Multilateral Tariff Cooperation under Fairness and Reciprocity

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

Abstract

Keywords:

JEL classification:

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

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

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. Specific Model

We consider a world with two countries, A and B, and two goods, a and b. Country J imports good j from country $-J$, where $J \in \{A, B\}$, $j \in \{a, b\}$, and $-J$ denotes the trading partner of J. Each country is endowed with zero units of its import good and 1 unit of its export good.

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 $-j$, where $-j$ denotes the good exported by J to $-J$. More specifically, the demand for good j in country J is given by $D\left(P_j^J\right) = \alpha - \beta P_j^J$, where P_j^J is the price of good j in country J.

Let τ_j be the tariff country J levies on its imports. The no-arbitrage condition is given by:

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

The export supply functions are:

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

Market clearing then implies:

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

3. Generalized Model

We now introduce a generic demand function. But the signs of the derivative of any function in what follows remain the same as hitherto.

3.1. Static Game

Let the reciprocal two-country trade game be denoted by $\Gamma^R(RW, w, \tau^f)$, where RW is the welfare function of a reciprocal country, w is the weight function a country places on its trading partner's self-interested welfare SW, and $\tau^f \equiv \left(\tau_j^f, \tau_{-j}^f\right)$ with τ_j^f being the fair-tariff perception of country $-J$ with regard to what the level of country J's import tariff should be. Let the self-interested two-country trade game be denoted by $\Gamma^S(SW)$.

The welfare of reciprocal country J is given by:

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

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:

$$SW^{-J} = \int_{P^{-J}_{-j}}^{\frac{\alpha}{\beta}} D(P)\, dP + \int_{P^{-J}_{j}}^{\frac{\alpha}{\beta}} D(P)\, dP + P^{-J}_{j} + \tau_{-j} X^{J}_{-j}.$$

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. This suggests 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:

$$\frac{\partial RW^J}{\partial \tau_j \partial \tau_{-j}} = \underbrace{-\underbrace{\frac{\partial D\left(P_j^J\right)}{\partial \tau_{-j}}}_{=0}\underbrace{\frac{\partial P_j^J}{\partial \tau_j}}_{\geq 0} - \underbrace{D\left(P_j^J\right)}_{\geq 0}\underbrace{\frac{\partial^2 P_j^J}{\partial \tau_j \partial \tau_{-j}}}_{=0} + \underbrace{\frac{\partial X_j^{-J}}{\partial \tau_{-j}}}_{=0} + \tau_j \underbrace{\frac{\partial^2 X_j^{-J}}{\partial \tau_j \partial \tau_{-j}}}_{=0}}_{=0}$$
$$+\gamma \left(\begin{array}{c} \underbrace{\underbrace{\frac{\partial w_j(\tau_{-j},\tau^f_{-j})}{\partial \tau_{-j}}}_{\leq 0}\underbrace{\left(1 - D\left(P_j^{-J}\right)\right)}_{\geq 0}\underbrace{\frac{\partial P_j^{-J}}{\partial \tau_j}}_{\leq 0}}_{\geq 0} \\ \underbrace{+ w_j(\tau_{-j},\tau^f_{-j})\left(1 - D\left(P_j^{-J}\right)\right)\underbrace{\frac{\partial^2 P_j^{-J}}{\partial \tau_j \partial \tau_{-j}}}_{=0} - w_j(\tau_{-j},\tau^f_{-j})\frac{\partial P_j^{-J}}{\partial \tau_j}\underbrace{\frac{\partial D\left(P_j^{-J}\right)}{\partial \tau_{-j}}}_{=0}}_{=0} \end{array} \right).$$

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, $\overline{\tau_j}$ and $\underline{\tau_j}$. Moreover, the strategy profiles $\overline{\tau} \equiv (\overline{\tau_j}, \overline{\tau_{-j}})$ 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., $\overline{\tau}^{NR} \equiv \left(\overline{\tau_j}^{NR}, \overline{\tau_{-j}}^{NR}\right)$ and $\underline{\tau}^{NR} \equiv \left(\underline{\tau_j^{NR}}, \underline{\tau_{-j}}^{NR}\right)$, 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 $\tau^{NR} \equiv \left(\tau_j^{NR}, \tau_{-j}^{NR}\right)$, 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_{-j}^f) *and* $\tau_{-j}^{NS} \geq \tau_{-j}^f$, *then* $\tau^{NR} \geq \tau^{NS}$ and for all J, $RW^J(\tau^{NR}, \tau_{-j}^f) \leq SW^J(\tau^{NS})$.

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 $\tau_{-j}^f = \tau_{-j}^{NS}$, then trivially $\tau^{NR} = \tau^{NS} = \tau^N$ and for all J, $RW^J(\tau^N, \tau_{-j}^f) = SW^J(\tau^N)$ since $w_j(\tau_{-j}^N, \tau_{-j}^f) = 0$. If for any J $\tau_{-j}^{NS} \geq \tau_{-j}^f$, then $\tau^{NR} \geq \tau^{NS}$ by Proposition 1. For all J, these two inequalities imply $\tau_{-j}^{NR} \geq \tau_{-j}^f$, which together with (1) imply $w_j(\tau_{-j}^{NR}, \tau_{-j}^f) \leq 0$. 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, $RW^J(\tau^{NR}, \tau_{-j}^f) \leq SW^J(\tau^{NS})$.

3.2. Dynamic Game

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

$$\overrightarrow{RW}^J = \sum_{t=0}^{\infty} RW^J(\tau_j, \tau_{-j}, \tau_{-j}^f)\delta^t.$$

Denote the infinitely repeated game with reciprocal players by $\Gamma^R_\infty(RW, w, \tau^f)$ and the infinitely repeated game with self-interested players by $\Gamma^S_\infty(SW)$. The incentive compatibility condition for a self-interested country j to play the cooperative action in $\Gamma^S_\infty(SW)$ using a grim trigger strategy is that the welfare from cooperation, $SW^J_C(\tau^C_j, \tau^C_{-j})/(1-\delta)$, must be no less than the welfare from defection which consists of the one period gain from deviating $SW^J_D(BR^J_S(\tau^C_{-j}), \tau^C_{-j})$ plus the discounted welfare of inducing Nash reversion forever $\delta SW^J_{NS}(\tau^{NS}_j, \tau^{NS}_{-j})/(1-\delta)$. That is,

$$SW^J_D(BR^J_S(\tau^C_{-j}), \tau^C_{-j}) + \frac{\delta}{1-\delta} SW^J_{NS}(\tau^{NS}_j, \tau^{NS}_{-j}) \leq \frac{1}{1-\delta} SW^J_C(\tau^C_j, \tau^C_{-j}).$$

Solving for δ we obtain

$$\delta^S_{\tau^C} = \frac{SW^J_D - SW^J_C}{SW^J_D - SW^J_{NS}} \leq \delta. \tag{2}$$

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

The same reasoning applies when countries have reciprocal welfare with the difference that the welfare depend on the tariffs chosen and on countries' perceptions of fairness. Thus, the incentive compatibility condition for a reciprocal country i to play the self-interested cooperative action in $\Gamma^R_\infty(RW, w, \tau^f)$ using a grim trigger strategy is

$$RW^J_D(BR^J_R(\tau^C_{-j}), \tau^C_{-j}, \tau^f_{-j}) + \frac{\delta}{1-\delta} RW^J_{NR}(\tau^{NR}_j, \tau^{NR}_{-j}, \tau^f_{-j}) \leq \frac{1}{1-\delta} RW^J_C(\tau^C_j, \tau^C_{-j}, \tau^f_{-j}).$$

Solving for δ we obtain

$$\delta^{R}_{\tau^C} = \frac{RW^J_D - RW^J_C}{RW^J_D - RW^J_{NR}} \leq \delta. \tag{3}$$

When countries have reciprocal welfare it follows that the self-interested cooperative outcome can be sustained if countries are patient enough such that $\delta^{R}_{\tau^C} \leq \delta$ where $\delta^{R}_{\tau^C}$ is the critical discount factor above which the self-interested cooperative outcome can be sustained by reciprocal countries. Note that the most collusive tariffs and the nash tariffs need not to be the same for the reciprocal countries, since fair tariff perceptions may differ. This possibility yields different levels of critical discount factors. We focus on $\delta^{R}_{\tau^C} = \max\left\{\delta^{R}_{\tau^C_j}, \delta^{R}_{\tau^C_{-j}}\right\}$, since both countries can sustain cooperation at the higher discount factor.

We will use (2) and (3) to characterize the impact that preferences for reciprocity have on cooperation. To perform this analysis we compare the critical discount factor above which the self-interested cooperative tariff can be sustained in the infinitely repeated game with self-interested countries to 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) with the exception of the reciprocal term at the countries' welfare. We say that welfare for reciprocity facilitate cooperation when the self-interested cooperative tariff 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 (i)* RW^J *has de-*

creasing differences in (τ_j, τ^f_{-j}) *for all* j, *(ii)* $\tau^f_{-j} \in [\tau^C_{-j}, \tau^{NS}_{-j}]$ *for all* j. *Let* τ^C *satisfy* $SW^J(\tau^C) > SW^J(\tau^{NS})$ *for all* J. *Under these conditions there is a sufficiently high discount factor such that there exists a subgame-perfect Nash equilibrium of* $\Gamma^R_\infty(RW, w, \tau^f)$ *at* τ^C.

Proof: If $\tau^f_{-j} \in [\tau^C_{-j}, \tau^{NS}_{-j}]$ *for all* j, then $w_j(\tau^C_{-j}, \tau^f_{-j}) > 0$ and $w_j(\tau^{NR}_{-j}, \tau^f_{-j}) < 0$, for all j. 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

$$RW^J(\tau^{NR}, \tau^f_{-j}) \leq SW^J(\tau^{NS}) \tag{6}$$

for all i. From (4), (5) and (6) we 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 $\Gamma^R(RW, w, \tau^f)$.

This result states that given the fair tariff profile, τ^f, for any τ^C such that the countries' welfare at the cooperative tariffs are higher than their welfare at any Nash equilibria of the stage game, cooperation can 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 (i)* RW^J *has decreasing differences in* (τ_j, τ^f_{-j}) *for all* j*, and (ii)* $\tau^f \in [\tau^C, \tau^{NS}]$ *and* γ *is sufficiently small. Let Nash punishments in* $\Gamma^R_\infty(RW, w, \tau^f)$ *and in* $\Gamma^S_\infty(SW)$ *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 all* 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 all j implies that $\delta^R_{\tau^C} = \frac{RW^J_D - RW^J_C}{RW^J_D - RW^J_{NR}} < \frac{SW^J_D - SW^J_C}{SW^J_D - SW^J_{NS}} = \delta^S_{\tau^C}$. To be able to do that we will show:
(i) $\tau^C_{-j} \leq \tau^f_{-j}$ for all $j \Rightarrow RW^J_D - RW^J_C \leq SW^J_D - SW^J_C$.
(ii) $\tau^C_{-j} \leq \tau^f_{-j} \leq \tau^{NS}_{-j}$ for all j and γ is sufficiently small $\Rightarrow RW^J_D - RW^J_{NR} > SW^J_D - SW^J_{NS}$.

Let's start with (i):

$$RW^J_C = SW^J_C(\tau^C_j, \tau^C_{-j}) + \gamma w_j(\tau^C_{-j}, \tau^f_{-j}) SW^{-J}_C(\tau^C_j, \tau^C_{-j})$$

and

$$RW^J_D = SW^J_D(BR^J_R(\tau^C_{-j}), \tau^C_{-j}) + \gamma w_j(\tau^C_{-j}, \tau^f_{-j}) SW^{-J}_C(\tau^D_j, \tau^C_{-j})$$

So

$$\begin{aligned} RW_D^J - RW_C^J &= SW_D^J(BR_R^J(\tau_{-j}^C), \tau_{-j}^C) - SW_C^J(\tau_j^C, \tau_{-j}^C) + ... \\ &\quad ... + \gamma w_j(\tau_{-j}^C, \tau_{-j}^f)\left(SW_C^{-J}(\tau_j^D, \tau_{-j}^C) - SW_C^{-J}(\tau_j^C, \tau_{-j}^C)\right) \quad (7) \\ &\leq SW_D^J(BR_R^J(\tau_{-j}^C), \tau_{-j}^C) - SW_C^J(\tau_j^C, \tau_{-j}^C) \\ &< SW_D^J(BR_S^J(\tau_{-j}^C), \tau_{-j}^C) - SW_C^J(\tau_j^C, \tau_{-j}^C) \end{aligned}$$

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 lower when country J deviates while country $-J$ continues to set the cooperative tariff than both countries set most cooperative tariff, that is, $SW_C^{-J}(\tau_j^D, \tau_{-j}^C) - SW_C^{-J}(\tau_j^C, \tau_{-j}^C) \leq 0$, which explains the first inequality. The strict inequality is by the fact that $BR_S^J(\tau_{-j}^C)$ is the best reply of the self-interested country J. Thus, if $\tau_{-j}^C \leq \tau_{-j}^f$, then $RW_D^J - RW_C^J \leq SW_D^J - SW_C^J$.

Now the case (ii): Let's rewrite the result we want to show:

$$\tau_{-j}^C \leq \tau_{-j}^f \leq \tau_{-j}^{NS} \text{ and } \gamma \text{ is sufficiently small} \Rightarrow \left(RW_D^J - SW_D^J\right) - \left(RW_{NR}^J - SW_{NS}^J\right) \geq 0$$

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

$$RW_{NR}^J = SW_{NR}^J(\tau_j^{NR}, \tau_{-j}^{NR}) + \gamma w_j(\tau_{-j}^{NR}, \tau_{-j}^f) SW_{NR}^{-J}(\tau_j^{NR}, \tau_{-j}^{NR}) \leq SW_{NS}^J(\tau_j^{NS}, \tau_{-j}^{NS})$$

Now we will show that $RW_D^J - SW_D^J \geq 0$. Taking a first-order Taylor series expansion of $RW_D^J(BR^J(\tau_{-j}^C), \tau_{-j}^C, \tau_{-j}^f)$ around $\gamma = 0$ we obtain

$$\begin{aligned} RW_D^J(BR_R^J(\tau_{-j}^C), \tau_{-j}^C, \tau_{-j}^f) &\approx SW_D^J(BR_S^J(\tau_{-j}^C), \tau_{-j}^C) + ... \\ & ... + \gamma w_j(\tau_{-j}^C, \tau_{-j}^f) SW_C^{-J}(BR^J(\tau_{-j}^C), \tau_{-j}^C) \\ RW_D^J(BR_R^J(\tau_{-j}^C), \tau_{-j}^C, \tau_{-j}^f) - SW_D^J(BR_S^J(\tau_{-j}^C), \tau_{-j}^C) &\approx \\ \gamma w_j(\tau_{-j}^C, \tau_{-j}^f) SW_C^{-J}(BR^J(\tau_{-j}^C), \tau_{-j}^C) &\geq 0 \end{aligned} \tag{8}$$

The inequality follows by the fact that $w_j(\tau_{-j}^C, \tau_{-j}^f) \geq 0$. Thus, the condition we seek is satisfied. By (i) and (ii), we have $\delta_{\tau^C}^R < \delta_{\tau^C}^S$.

Proposition 3: *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_{-j}^f) for all j, and (ii) $\tau^f \in [\tau^{CS}, \tau^{NS}]$ and γ is sufficiently small. Let $\delta \in [\underline{\delta}, \overline{\delta},]$ be such that both standard and reciprocal countries will cooperate and there is no free trade, then the most cooperative tariff of $\Gamma_\infty^s(SW)$ is higher than the most cooperative tariff of $\Gamma_\infty^R(RW, w, \tau^f)$, that is, $\tau^{CS} > \tau^{CR}$.*

Proof: By proposition 2, we know that for any cooperative tariff τ^C, the $\delta_{\tau^C}^R < \delta_{\tau^C}^S$. So, it is also true for the most cooperative tariff of standard country, τ^{CS}, $\delta_{\tau^{CS}}^R < \delta_{\tau^{CS}}^S$. Note that both standard and reciprocal countries can cooperate at the discount factor $\delta_{\tau^{CS}}^S$, but only the reciprocal country can cooperate at $\delta_{\tau^{CS}}^R$. By (2) we write

$$\begin{aligned} SW_D^J - SW_C^J &= \delta_{\tau^{CS}}^S \left(SW_D^J - SW_{NS}^J\right) \\ RW_D^J - RW_C^J &= \delta_{\tau^{CS}}^R \left(RW_D^J - RW_{NR}^J\right) \end{aligned}$$

Since $\delta^R_{\tau^{CS}} < \delta^S_{\tau^{CS}}$,

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

First note that RW^J_{NR} does depend only on the nash tariff, τ^{NR}. Moreover, for any most cooperative tariff lower than the most cooperative tariff of the standard country, 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,

$$RW^J_D(BR^J_R(\tau^C_{-j}), \tau^C_{-j}, \tau^f_{-j}) > RW^J_D(BR^J_R(\tau^{CS}_{-j}), \tau^{CS}_{-j}, \tau^f_{-j})$$

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},

$$RW^J_C(\tau^C, \tau^f_{-j}) > RW^J_C(\tau^{CS}, \tau^f_{-j}).$$

By the continuity of RW(.), for any discount factor $\delta \in [\underline{\delta}, \overline{\delta},]$, the most cooperative tariff of the standard country is higher than the most cooperative tariff of the reciprocal country, $\tau^{CS} > \tau^{CR}$.

4. 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.