

WTO as Moral Support

Taiji Furusawa*

Abstract

International cooperation in trade policies under the auspices of the WTO makes countries “feel” more obliged to uphold agreements. The paper emphasizes the role of the WTO to give moral support: countries incur “psychological costs” when they renege on the agreements that are formally signed under the WTO. Using the concept of Kandori’s (2003) “morale equilibrium,” we formalize this idea and show that countries can agree on a cooperative level of the binding tariffs but they occasionally deviate from the agreement, which lowers the morale and invites further deviations in the future.

1. Introduction

After World War II, countries have cooperated in their trade relationships more-or-less successfully within the system of the GATT/WTO. There is perhaps no doubt that the GATT/WTO system facilitates trade liberalization and sustaining international cooperation in the exchange of goods and services. But the question is how the GATT/WTO system helps countries cooperate. It is important to know the answer to this question in order to build a better world trade system.

International cooperation in trade policies can be sustained with the threat of punishment that is triggered when countries renege on agreements. Dixit (1987) formalizes this idea in a repeated-game setting and shows that countries can sustain a cooperative tariff profile in a self-enforcing subgame-perfect equilibrium. Since any agreements in that framework are self-enforcing, however, there is little room left for the WTO to actively participate in international cooperation. Maggi (1999) emphasizes that the WTO enhances a multilateral enforcement mechanism so that it facilitates international cooperation when there exists some kind of imbalances in bilateral trade relationships. Bagwell and Staiger (1999, 2005) show that the two pillars of the GATT/WTO system, reciprocity and nondiscrimination, together help countries sustain a Pareto-optimal tariff profile. They emphasize the role of GATT Article XXVIII that allows contracting parties to withdraw substantially equivalent concessions when other countries have reneged on the agreement. So they also appeal effectively to the threat of punishment as a cause of international cooperation.

The above literature emphasizes the role of punishment in the international cooperation in trade policies, and it is true that the WTO can authorize retaliation. As Jackson (1997, p. 116) notes, however, retaliation has not been carried out often, especially before the birth of the WTO. Yet, “the GATT system held together rather well” (Schwartz and Sykes, 2002). Schwartz and Sykes (2002) also write “[w]hat is

* Furusawa: Graduate School of Economics, Hitotsubashi University, Kunitachi, Tokyo 186-8601, Japan. E-mail: furusawa@econ.hit-u.ac.jp. I am grateful to Katsuaki Matsutani, Hiroshi Mukunoki, Yoichi Sugita, and participants of the spring 2002 meeting of the Japanese Economic Association, the conference of International Trade and Factor Mobility at Kansai University and seminars at the Development Bank of Japan, Hitotsubashi University, and Kyoto University for helpful comments. I am also grateful to an anonymous referee for helpful comments and to the Ministry of Education, Culture, Sports, Science, and Technology of Japan for financial support provided to the Hitotsubashi COE/RES.

remarkable about the WTO/GATT system is how unimportant formal sanctions have been in encouraging compliance with trade commitments throughout its history.” They list three considerations that explain why the GATT/WTO works well. One of them is the reputation effect: countries that renege on the agreement may face some reputational cost. That reputation effect has been also mentioned elsewhere. Kovenock and Thursby (1997) quote “twin engines of international obligation and retaliation” (Hudec, 1990). Kovenock and Thursby (1992) formalize this idea of twin engines by introducing in their model the disutility, which can be considered as “guilty conscience” when countries renege on the agreement.

In this paper, we extend the idea introduced by Kovenock and Thursby (1992), but focus on the role of international obligation. As they formalize, countries may well feel more obliged to uphold an agreement under the auspices of the WTO and incur “psychological costs” if they renege on the agreement. Unlike Kovenock and Thursby (1992), however, we allow the morale, which can be thought of as the marginal psychological costs, to evolve reflecting countries’ past and present actions. Our mechanism that sustains international tariff cooperation does not involve explicit punishment that follows immediately and automatically from a breach of the agreement. The model accords with the observation that retaliation against a deviation does not often occur in reality. The WTO facilitates international cooperation simply by providing the contracting parties with the cooperation framework. But the fact that agreements are formally signed under the WTO makes the countries feel more obliged to uphold the agreement; the WTO gives moral support.

To analyze this function of the WTO, we adopt the concept of the “morale equilibrium,” developed by Kandori (2003). We find that countries will select low, cooperative, tariff rates even without any explicit punishment scheme since countries would feel guilty when they renege on the agreement. How much guilt they feel depends on the current morale, which evolves if countries occasionally renege on the agreement. Deviations from the agreement lower the morale, which invites further deviations in the future. This domino effect induces the morale to deteriorate but gives countries more incentive to cooperate. A similar concept to the morale has also been introduced by Benchenkroun and Long (2008) who study a public-good contribution game in which agents condition their individual contributions on the “stock of cooperation” that summarizes the history of cooperation.

Common in these studies is the feature that cooperation level higher than that of the static Nash equilibrium can be sustained. Kandori (2003) introduces a psychological factor into players’ payoff functions to derive this result. He analyzes how the effort levels of myopic agents evolve over time, and derives the long-run stochastic stable set of efforts (Kandori et al., 1993; Young, 1993). Agents would incur psychological costs when selecting a lower effort than the norm, which is a median of the agents’ effort in the last period. In the presence of these costs, myopic agents sustain cooperative levels of effort. In this paper, the norm is the agreed-upon tariffs. Countries feel guilty when they violate the agreement. Unlike Kandori, players (i.e. countries) are far-sighted, so they can sustain even higher levels of cooperation due to the aforementioned domino effect. Benchenkroun and Long’s (2008) mechanism to sustain a cooperative level of contribution is similar to the domino effect. The domino effect in this paper works through the evolution of the payoff-relevant morale, whereas that of Benchenkroun and Long works through the payoff-irrelevant belief that agents possess about other agents’ behavior. More specifically, Benchenkroun and Long find the equilibrium in which agents condition their actions on the stock of cooperation; cooperation is enhanced because agents expect their additional contributions to increase the stock of cooperation, which in turn induce

all agents to contribute more in the future. The mechanism to enhance cooperation in this paper is similar to Benchechroun and Long's, but it is made more explicit by the introduction of psychological factor, just as Kandori.

We also examine the roles of safeguards and continuing GATT/WTO negotiation rounds in this context. We find that the safeguards policy facilitates cooperation as Bagwell and Staiger (1990) and Ethier (2002) argue in different contexts. Continuation of multilateral trade negotiation rounds is also expected to facilitate international cooperation by refreshing the morale of countries.

2. The Basic Model

We consider tariff settings by n symmetric countries that produce and consume a competitively produced *numéraire* good. Each country i also consumes the competitively produced non-*numéraire* good i , which is produced by all countries. Countries are populated by the same number of identical consumers whose preferences are characterized by an additively-separable quasilinear utility function in which the consumption of the *numéraire* good enters linearly. Owing to the quasilinearity of the utility function, social welfare of each country can be measured by the total surplus derived in the markets of the non-*numéraire* goods.

We consider the situation in which each country i imports good i and exports all other $n - 1$ goods, deriving the import surplus that is a function of its own tariff rate and the export surplus from exporting good $j \neq i$ to every other country j that is a function of country j 's tariff rate. Social surplus, which is the objective function of the government of each country, can be measured by the sum of the import surplus and the aggregate export surplus.¹ We also suppose that every country is exposed to a country-specific demand or supply shock in each period of infinite horizon. We assume for simplicity that any shock in a country only affects that country's import surplus.²

Each country i imposes an import tariff on the non-*numéraire* good at a specific rate of τ_i^j . We assume for simplicity that countries do not impose tariffs on the *numéraire* good. The country-specific shock θ_i^j is i.i.d. over time (t) and countries (i), having a common probability distribution over the support $[0, \bar{\theta}]$. Then we express country i 's import surplus by $M(\tau_i^j, \theta_i^j)$ and its export surplus derived from its export to country $j \neq i$ by $X(\tau_j^i, \theta_j^i)$. Notice that the functions M and X are common to all countries by symmetry. We assume that M is concave and has a unique maximum with respect to τ_i^j for any θ_i^j , which reflects the terms-of-trade effect, and that θ_i^j is defined so that M is increasing in θ_i^j . We further assume that $\partial M / \partial \tau_i^j$ is also increasing in θ_i^j to capture the idea that an increase in import demands raises country i 's incentive to impose a higher tariff. On the other hand, the export surplus derived from country $j \neq i$, $X(\tau_j^i, \theta_j^i)$, is decreasing in τ_j^i and increasing in θ_j^i .

With these functions, we express country i 's social welfare by

$$W(\{\tau_i^j\}_{j=1}^n, \{\theta_i^j\}_{j=1}^n) = M(\tau_i^i, \theta_i^i) + \sum_{j \neq i} X(\tau_j^i, \theta_j^i).$$

The final assumption that we make on W is that W is jointly decreasing in all its tariff arguments:

$$\frac{\partial}{\partial \tau} \left[M(\tau, \theta_i^i) + \sum_{j \neq i} X(\tau, \theta_j^i) \right] < 0, \quad \text{for any } \{\theta_i^j\}_{j=1}^n.$$

That is, mutual tariff reduction is Pareto-improving.

In addition to the material payoff described above, the per-period payoff for the government of each country involves a psychological factor. Let α and k_t , respectively, denote an agreed-upon binding rate of tariff and the morale, which we define as the psychological marginal cost in period t . Then, we express country i 's per-period payoff by

$$u\left(\{\tau_t^j\}_{j=1}^n, k_t, \alpha, \{\theta_t^j\}_{j=1}^n\right) = W\left(\{\tau_t^j\}_{j=1}^n, \{\theta_t^j\}_{j=1}^n\right) - k_t[\tau_t^i - \alpha]_+,$$

where $[x]_+ = \max\{x, 0\}$. A country will incur psychological costs if it selects its tariff rate above the agreed-upon level α . The size of such costs depends on the morale and the size of deviation. The country feels more guilty when the morale is high and when its (upward) deviation from the agreed-upon tariff rate is large.

The morale k_t evolves in general, reflecting the history of all countries' tariff-setting behavior. Countries' deviations from the agreed-upon tariff rate are expected to lower k_t as the morale deteriorates in the international society. However, the morale may be enhanced again if the world observes fewer or no deviations. We formalize this idea by specifying the function K that determines the law of motion for k_t . Let $d_t \equiv \sum_{j=1}^n [\tau_t^j - \alpha]_+$ denote the aggregate deviation in period t . Restricting k_t to the closed interval $[\underline{k}, \bar{k}]$ for any $t = 1, 2, \dots$, we specify the law of motion for k_t by

$$k_{t+1} = K(d_t, k_t),$$

where K is decreasing in d_t and increasing in k_t . It is natural to assume that the morale is enhanced if no deviation is observed in the current period. The uppermost schedule in Figure 1 indicates this situation, where $k_{t+1} \geq k_t$ for the entire range of k_t with strict

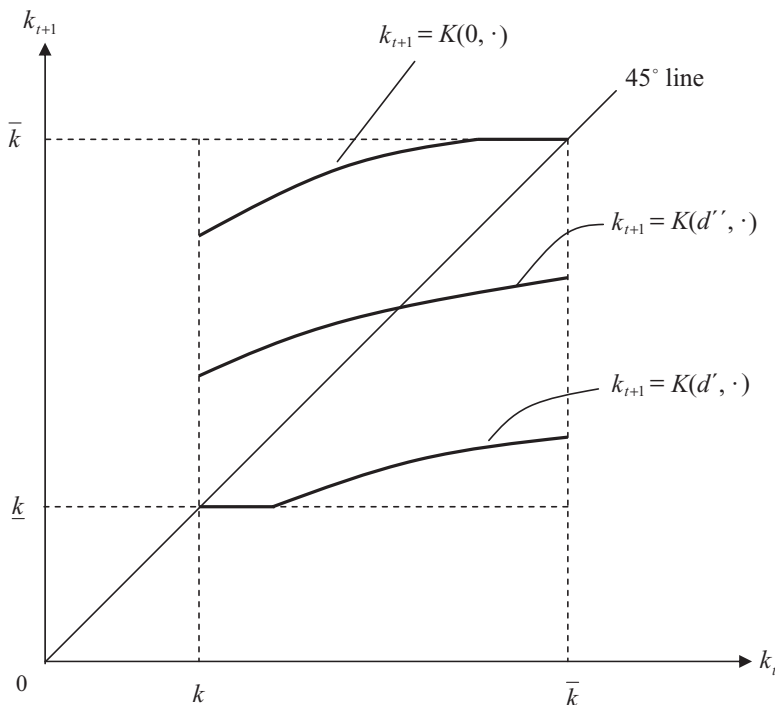


Figure 1. The Law of Motion for the Morale

inequality when k_i is smaller than a critical value. The morale is expected to be eroded, on the other hand, if the current aggregate deviation is very large. The lowermost schedule in Figure 1 shows the case in which a significantly large aggregate deviation, d' , will bring the morale down to the lowest level if the current morale is small enough. The middle schedule depicts the case in which the aggregate deviation is an intermediate level, d'' . In that case, the morale is eroded if the current morale is high but it is enhanced if the current morale is low. The same aggregate deviation may be viewed differently depending on the current morale: it is viewed as a significant deviation if the current morale is high but is viewed as a minor deviation otherwise. The higher the morale, the more easily is the morale eroded.

3. Morale Equilibrium of the Tariff-Setting Game

This section derives the one-shot morale equilibrium (Kandori, 2003) and shows that countries may sometimes violate the agreement by setting a higher tariff rate than the agreed-upon level and may even select a lower tariff rate some other time. We fix the morale at a certain level for the entire analysis of this section.

We define the morale equilibrium in this context such that $\{\hat{\tau}^j\}_{j=1}^n$ is the morale equilibrium if for any i and $\tau_i^i, u(\hat{\tau}^i, \{\hat{\tau}^j\}_{j \neq i}, k_i, \alpha, \{\theta_i^j\}_{j=1}^n) \geq u(\tau_i^i, \{\hat{\tau}^j\}_{j \neq i}, k_i, \alpha, \{\theta_i^j\}_{j=1}^n)$. To derive the morale equilibrium, we first observe that

$$\frac{\partial u}{\partial \tau_i^i}(\{\tau_i^j\}_{j=1}^n, \{\theta_i^j\}_{j=1}^n) = \begin{cases} \frac{\partial M}{\partial \tau_i^i}(\tau_i^i, \theta_i^i) - k_i & \text{if } \tau_i^i > \alpha \\ \frac{\partial M}{\partial \tau_i^i}(\tau_i^i, \theta_i^i) & \text{if } \tau_i^i \leq \alpha. \end{cases} \quad (1)$$

The function u is not differentiable with respect to τ_i^i at $\tau_i^i = \alpha$. But it is immediate that the right-hand derivative equals the first line of (1) and the left-hand derivative equals the second line of (1), both evaluated at $\tau_i^i = \alpha$.

To express the morale equilibrium, we define $\tau^*(k_i, \theta_i^i)$ as τ_i^i that satisfies $(\partial M / \partial \tau_i^i)(\tau_i^i, \theta_i^i) = k_i$. Since $\partial M / \partial \tau_i^i$ is decreasing in τ_i^i and increasing in θ_i^i , the function τ^* is decreasing in k_i and increasing in θ_i^i . Figure 2 depicts two such tariffs, $\tau^*(k_i, \theta_i^i)$ and $\tau^*(0, \theta_i^i)$, which are given as the tariff rates that assign k_i and 0 to the values of $(\partial M / \partial \tau_i^i)(\tau_i^i, \theta_i^i)$. We also define $\hat{\theta}(k_i, \alpha)$ as θ_i^i that satisfies $\tau^*(k_i, \theta_i^i) = \alpha$. That is, the marginal import surplus at $\tau_i^i = \alpha$ is exactly equal to k_i when $\theta_i^i = \hat{\theta}(k_i, \alpha)$. Since τ^* is decreasing in k_i and increasing in θ_i^i , $\hat{\theta}$ is increasing in both k_i and α .

Figure 2 shows $\partial M / \partial \tau_i^i$ and the marginal psychological costs, which is k_i , for $\tau_i^i > \alpha$ and 0 for $\tau_i^i \leq \alpha$. As the figure suggests, country i selects α for multiple contingencies. Since the best responses do not depend on other countries' tariff selection as (1) shows, we find that in the morale equilibrium, each country i chooses its tariff such that

$$\hat{\tau}^i = \begin{cases} \tau^*(0, \theta_i^i) & \text{if } \theta_i^i < \hat{\theta}(0, \alpha) \\ \alpha & \text{if } \hat{\theta}(0, \alpha) \leq \theta_i^i \leq \hat{\theta}(k_i, \alpha) \\ \tau^*(k_i, \theta_i^i) & \text{if } \theta_i^i > \hat{\theta}(k_i, \alpha). \end{cases}$$

Note that $\tau^*(0, \theta_i^i)$ is the Nash-equilibrium tariff rate when countries only care about the material payoffs. In the presence of psychological factors, a country selects a tariff rate lower than this level in general. If the import demand level is low ($\theta_i^i < \hat{\theta}(0, \alpha)$) in the current period, the Nash-equilibrium tariff rate with only material payoffs falls short of the agreed-upon tariff rate. In this case, the country will choose the

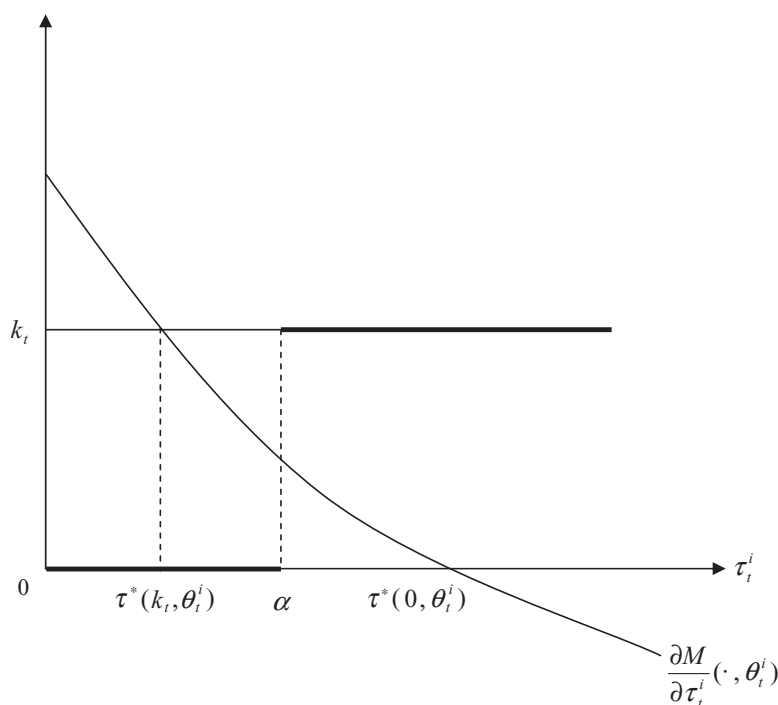


Figure 2. The Morale Equilibrium Strategy

Nash-equilibrium tariff rate. If the import demand level is in an intermediate range ($\hat{\theta}(0, \alpha) \leq \theta_t^i \leq \hat{\theta}(k_t, \alpha)$), the country optimally selects the agreed-upon tariff rate. If the import demand level is high ($\theta_t^i > \hat{\theta}(k_t, \alpha)$), the country is better off to violate the agreement. Even in this case, however, the country's tariff rate is lower than the Nash-equilibrium tariff rate when countries care only about their material payoffs.

PROPOSITION 1. *When countries feel guilty if and only if they choose a tariff higher than an agreed-upon level, they optimally select their individual agreed-upon tariff rates in multiple contingencies. If the positive shock to the import demand function is large, however, a country may violate the agreement. If the shock is very small, a country may choose a tariff rate that is smaller than the agreed-upon level.*

4. Sustainability of International Tariff Agreements

The current morale level affects countries' tariff settings, which in turn affect the morale in the future. Countries select their individual tariffs taking account of this chain reaction. In this section, we first derive the Markov-perfect morale equilibrium in which each country i 's tariff rate depends only on the current levels of the state variables (the morale, agreed-upon tariff rate, and country-specific shock) and the discount factor δ common to all countries, i.e. $\tau_t^i = \tilde{\tau}(k_t, \alpha, \theta_t^i, \delta)$, provided that the agreed-upon tariff rate is exogenously given at α . We then consider the optimal binding rate of the tariff itself.

The value function that represents a country's average discounted payoff in the Markov-perfect morale equilibrium is given by

$$\begin{aligned}
& v(k_t, \alpha, \{\theta_t^i\}_{i=1}^n, \delta) \\
&= \max_{\tau_t^i} \left\{ (1-\delta) u(\tau_t^i, \{\tilde{\tau}(k_t, \alpha, \theta_t^i, \delta)\}_{i \neq t}, k_t, \alpha, \{\theta_t^i\}_{i=1}^n) \right. \\
&\quad \left. + \delta E \left[v \left(K \left([\tau_t^i - \alpha]_+, \sum_{j \neq i} [\tilde{\tau}(k_t, \alpha, \theta_t^j, \delta) - \alpha]_+ \right), k_t, \alpha, \{\theta_t^j\}_{j=1}^n, \delta \right) \right] \right\}, \quad (2)
\end{aligned}$$

where the expectation is taken with respect to the joint probability distribution of $\{\theta_t^i\}_{i=1}^n$.

Letting $d(k_t, \alpha, \{\theta_t^i\}_{i=1}^n, \delta) = \sum_{i=1}^n [\tilde{\tau}(k_t, \alpha, \theta_t^i, \delta) - \alpha]_+$ denote the function that represents the aggregate upward deviation in equilibrium, we can write the derivative of the expression to be maximized in (2), multiplied by $1/(1-\delta)$, as

$$\begin{aligned}
& \frac{\partial M}{\partial \tau_t^i}(\tau_t^i, \theta_t^i) - k_t + \frac{\delta}{1-\delta} E \left[\frac{\partial v}{\partial k_{t+1}} \left(K(d(k_t, \alpha, \{\theta_t^i\}_{i=1}^n, \delta), k_t), \alpha, \{\theta_t^i\}_{i=1}^n, \delta \right) \right. \\
& \quad \left. \times \frac{\partial K}{\partial d_t} \left(d(k_t, \alpha, \{\theta_t^i\}_{i=1}^n, \delta), k_t \right) \right],
\end{aligned}$$

if $\tau_t^i > \alpha$ and $\partial M / \partial \tau_t^i$ if $\tau_t^i < \alpha$. They are also the right-hand and left-hand derivatives at $\tau_t^i = \alpha$, respectively. Compared with (1), the derivatives are the same if $\tau_t^i < \alpha$, but if $\tau_t^i \geq \alpha$ the marginal cost of raising the tariff is higher in the current intertemporal setting by

$$\begin{aligned}
c(k_t, \alpha, \delta) &\equiv -\frac{\delta}{1-\delta} E \left[\frac{\partial v}{\partial k_{t+1}} \left(K(d(k_t, \alpha, \{\theta_t^i\}_{i=1}^n, \delta), k_t), \alpha, \{\theta_t^i\}_{i=1}^n, \delta \right) \right. \\
&\quad \left. \times \frac{\partial K}{\partial d_t} \left(d(k_t, \alpha, \{\theta_t^i\}_{i=1}^n, \delta), k_t \right) \right],
\end{aligned}$$

which is positive at least in the case where α is set at a relatively high level so that countries seldom violate the agreement. An increase in the marginal psychological costs induces more cooperation, which is beneficial to all countries. The only negative factor is that an increase in the marginal psychological costs may increase total psychological costs upon deviation. The former positive effect outweighs the latter negative one if the probability of deviation is small. Moreover, the total psychological costs may even decrease as a consequence of an increase in the marginal psychological costs if the induced reduction of deviation outweighs the direct effect of raising the marginal psychological costs. We will henceforth restrict our attention to the case where $\partial c / \partial k_t > 0$.

Now, the Markov-perfect morale equilibrium strategy is given by

$$\tilde{\tau}(k_t, \alpha, \theta_t^i, \delta) = \begin{cases} \tau^*(0, \theta_t^i) & \text{if } \theta_t^i < \theta(0, \alpha) \\ \alpha & \text{if } \theta(0, \alpha) \leq \theta_t^i \leq \theta(k_t + c(k_t, \alpha, \delta), \alpha) \\ \tau^*(k_t + c(k_t, \alpha, \delta), \theta_t^i) & \text{if } \theta_t^i > \theta(k_t + c(k_t, \alpha, \delta), \alpha). \end{cases}$$

Countries' deviations would lower the morale, which further invites more future deviations. Moreover, the size of the deviation becomes larger as the morale decreases, which in turn decreases the morale more than otherwise. Correctly recognizing these effects, countries deviate in fewer occasions with smaller size of deviation when the morale may decay intertemporally. We also note that deviations from the agreed-upon tariff have a domino effect; a deviation will induce more future deviation and the morale erosion induces further morale erosion.

PROPOSITION 2. *Given an agreed-upon tariff rate, countries may occasionally renege on the agreement, which induces the erosion of the morale and invites future deviations from the agreement. Correctly recognizing this domino effect, however, countries have more incentive to uphold the agreement.*

Having derived the Markov-perfect morale equilibrium, we turn to the discussion of the optimal choice of the tariff rate in the international agreement. We first derive the Pareto-optimal state-contingent tariff agreement, in which the binding rate of tariff is expressed as a function of each country's shock, i.e. the agreed-upon tariff rate of country i in period t can be written as $\alpha^s(\theta_t^i)$. Letting $\tilde{\tau}^s(k_t, \theta_t^i, \delta)$ denote country i 's Markov-perfect morale equilibrium strategy and v^s denote the value function in this case, we have

$$\begin{aligned} & v^s(k_t, \alpha^s, \{\theta_t^j\}_{j=1}^n, \delta) \\ &= \max_{\tau_t^i} \left\{ (1-\delta)u\left(\tau_t^i, \{\tilde{\tau}^s(k_t, \theta_t^j, \delta)\}_{j \neq i}, k_t, \alpha^s(\theta_t^i), \{\theta_t^j\}_{j=1}^n\right) \right. \\ & \quad \left. + \delta E \left[v^s \left(K \left([\tau_t^i - \alpha^s(\theta_t^i)]_+, \sum_{j \neq i} [\tilde{\tau}^s(k_t, \theta_t^j, \delta) - \alpha^s(\theta_t^j)]_+, k_t \right), \alpha^s, \{\theta_t^j\}_{j=1}^n, \delta \right) \right] \right\}. \end{aligned}$$

Given the Markov-perfect morale equilibrium strategy characterized in this equation, it is optimal to set $\alpha^s(\theta_t^i)$ at the smallest tariff rate from which country i will not deviate. Then there will be no deviation in equilibrium, so that the marginal psychological costs can be kept at the highest level $\bar{k}$. Consequently, the marginal future costs of deviation can be written as

$$c^s(\bar{k}, \delta) \equiv -\frac{\delta}{1-\delta} E \left[\frac{\partial v^s}{\partial k_{t+1}}(\bar{k}, \alpha^s, \{\theta_t^j\}_{j=1}^n, \delta) \frac{\partial K}{\partial d_t}(0, \bar{k}) \right].$$

Thus, we find that the optimal state-contingent agreement $\alpha(\theta^i)$ can be given by

$$\alpha^s(\theta_t^i) = \tau^*(\bar{k} + c^s(\bar{k}, \delta), \theta_t^i).$$

Countries always conform to the optimal state-contingent agreement, which enables them to keep the highest morale, which in turn enables them to engage in a high level of cooperation. In reality, however, it is almost impossible to implement this agreement, since to do so countries need to know other countries' country-specific shocks and to monitor their state-contingent tariff policies.

A more practical tariff agreement is to select α irrespective of the demand shocks. Let us suppose now that the morale starts at the highest level in period 1, i.e. $k_1 = \bar{k}$, when countries agree on the level of α and implement this agreement. Then, it is optimal to choose α so as to maximize $E[v(\bar{k}, \alpha, \{\theta_t^j\}_{j=1}^n, \delta)]$, where v is characterized by (2). Countries can agree on choosing α at a high level so that no country would deviate on any occasion. Although they can keep the highest morale by doing so, this agreement may be inefficient since the fruit from cooperation is limited. The optimal agreement, therefore, may allow each country to occasionally deviate from the agreement in order to enjoy more cooperation most of the time with a more ambitious agreement. The morale in such agreements deteriorates occasionally, and the deterioration may accelerate in some occasions due to the aforementioned domino effect.

The question is whether it is worthwhile to lower α despite it worsening the probability distribution of the morale in the future. Once $K(d(k_t, \alpha, \{\theta_t^j\}_{j=1}^n, \delta), k_t)$ is

given, the invariant probability distribution of the morale, which we represent by $F[\alpha]$ as a function of α defined on the support $[\underline{k}, \bar{k}]$ is determined. Regardless of the current morale level, the probability distribution of the morale converges to this invariant distribution. It is expected that if a small decrease of α dramatically worsens the invariant distribution of the morale, α should be kept high.

To find whether or not α should be chosen so as to avoid any deviation, we define $\bar{d}$ such that $K(d, \bar{k}) = \bar{k}$ for any $d \leq \bar{d}$, and define $\underline{d}$ similarly such that $K(d, \underline{k}) = \underline{k}$ for any $d > \underline{d}$. We also let α^* denote the smallest α such that $\text{Prob}\{d(\bar{k}, \alpha, \{\theta_i^j\}_{j=1}^n, \delta) \leq \bar{d}\} = 1$. The morale is kept at the highest level also in the next period if the current aggregate deviation falls short of $\bar{d}$, while it is kept at the lowest level if the aggregate deviation exceeds $\underline{d}$. If countries agree on the binding rate of tariff at a level higher than or equal to α^* , they can keep the highest morale perpetually even though they may occasionally deviate from the agreement.

We define the expected per-period payoff for a country by

$$w(k_t, \alpha) = E\left[u\left(\{\tilde{\tau}(k_t, \alpha, \theta_i^j, \delta)\}_{j=1}^n, k_t, \alpha, \{\theta_i^j\}_{j=1}^n\right)\right].$$

Then we have

$$\lim_{t \rightarrow \infty} v(k_t, \alpha, \{\theta_i^j\}_{j=1}^n, \delta) = \int_{\underline{k}}^{\bar{k}} w(k, \alpha) dF[\alpha](k),$$

which we call $\omega(\alpha)$. Assuming that F is differentiable with respect to α to avoid unnecessary complications, we have

$$\omega'(\alpha) = \int_{\underline{k}}^{\bar{k}} \frac{\partial w}{\partial \alpha}(k, \alpha) dF[\alpha](k) + \int_{\underline{k}}^{\bar{k}} w(k, \alpha) dF'[\alpha](k), \quad (3)$$

where $\int_{\underline{k}}^{\bar{k}} dF'[\alpha](k) = 0$ as $\int_{\underline{k}}^{\bar{k}} dF[\alpha](k) = 1$ for any α . The first term on the right-hand side of (3) represents the negative impact of an increase in α on the expected per-period payoff, while the second term shows the positive impact on the probability distribution of the morale.

Now, we show that $\omega(\alpha) < 0$, i.e. countries are better off by lowering α , if $\underline{k}$ and $\bar{k}$ are close to each other. As $\underline{k} \rightarrow \bar{k}$,

$$\int_{\underline{k}}^{\bar{k}} w(k, \alpha) dF'[\alpha](k) \rightarrow 0,$$

from $\int_{\underline{k}}^{\bar{k}} dF'[\alpha](k) = 0$. Thus, $\omega'(\alpha)$ converges to $(\partial w / \partial \alpha)(\bar{k}, \alpha) < 0$, and our claim follows immediately. If the range of the morale is small, countries should not be concerned much about the erosion of the morale. Indeed, if $\underline{k}$ and $\bar{k}$ are close enough, $\omega'(\alpha) < 0$ for any α . If this is the case, the optimal α is zero and countries constantly violate the agreement.

Countries may want to set $\alpha = \alpha^*$, on the other hand, if lowering α from α^* dramatically worsens the distribution of the morale. Let us consider the case in which $\text{Prob}\{d(\underline{k}, \alpha^*, \{\theta_i^j\}_{j=1}^n, \delta) > \underline{d}\} = 1$, i.e. once the morale reaches $\underline{k}$, it will stay there indefinitely. In this case, the invariant distribution of the morale puts all the probability on $\underline{k}$ if k_t reaches $\underline{k}$ in finite time with a positive probability. Thus, lowering α from α^* by an infinitesimal amount of ε will cause a discrete drop of the expected per-period payoff from $w(\bar{k}, \alpha^*)$ to $w(\bar{k}, \alpha^* - \varepsilon)$. The agreement should specify α^* as the binding rate if δ is large enough in such cases.

PROPOSITION 3. *Countries agree on specifying the binding rate of tariff as a level that is small enough to invite their occasional deviations and morale erosion, if the range of morale is small enough that the morale erosion would not lead to a serious retreat from the agreement.*

5. WTO Rules to Facilitate International Cooperation

We have shown that the WTO gives moral support to member countries in tariff cooperation by making them feel obliged to keep promises in the agreement under the auspices of the WTO. In this section, we show that the safeguards policy and continuing GATT/WTO trade negotiation rounds also play positive roles in international tariff cooperation.

The safeguards policy in our context is considered as a system to allow countries to deviate from an agreed-upon tariff if and only if they observe high import demand shocks. More specifically, the agreement α is tailored so that α takes a fixed number, say $\bar{\alpha}$, if the demand shock is smaller than a critical level, but it is an aforementioned state-contingent plan if the demand shock exceeds the critical level. Countries would not feel guilty as long as they select a tariff that is smaller than or equal to the prescribed level since “deviations” from $\bar{\alpha}$ when the demand shocks are high are now considered to be legal in the WTO framework, but continue to feel guilty if they select a higher tariff. Since “deviations” are authorized by the WTO, they are not likely to lower the morale and countries can keep the highest morale enjoying the fruit from deeper cooperation.

Continuing GATT/WTO negotiation rounds also play a positive role. Here we focus on two aspects of GATT/WTO rounds. First, the conclusion of a round may refresh countries’ attitudes to international cooperation. As Kandori (2003) also mentioned, Andreoni (1988) reported that in his experiment of a public-good contribution game, subjects return to selecting high levels of contribution in the restart of the repeated play. As we have seen in the last section, the morale can be eroded over time. Continuation of the GATT/WTO rounds is likely to make possible the cooperation framework to restart with the morale reset at the highest level. Secondly, a new round may include new issues to be negotiated, as the GATT/WTO history suggests. The morale is affected by any deviation in any industry that is included in the agreement. As countries add new agreements on new issues, such as tariff cooperation in industries that have not been included in the past agreement, the number of industries that would be affected by an induced erosion of the morale increases. A resulting rise in the future costs of deviation gives countries more incentive to keep agreements. Consequently, countries can even lower the tariff rates that have been agreed-upon in previous negotiations. This last phenomenon accords with the history of the actual GATT rounds.

6. Concluding Remarks

Countries have cooperated in trade policies more-or-less successfully in these decades, owing to the “twin engines of international obligation and retaliation” (Hudec, 1990). Although the WTO authorizes retaliation on some occasions, countries need not be approved to retaliate against other countries’ deviations. Thus, the role of the WTO can be considered to be greater in the first engine of international obligation rather than in the second engine of retaliation. We have formalized the cooperation mechanism in which the mere existence of the WTO helps countries cooperate as (1) they feel guilty when they deviate and (2) their deviations may lead to the erosion of the morale inviting the implicit future punishment of further deviations by all countries.

The morale may be eroded if countries fail to punish other countries for their deviations if the WTO has authorized the retaliation. Twin engines are related to each other in interesting ways, which is beyond the scope of the present paper and is left for future research.

References

- Andreoni, James, "Why Free Ride? Strategies and Learning in Public Goods Experiments," *Journal of Public Economics* 37 (1988):291–304.
- Bagwell, Kyle and Robert W. Staiger, "A Theory of Managed Trade," *American Economic Review* 80 (1990):779–95.
- , "An Economic Theory of GATT," *American Economic Review* 89 (1999):215–48.
- , "Multilateral Trade Negotiations, Bilateral Opportunism and the Rules of GATT/WTO," *Journal of International Economics* 67 (2005):268–94.
- Benckekroun, Hassan and Ngo Van Long, "The Build-Up of Cooperative Behavior among Non-Cooperative Selfish Agents," *Journal of Economic Behavior and Organization* 67 (2008):239–52.
- Dixit, Avinash, "Strategic Aspects of Trade Policy," in Truman F. Bewley (ed.), *Advances in Economic Theory: Fifth World Congress*, Cambridge: Cambridge University Press (1987).
- Ethier, Wilfred J., "Escape and Entry Mechanism in the Multilateral Trade System," PIER working paper 02-009, University of Pennsylvania (2002).
- Hudec, Robert E., "Dispute Settlement," in Jeffrey J. Schott (ed.), *Completing the Uruguay Round: A Result-Oriented Approach to the GATT Trade Negotiations*, Washington, DC: Institute for International Economics (1990).
- Jackson, John H., *The World Trading System: Law and Policy for International Economic Relations*, 2nd edn, Cambridge, MA: MIT Press (1997).
- Kandori, Michihiro, "The Erosion and Sustainability of Norms and Morale," *Japanese Economic Review* 54 (2003):29–48.
- Kandori, Michihiro, George J. Mailath, and Rafael Rob, "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica* 61 (1993):29–56.
- Kovenock, Dan and Marie Thursby, "GATT, Dispute Settlement and Cooperation," *Economics and Politics* 4 (1992):151–70.
- , "GATT, Dispute Settlement and Cooperation: A Reply," *Economics and Politics* 9 (1997):95–98.
- Maggi, Giovanni, "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review* 89 (1999):190–214.
- Schwartz, Warren F. and Alan O. Sykes, "The Economic Structure of Renegotiation and Dispute Resolution in the World Trade Organization," *Journal of Legal Studies* 31 (2002):S179–S204.
- Young, H. Peyton, "The Evolution of Conventions," *Econometrica* 61 (1993):57–84.

Notes

1. What is important to simplifying our analysis is that the total surplus for a country can be written as the sum of the import surplus as a function of its own tariff and the aggregate export surplus whose component is a function of the importing country's tariff rate. This requirement is also satisfied in an alternative model in which the non-numéraire good industry is imperfectly competitive and the markets are segmented by national borders. In this case, the surplus derived from the domestic market is considered as the import surplus while the surplus derived from the foreign markets is considered as the export surplus. We should emphasize, however, that our analysis can be applied to many other trade situations with appropriate modifications. In particular, we make the assumption of symmetry only to simplify the exposition of the analysis.
2. Shocks can be on either production or consumption in the import-good industry. In the alternative model suggested in note 1, shocks should be on consumption since domestic production shocks would also affect the export surplus.