left(\underline{\tau}_{NR}^J, \underline{\tau}_{NR}^{-J}\right)$, are nonincreasing functions of τ_f.

Proof: Immediate from Theorem 6 in Milgrom and Roberts (1990).

For the rest of the paper, we focus on the generic equilibrium $\tau_{NR} \equiv \left(\tau_{NR}^J, \tau_{NR}^{-J}\right)$, since all our arguments apply for both the largest and the smallest Nash equilibrium.

Corollary 1: *If* $\Gamma^R(RW, w, \tau_f)$ *is a supermodular game and for any* J, RW^J has decreasing differences in (τ^J, τ_f^{-J}) *and* $\tau_{NS}^{-J} \geq \tau_f^{-J}$, *then* $\tau_{NR} \geq \tau_{NS}$ and for all J, $RW^J(\tau_{NR}, \tau_f^{-J}) \leq SW^J(\tau_{NS})$.

Proof: The stage game $\Gamma^S(SW)$ can be obtained from the stage game $\Gamma^R(RW, w, \tau_f)$ by setting $\gamma = 0$. Thus, if $\Gamma^R(RW, w, \tau_f)$ is a supermodular game, so is $\Gamma^S(SW)$. This means that $\Gamma^S(SW)$ also has a smallest and a largest Nash equilibrium in pure strategies. Denote the generic equilibrium by τ_{NS}. If for any J $\tau_f^{-J} = \tau_{NS}^{-J}$, then trivially $\tau_{NR} = \tau_{NS} = \tau_N$ and for all J, $RW^J(\tau_N, \tau_f^{-J}) = SW^J(\tau_N)$ since $w^J(\tau_N^{-J}, \tau_f^{-J}) = 0$. If for any J $\tau_{NS}^{-J} > \tau_f^{-J}$, then $\tau_{NR} \geq \tau_{NS}$ by Proposition 1. For all J, these two inequalities imply $\tau_{NR}^{-J} > \tau_f^{-J}$, which together with (1) imply $w^J(\tau_{NR}^{-J}, \tau_f^{-J}) < 0$. Moreover, for $\tau_{NR} \geq \tau_{NS}$,

$SW^J(\tau_{NR}) \leq SW^J(\tau_{NS})$ for all J. But then it follows that for all J, $RW^J(\tau_{NR}, \tau_f^{-J}) < SW^J(\tau_{NS})$.

2.2. Dynamic Game

We assume that countries' welfare and perceptions of the fair tariff level of their partners are common knowledge. Countries discount the future at rate $\delta \in (0,1)$. The repeated-game welfare is given by:

$$\overrightarrow{RW}^J = \sum_{t=0}^{\infty} RW^J(\tau^J, \tau^{-J}, \tau_f^{-J})\delta^t.$$

Denote the infinitely repeated game with reciprocal players by $\Gamma_\infty^R(RW, w, \tau_f)$, and the infinitely repeated game with self-interested players by $\Gamma_\infty^S(SW)$. The incentive-compatibility condition for a self-interested country J to adhere to the cooperative path in $\Gamma_\infty^S(SW)$ using a grim trigger strategy is that the welfare under cooperation, $SW_C^J(\tau_C^J, \tau_C^{-J})/(1-\delta)$, must be no less than the welfare under defection, which consists of the one-time gain from cheating, $SW_D^J(BR_S^J(\tau_C^{-J}), \tau_C^{-J})$, plus the discounted welfare from playing Nash thereafter, $\delta SW_{NS}^J(\tau_{NS}^J, \tau_{NS}^{-J})/(1-\delta)$. That is,

$$SW_D^J(BR_S^J(\tau_C^{-J}), \tau_C^{-J}) + \frac{\delta}{1-\delta} SW_{NS}^J(\tau_{NS}^J, \tau_{NS}^{-J}) \leq \frac{1}{1-\delta} SW_C^J(\tau_C^J, \tau_C^{-J}).$$

Solving for δ we obtain:

$$\delta^{S}_{\tau_C} = \frac{SW^J_D - SW^J_C}{SW^J_D - SW^J_{NS}} \leq \delta. \qquad (2)$$

When countries are self-interested, it follows that cooperation can be sustained if countries are patient enough so that $\delta^S_{\tau_C} \leq \delta$, where $\delta^S_{\tau_C}$ is the critical discount factor above which the self-interested cooperative tariff vector can be sustained by self-interested countries.

The same reasoning applies to countries with reciprocal preferences, the only difference being that their welfare depends on their perceptions of fairness. Thus, the incentive-compatibility condition for a reciprocal country J to impose its self-interested cooperative tariff in $\Gamma^R_\infty(RW, w, \tau^f)$ using a grim trigger strategy is that:

$$RW^J_D(BR^J_R(\tau^{-J}_C), \tau^{-J}_C, \tau^{-J}_f) + \frac{\delta}{1-\delta} RW^J_{NR}(\tau^J_{NR}, \tau^{-J}_{NR}, \tau^{-J}_f) \leq \frac{1}{1-\delta} RW^J_C(\tau^J_C, \tau^{-J}_C, \tau^{-J}_f).$$

Solving once again for δ, we obtain:

$$\delta^{R}_{\tau_C} = \frac{RW^J_D - RW^J_C}{RW^J_D - RW^J_{NR}} \leq \delta. \qquad (3)$$

When countries have reciprocal preferences, it follows that the self-interested cooperative tariff vector can be sustained if countries are patient enough so that $\delta^R_{\tau_C} \leq \delta$, where $\delta^R_{\tau_C}$ is the critical discount factor above which the self-interested cooperative tariff vector can be sustained by reciprocal countries. Note that the cooperative tariffs (or the Nash tariffs) need not be the same across reciprocal countries, since fair-tariff perceptions might differ, implying that the critical discount factors might also differ across them. We focus on $\delta^R_{\tau_C} = \max\left\{\delta^R_{\tau^J_C}, \delta^R_{\tau^{-J}_C}\right\}$, since both countries can sustain cooperation at this discount factor.

We will use (2) and (3) to characterize the impact that preferences for reciprocity have on cooperation. To this end, we compare the critical discount factor above which the self-interested cooperative tariff vector can be sustained in the infinitely repeated game with self-interested countries against that in the infinitely repeated game with reciprocal players. We assume that these two games are identical in all respects (producer surplus, consumer surplus, tariff revenues, and number of players) except for the reciprocity term in the countries' welfare functions. We say that preferences for reciprocity facilitate cooperation when the self-interested cooperative tariff vector can be sustained at a lower critical discount factor in the game with reciprocal countries than in the game with self-interested countries. If the opposite happens, we say that preferences for reciprocity make cooperation harder.

Lemma 2: *Let* $\Gamma^R(RW, w, \tau_f)$ *be a supermodular game where for any* J *(i)* RW^J *has decreasing differences in* (τ^J, τ_f^{-J}), and *(ii)* $\tau_f^{-J} \in [\tau_C^{-J}, \tau_{NS}^{-J}]$. *Let* τ_C *satisfy* $SW^J(\tau_C) > SW^J(\tau_{NS})$ *for all* J. *Under these conditions, there exists a sufficiently high discount factor such that a subgame-perfect Nash equilibrium of* $\Gamma_\infty^R(RW, w, \tau_f)$ *occurs at* τ_C.

Proof: If $\tau_f^{-J} \in [\tau_C^{-J}, \tau_{NS}^{-J}]$ *for all* J, then $w^J(\tau_C^{-J}, \tau_f^{-J}) \geq 0$ and $w^J(\tau_{NR}^{-J}, \tau_f^{-J}) \leq 0$. This in turn implies that:

$$RW^J(\tau_C, \tau_f^{-J}) \geq SW^J(\tau_C). \tag{4}$$

We also know that:

$$SW^J(\tau_C) > SW^J(\tau_{NS}). \tag{5}$$

If $\tau_{NS}^{-J} \geq \tau_f^{-J}$ for all J, then we know from Corollary 1 that:

$$RW^J(\tau_{NR}, \tau_f^{-J}) \leq SW^J(\tau_{NS}). \tag{6}$$

From (4), (5), and (6), we then obtain:

$$RW^J(\tau_C, \tau_f^{-J}) > RW^J(\tau_{NR}, \tau_f^{-J})$$

for all J, which by Friedman (1971) implies that there exists a discount factor such that τ_C is a subgame-perfect Nash equilibrium of $\Gamma_\infty^R(RW, w, \tau_f)$.

This result states that given the fair tariff profile, τ_f, for any τ_C such that countries' welfare at the cooperative tariffs is higher than their welfare at any Nash equilibrium of the stage game, cooperation can also be sustained by reciprocal countries at the strategy profile τ_C. We are now ready to state our first result about the impact of fairness and reciprocity on cooperation.

Proposition 2: *Let $\Gamma^R(RW, w, \tau_f)$ and $\Gamma^S(SW)$ be supermodular games where for any J (i) RW^J has decreasing differences in (τ^J, τ_f^{-J})* and *(ii) $\tau_f^{-J} \in [\tau_C^{-J}, \tau_{NS}^{-J}]$, and γ is sufficiently small. Let the Nash punishments in $\Gamma^R_\infty(RW, w, \tau_f)$ and $\Gamma^S_\infty(SW)$ be either at the smallest or largest pure strategy Nash equilibria of $\Gamma^R(RW, w, \tau_f)$ and $\Gamma^S(SW)$, respectively. Let τ_C satisfy $SW^J(\tau_C) > SW^J(\tau_{NS})$ for any J. Under these assumptions, the critical (minimum) discount factor needed to sustain cooperation at τ_C is lower in $\Gamma^R_\infty(RW, w, \tau_f)$ than in $\Gamma^S_\infty(SW)$, that is, $\delta^R_{\tau_C} < \delta^S_{\tau_C}$.*

Proof: We want to show that $\tau_C^{-J} \leq \tau_f^{-J} \leq \tau_{NS}^{-J}$ for any J implies that $\delta^R_{\tau_C} = \frac{RW^J_D - RW^J_C}{RW^J_D - RW^J_{NR}} < \frac{SW^J_D - SW^J_C}{SW^J_D - SW^J_{NS}} = \delta^S_{\tau_C}$. In order to do that, we will show:

(i) If $\tau_C^{-J} \leq \tau_f^{-J}$ for all $J \Rightarrow RW^J_D - RW^J_C \leq SW^J_D - SW^J_C$.

(ii) If $\tau_C^{-J} \leq \tau_f^{-J} \leq \tau_{NS}^{-J}$ for any J, and γ is sufficiently small $\Rightarrow RW^J_D - RW^J_{NR} > SW^J_D - SW^J_{NS}$.

Let us start with (i). We have that:

$$RW^J_C = SW^J_C(\tau^J_C, \tau^{-J}_C) + \gamma w^J(\tau^{-J}_C, \tau^{-J}_f) SW^{-J}_C(\tau^J_C, \tau^{-J}_C) \text{ and}$$

$$RW^J_D = SW^J_D(BR^J_R(\tau^{-J}_C), \tau^{-J}_C) + \gamma w^J(\tau^{-J}_C, \tau^{-J}_f) SW^{-J}_C(BR^J_R(\tau^{-J}_C), \tau^{-J}_C).$$

Therefore:

$$\begin{aligned} RW^J_D - RW^J_C &= SW^J_D(BR^J_R(\tau^{-J}_C), \tau^{-J}_C) - SW^J_C(\tau^J_C, \tau^{-J}_C) \qquad (7) \\ &+\gamma w^J(\tau^{-J}_C, \tau^{-J}_f)\left(SW^{-J}_C(BR^J_R(\tau^{-J}_C), \tau^{-J}_C) - SW^{-J}_C(\tau^J_C, \tau^{-J}_C)\right) \\ &\leq SW^J_D(BR^J_R(\tau^{-J}_C), \tau^{-J}_C) - SW^J_C(\tau^J_C, \tau^{-J}_C) \\ &\leq SW^J_D(BR^J_S(\tau^{-J}_C), \tau^{-J}_C) - SW^J_C(\tau^J_C, \tau^{-J}_C). \end{aligned}$$

We know that $\gamma > 0$, and $w^J(\tau^{-J}_C, \tau^{-J}_f) \geq 0$ if $\tau^{-J}_C \leq \tau^{-J}_f$. It is clear that the self-interested welfare of country $-J$ is (weakly) lower when country J deviates while country $-J$ continues to set the cooperative tariff than when both countries set the most cooperative tariff, that is, $SW^{-J}_C(BR^J_R(\tau^{-J}_C), \tau^{-J}_C) - SW^{-J}_C(\tau^J_C, \tau^{-J}_C) \leq 0$, which explains the first inequality. The second inequality is by the fact that $BR^J_S(\tau^{-J}_C)$ is the best reply of the self-interested country J. Thus, if $\tau^{-J}_C \leq \tau^{-J}_f$, then $RW^J_D - RW^J_C \leq SW^J_D - SW^J_C$.

We now turn to (ii). Let us rewrite the result we want to show:

$$\tau_C^{-J} \leq \tau_f^{-J} \leq \tau_{NS}^{-J} \text{ and } \gamma \text{ is sufficiently small} \Rightarrow \big(RW_D^J - SW_D^J\big) - \big(RW_{NR}^J - SW_{NS}^J\big) > 0.$$

By corollary 1 we know that if $\tau_f^{-J} \leq \tau_{NS}^{-J}$ for all J, then the Nash equilibrium of $\Gamma^S(SW)$ is (weakly) smaller than that of $\Gamma^R(RW, w, \tau_f)$, that is, $\tau_{NR} \geq \tau_{NS}$. Thus, $\tau_f^{-J} \leq \tau_{NS}^{-J} \leq \tau_{NR}^{-J}$ for all J, and this implies that $w^J(\tau_{NR}^{-J}, \tau_f^{-J}) \leq 0$. Therefore, we can write the following inequality:

$$\begin{aligned} RW_{NR}^J &= SW_{NR}^J(\tau_{NR}^J, \tau_{NR}^{-J}) + \gamma w^J(\tau_{NR}^{-J}, \tau_f^{-J}) SW_{NR}^{-J}(\tau_{NR}^J, \tau_{NR}^{-J}) \\ &\leq SW_{NR}^J(\tau_{NR}^J, \tau_{NR}^{-J}) \leq SW_{NS}^J(\tau_{NS}^J, \tau_{NS}^{-J}). \end{aligned}$$

Now we will show that $RW_D^J - SW_D^J \geq 0$. Taking a first-order Taylor series expansion of $RW_D^J(BR_R^J(\tau_C^{-J}), \tau_C^{-J}, \tau_f^{-J})$ around $\gamma = 0$, we obtain:

$$\begin{aligned} RW_D^J(BR_R^J(\tau_C^{-J}), \tau_C^{-J}, \tau_f^{-J}) &\approx SW_D^J(BR_S^J(\tau_C^{-J}), \tau_C^{-J}) \\ &+\gamma w^J(\tau_C^{-J}, \tau_f^{-J}) SW_C^{-J}(BR_S^J(\tau_C^{-J}), \tau_C^{-J}) \Rightarrow \\ \Rightarrow RW_D^J(BR_R^J(\tau_C^{-J}), \tau_C^{-J}, \tau_f^{-J}) &- SW_D^J(BR_S^J(\tau_C^{-J}), \tau_C^{-J}) \approx \\ &\gamma w^J(\tau_C^{-J}, \tau_f^{-J}) SW_C^{-J}(BR_S^J(\tau_C^{-J}), \tau_C^{-J}) \geq 0. \end{aligned} \tag{8}$$

The inequality follows by the fact that $w^J(\tau_C^{-J}, \tau_f^{-J}) \geq 0$. Thus, the condition we seek is satisfied (since $\tau_C < \tau_{NS}$). By (i) and (ii), we have $\delta^R_{\tau_C} < \delta^S_{\tau_C}$.

Proposition 3: *Let* $\Gamma^R(RW, w, \tau_f)$ *and* $\Gamma^S(SW)$ *be supermodular games where* for all J *(i)* RW^J *has decreasing differences in* (τ^J, τ_f^{-J}) *and (ii)* $\tau_f^{-J} \in [\tau_{CS}^{-J}, \tau_{NS}^{-J}]$, *and* γ *is sufficiently small. Let* $\delta \in [\underline{\delta}, \overline{\delta},]$ *be such that both self-interested and reciprocal countries can sustain cooperative tariff policies but free trade is unenforceable. Then the most cooperative tariff vector of* $\Gamma^s_\infty(SW)$ *is higher than the most cooperative tariff vector of* $\Gamma^R_\infty(RW, w, \tau_f)$, *that is,* $\tau_{CS} > \tau_{CR}$.

Proof: By proposition 2, we know that for any cooperative tariff vector τ_C, $\delta^R_{\tau_C} < \delta^S_{\tau_C}$. So, it is also true for the most cooperative tariff vector of the repeated game with self-interested countries, τ_{CS}: $\delta^R_{\tau_{CS}} < \delta^S_{\tau_{CS}}$. Note that both self-interested and reciprocal countries can cooperate at the discount factor $\delta^S_{\tau_{CS}}$, but only reciprocal countries can cooperate at $\delta^R_{\tau_{CS}}$. From (2) and (3), we have:

$$
\begin{aligned}
SW_D^J - SW_C^J &= \delta^S_{\tau_{CS}} \left(SW_D^J - SW_{NS}^J \right) \\
RW_D^J - RW_C^J &= \delta^R_{\tau_{CS}} \left(RW_D^J - RW_{NR}^J \right)
\end{aligned}
$$

Since $\delta^R_{\tau_{CS}} < \delta^S_{\tau_{CS}}$:

$$\begin{aligned} RW^J_D - RW^J_C &< \delta^S_{\tau_{CS}} \left(RW^J_D - RW^J_{NR} \right) \\ \left(1 - \delta^S_{\tau_{CS}}\right) RW^J_D(BR^J_R(\tau^{-J}_{CS}), \tau^{-J}_{CS}, \tau^{-J}_f) &< RW^J_C(\tau^J_{CS}, \tau^{-J}_f) - \delta^S_{\tau_{CS}} RW^J_{NR} \end{aligned}$$

First note that RW^J_{NR} does only depend on the Nash tariff profile, τ_{NR}. Moreover, for any cooperative tariff profile lower than the most cooperative tariff profile of $\Gamma^s_\infty(SW)$, say $\tau_C < \tau_{CS}$, the gain from cheating at τ_C will be higher than the gain from cheating at τ_{CS}, that is:

$$RW^J_D(BR^J_R(\tau^{-J}_C), \tau^{-J}_C, \tau^{-J}_f) > RW^J_D(BR^J_R(\tau^{-J}_{CS}), \tau^{-J}_{CS}, \tau^{-J}_f).$$

On the other hand, for any $\tau_C < \tau_{CS}$, the gain from cooperation will be also higher at τ_C than at τ_{CS}:

$$RW^J_C(\tau_C, \tau^{-J}_f) > RW^J_C(\tau_{CS}, \tau^{-J}_f).$$

By the continuity of $RW(\bullet)$, for any discount factor $\delta \in [\underline{\delta}, \overline{\delta},]$, the most cooperative equilibrium tariff vector of $\Gamma^s_\infty(SW)$ is higher than the most cooperative equilibrium tariff profile of $\Gamma^R_\infty(RW, w, \tau_f)$, $\tau_{CS} >$ τ_{CR}. ■

Proposition 4: *Let* $\Gamma^R(RW, w, \tau_f)$ *and* $\Gamma^S(SW)$ *be supermodular games where:*

(i) RW^J *has decreasing differences in* (τ^J, τ_f^{-J}) *for all* J*,*

(ii) $\tau_f^{-J} \in \left(0, \tau_{CS}^{-J}\right)$ *for all* J *and* γ *is sufficiently small,*

(iii) Let $\delta \in [\underline{\delta}, \overline{\delta},]$ *be such that both self-interested and reciprocal countries can sustain cooperation but not free trade.*

Then the most cooperative equilibrium tariff profile of $\Gamma_\infty^R(RW, w, \tau_f)$ *can be either higher or lower than the one of* $\Gamma_\infty^S(SW)$*.*

Proof: Proposition 2 holds for any most cooperative tariff profile τ_C lower than the fair-tariffs vector, $\tau_C \leq \tau_f$. For any $\tau_C > \tau_f$, it is ambiguous which critical discount factor is higher by (7) and (8), since the weight function is negative at τ_C. Hence, it is possible that the minimum discount factor required for countries with reciprocal preferences to sustain cooperation at τ_C is higher than that for self-interested countries, i.e., $\delta_{\tau_C}^R > \delta_{\tau_C}^S$. In this case, both types of countries could sustain cooperation at τ_C with a level of discount factor $\delta_{\tau_C}^R$ or above. Moreover, τ_C is the most cooperative tariff vector for $\Gamma_\infty^R(RW, w, \tau_f)$ when the discount factor equals $\delta_{\tau_C}^R$, so we denote them τ_{CR} and $\delta_{\tau_{CR}}^R$ henceforth. From (3), we have:

$$RW_D^J - RW_C^J = \delta_{\tau_{CR}}^R \left(RW_D^J - RW_{NR}^J\right)$$

Since $\delta^R_{\tau_{CR}} > \delta^S_{\tau_{CR}}$:

$$\begin{aligned} SW^J_D - SW^J_C &< \delta^R_{\tau_{CR}} \left(SW^J_D - SW^J_{NS}\right) \\ \left(1 - \delta^R_{\tau_{CR}}\right) SW^J_D(BR^J_S(\tau^{-J}_{CR}), \tau^{-J}_{CR}) &< SW^J_C(\tau^J_{CR}, \tau^{-J}_{CR}) - \delta^R_{\tau_{CR}} SW^J_{NS} \end{aligned}$$

Note that SW^J_{NS} depends only on the Nash tariff τ_{NS}. Moreover, for any most cooperative tariff vector lower than the most cooperative equilibrium tariff profile of $\Gamma^R_\infty(RW, w, \tau_f)$, say $\widehat{\tau}_C < \tau_{CR}$, the gain from cheating at $\widehat{\tau}^C$ will be higher than the gain from cheating at τ_{CR}, that is:

$$SW^J_D(BR^J_S(\widehat{\tau}^{-J}_C), \widehat{\tau}^{-J}_C) > SW^J_D(BR^J_S(\tau^{-J}_{CR}), \tau^{-J}_{CR}).$$

On the other hand, for any $\widehat{\tau}_C < \tau_{CR}$, the gain from cooperation will be also higher at $\widehat{\tau}_C$ than at τ_{CR}:

$$SW^J_C(\widehat{\tau}_C) > SW^J_C(\tau_{CR}).$$

By the continuity of $SW(\bullet)$, for any discount factor $\delta \in [\underline{\delta}, \overline{\delta},]$, the most cooperative equilibrium tariff vector of $\Gamma^S_\infty(SW)$ is lower than the most cooperative equilibrium tariff profile of $\Gamma^R_\infty(RW, w, \tau_f)$, i.e., $\tau_{CS} < \tau_{CR}$.

Nevertheless, $\delta^{R}_{\tau_{CR}} < \delta^{S}_{\tau_{CR}}$ is also possible by (7) and (8). Because of the ambiguity of which critical discount factor is higher when $\tau^{f} < \tau^{CS}$, it is also ambiguous which of the most cooperative tariffs is higher than the other.

3. Conclusions

This paper explores trade policy and trade agreements between governments with reciprocal preferences. Such governments respond kindly to actions that are perceived to be kinder than expected and they retaliate when others are perceived to engage in unkind behavior. In other words, these governments have some beliefs over what a "fair" trade policy is and they react with lower tariffs when trade partners set their tariffs lower than the "fair" tariff and vice versa. We then compare the results with the standard case where governments maximize national welfare.

We find that reciprocity has a significant impact on trade agreements. We show that if the "fair" tariff is higher than a critical level, governments with reciprocal preferences can achieve lower cooperative tariffs in a dynamic infinitely repeated tariff game. However, if governments believe that "fair" tariffs are too low then it is possible that they end up with higher cooperative tariffs. We, also, show that these results apply to very general social welfare functions.

The intuition gained from these results could provide insights for the failure of the Doha round. It is conceivable that the success of the previous rounds and the overall economic environment at the time may have raised expectations too high (i.e., lowered the general perception of "fair" tariffs significantly), hindering the efforts for further trade liberalization and more multilateral trade cooperation.