

# Dispute Settlement and Power Asymmetries in International Trade: Regional Trade Agreements and the WTO

Yoo Sun Jung

*University of California San Diego*

*ysjung@ucsd.edu*

January, 2021

## **Abstract**

Regional trade agreements (RTAs) have become increasingly common in the international trading system since the end of the Second World War. The number of RTAs has increased rapidly during the past two decades. In particular, it doubled to reach more than 400 RTAs in force from 2000 to 2016. Notwithstanding the recent surge in regionalism, dispute settlement mechanisms (DSMs) are much less in use in RTAs than in the World Trade Organization (WTO). I argue that the frequent use of the WTO-DSM results from its superiority in avoiding systemic breakdown and leveling the playing field for less powerful countries in disputes. The extension of Rosendorff's (2005) model of the WTO to RTAs demonstrates that RTA-DSMs increase the power disparity by providing more flexibility to stronger countries at the expense of weaker countries. Therefore, DSMs generate a trade-off in gains between large and small countries in RTAs instead of the trade-off between "rigidity and stability" in the WTO. The significant loss sustained by weak countries in RTA-DSMs generate the loss in the stability of the regional trading system. The findings imply that less powerful members insulate power politics in the multilateral trading system under the WTO more efficiently than in regional trade agreements.

**Word Count**— 201 (abstract), approx. 9,450(manuscript)

**Keywords**— International Organization, WTO, Regional trade agreements, RTAs, Dispute Settlement, Power Asymmetries

## 1 Introduction

Along with the rise of global economic interdependence, regional trade agreement (RTA) have become increasingly common in the international trade system after the end of the Second World War.<sup>1</sup> The number of RTAs has increased rapidly during the past two decades. From 2000 to 2016, the number of RTAs doubled to reach more than 400 RTAs in force (WTO 2016). All WTO members are now involved in RTAs in force following the agreement between Mongolia and Japan in June 2016. The recent surge in regionalism is also noticeable. In the years between 2011 and 2016, 16 RTAs have been newly created per year, on average.<sup>2</sup> The proliferation of RTAs is neither limited by geographic boundaries nor the level of economic development. RTAs have spread throughout the world between and within continents, and among developed and developing states. Most nations are involved in one or more RTAs; even nations without any RTAs are not completely outside the range of RTAs. As Bhagwati and Krueger (1995) have put it, we now live in a world like a “spaghetti bowl” where every nation is connected with different types of RTAs in a very complex way.

Although there is a great variation, almost all RTAs have systems of dispute resolution. Despite the continuing surge in RTAs with DSMs, however, previous studies found that the WTO has been used more frequently than RTA for dispute resolution (Chase et al., 2013). Consider the United States-Korea Free Trade Agreement (KORUS FTA) which entered into force on March 15, 2012. The bilateral agreement contains dispute settlement mechanisms (DSMs) in two formal ways: state-to-state disputes about the imposition of tariffs or non-tariff trade barriers or investor-state disputes over breach of an investment agreement. Since its establishment, the bilateral trade agreement between the U.S. and Korea has not listed any dispute which

---

<sup>1</sup>In this paper, regional trade agreement (RTA) refers to all types of non-multilateral agreements: free trade areas, customs unions, common markets, and economic unions. Most RTAs in force are free trade agreements (FTAs) and partial scope agreements, which account for 90% of the total 423 RTAs in force and notified to the WTO. Customs unions account for only 10% of these RTAs. RTAs also include international trade pacts across different regions such as EU-Mexico and Korea-Chile.

<sup>2</sup>Fifteen new agreements were notified to the WTO in 2011, twenty-four in 2012, twenty-two in 2013, twelve in 2014, ten in 2015, and thirteen in 2016 (the WTO RTA Database, <http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx>. Accessed May 19, 2018).

came through panel processes.<sup>3</sup> In comparison, Korea has brought two cases against the U.S. to the WTO which reached panel establishment since 2012.<sup>4</sup> Before the entry into force of the U.S.-Korea trade agreement, there were a total of fourteen cases between the U.S. and South Korea brought to the WTO. Out of fourteen, five cases were brought by the U.S. and nine cases were brought by Korea.<sup>5</sup>

Like the KORUS-FTA, most dispute settlement mechanisms found in regional trade agreements (RTAs) are inactive.<sup>6</sup> For example, there are only four cases brought to the Caribbean Common Market (CARICOM) since its establishment in 1973. Those four disputes are between individual investors and states.<sup>7</sup> In Africa, the Economic Community of West African States (ECOWAS), the Southern African Customs Union (SACU), and the Southern African Development Community (SADC) oversaw a few cases which are either investor-state disputes or resolved before panel process. In Asia, the Association of Southeast Asian Nations (ASEAN) has never witnessed a case brought to its DSM.

The empirical irregularity of disuse of RTA DSMs raises a series of questions: Why do states choose the WTO over regional trade agreements to resolve their disputes? Does the states' choice of forum indicate that the WTO's dispute settlement mechanism (DSM) is more legitimate and effective than that of RTAs? The relative disuse of RTA DSMs does not mean that an analysis of RTA DSMs is unimportant, rather, it leads to the puzzle of this paper: Why do states create DSMs in regional trade agreements if they do not use these DSMs?

---

<sup>3</sup>The Joint Committee has met three times at the Ministerial level. However, there does not exist any official records beyond the state.

<sup>4</sup>DS464 "Anti-dumping and countervailing measures on large residential washers from Korea" (requested consultations on 29 August 2013) and DS468 "Anti-dumping measures on certain oil country tubular goods from Korea (requested consultations on 22 December 2014)." The panel was composed for both cases and the U.S. appealed to the Appellate Body for DS464.

<sup>5</sup>WTO RTA Database. <http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx>. Accessed May 20, 2018.

<sup>6</sup>There are some active DSMs such as the European Court of Justice (ECJ), the European Free Trade Association (EFTA), Preferential Trade Area for Eastern and Southern American Common Market (Mercosur), the Andean Community, and the Central American Common Market (CACM). For example, the North American Free Trade Agreement (NAFTA) lists more than 200 disputes and the Preferential Trade Area for Eastern and Southern American Common Market (Mercosur) lists more than 300 cases. I will come back to these cases later in section 6.

<sup>7</sup>Before the inauguration of the Caribbean Court of Justice (CCJ) in 2005, an ad hoc tribunal was considered as the dispute settlement mechanism of CARICOM, which has never been used (Chase et al., 2013, 46).

Seeking answers to the puzzle, this paper compares the design and effects of dispute settlement systems between RTAs and the WTO. Rosendorff (2005) shows that the WTO's dispute settlement system brings more flexibility to the multilateral trading system and hence, enhances systematic stability. In this chapter, I extend Rosendorff's game theoretic model of the WTO to RTAs to find out what leads states to use the WTO over RTAs to resolve their disputes. I examine the role of dispute settlement mechanisms across trade regimes in shaping states' bargaining strategies. The model demonstrates that states will face losses of their strategic flexibility in RTAs more than in the WTO. The greater loss of flexibility states face in RTAs compared to the WTO leads to an increase in the per period probability of breakdown and, consequently, reduces stability of the trading system.

Additionally, RTA's dispute settlement mechanisms reinforce inequalities more than the WTO's dispute settlement mechanism. Small and poor countries can more efficiently insulate themselves from power politics under the WTO than RTAs. This result brings us to what lies at the root of the relative disuse of RTA DSMs. The frequent use of the WTO results from its superiority in avoiding system breakdown and leveling the playing field for small and poor countries, compared to RTAs.

## 2 Previous Research on Dispute Settlement Mechanisms

While regionalism gathered momentum over the last two decades, the legalization of international institutions also attracted the attention of academics as another salient trend in international trade systems. The abundant literature on international adjudication examines whether legalization of international institutions increases systemic stability or promote international cooperation (e.g. Maggi, 1999; Keohane et al., 2000). Previous studies on DSMs have been largely carried out on the multilateral trade system, the WTO (e.g. Kucik and Pelc, 2016; Rosendorff, 2005).

Some argue that the WTO, vested with more legal power through the dispute settlement system, increases the level of international cooperation and promotes trade liberalization. There

are different views on how the procedure facilitates cooperation. On the one hand, legalization of the WTO with the stick of the highly legalized dispute settlement mechanism prevents member states from defection and increases the compliance rates (e.g. Smith, 2000; Goldstein et al., 2000). A harsher punishment against defection by member states will strengthen the bargaining power of developing countries (e.g. Park, 2000).<sup>8</sup> Legalization of international organizations allows the transmission of information to voters about a state's trade behavior, which increases the compliance rate of member states (e.g. Chaudoin, 2010; Fang, 2008).

On the other hand, the legalized dispute settlement mechanism provides opportunities for members to manipulate the system, acting as a carrot rather than a stick (e.g. Rosendorff and Milner, 2001; Fearon, 1998). For instance, Rosendorff (2005) shows that the WTO's dispute settlement mechanism enhances the stability of the multilateral trading system by bringing flexibility to the system. The underlying logic is that the dispute settlement mechanism allows states to temporarily violate in response to domestic political pressure for trade protection.

Unlike this positive view, other scholars argue that the WTO's strong dispute settlement mechanism deters the use of harsh punishment and causes inefficient outcomes in trade negotiations. Ludema (2001) argues that the dispute settlement system disables the use of strong punishment against defections, which will reduce the level of national welfare. Bown (2005) shows that under the WTO, the system of dispute settlement discourages developing countries' participation in legal disputes. The WTO's dispute settlement mechanism has institutional bias due to the self-representative and self-enforcing features. Small countries lack resources to monitor violations and participate in costly legal proceedings. Tariff retaliation is also useless due to their lack of economic influence. Busch and Reinhardt (2003) argue that the dispute settlement system not only reinforces the power disparity of the trade regime but also expands the gap between rich and poor member states in their use of the WTO DSM.

In addition to the WTO DSM, RTAs usually have dispute settlement mechanisms. In all

---

<sup>8</sup>"Irreversible investment" strengthens the punishment power of the small countries disproportionately, which helps raise the bargaining position for small countries (Park, 2000).

the overlapping areas between the WTO and RTAs, such as intellectual property, services, and investment, RTAs include chapters which discuss dispute settlement procedures. Empirical studies on RTA DSMs have continuously tried to understand the role of DSMs used in RTAs and the factors which determine the level of legalism in RTAs (e.g. Smith, 2000; Chase et al., 2013; Jo and Namgung, 2012).

Yet, studies on RTA DSMs have not yet fully explained their dynamics but largely focus on potential conflicts between dispute settlement mechanisms across trade regimes (e.g. Hillman, 2009; Marceau and Wyatt, 2010). In the study of regionalism and multilateralism, there have been a handful of game theoretical analyses (e.g. Riezman, 1985; Aghion et al., 2007). In this chapter, I construct a model of trade disputes in RTAs building on Rosendorff's (2005) model of WTO dispute settlement mechanism. In the framework of a repeated Prisoners' Dilemma game, I analyze how dispute settlement mechanisms shape state behavior across trade regimes. I examine structural differences in the institutional design of dispute settlement system between RTAs and the WTO and the consequences of the differences in terms of states' strategic choices. This study deepens our understanding of the dynamics and the new features of the current global trading system.

### 3 Prisoners' Dilemma: Trade Game with Dispute Settlement Mechanism

I extend Rosendorff's analysis of the WTO DSM which shows that the introduction of the DSM to the WTO system allows states to "temporarily suspend their obligations in periods of unexpected, but heightened, domestic political pressure for protection" (389). I briefly summarize the setup of his model and then move to the extended model for RTA DSMs.

For the sake of simplicity, the model focuses on bilateral trade relations between two state actors. Players simultaneously choose their tariff levels in response to domestic political pressures. Along with the choice of tariffs, players have three action profiles: cooperation,

settlement, and defection. When both players keep their tariffs as low as the agreed level, i.e., both players choose cooperation strategies, they both receive gains from free trade. If a player imposes a tariff higher than the binding tariff level, the cooperative party suffers from its loss from trade. The amount of compensation paid by the violating party is calculated based on the difference in the complainant's expected payoff when both players cooperate and when only the complainant cooperates. The violating player decides whether it pays compensation or not if it loses the case. Why does the player choose to pay compensation? By paying compensation, the violating player stays in the system and benefits from free trade. Given that the compensation is made only if the panel finds in favor of the complainant, the loss function is weighted by the probability the complainant prevails, denoted by  $\theta$  which is a common probability.

A state that chooses the defection strategy imposes a tariff which is higher than the agreed tariff, i.e.,  $t > t^c$  or  $\tau > \tau^c$ , and is followed by non-compliance and exit from a treaty. Therefore, the defection strategy leads to the breakdown of the bilateral trade agreement which eliminates gains from cooperation. If a player chooses settlement as a strategy, the player sets a higher tariff than the agreed tariff level, complies with panel rulings (by paying compensation if it loses the case), and stays in the trading system. This game is infinitely repeated.<sup>9</sup> Also, this game assumes that each player adopts the *Grim Trigger* strategy so that a deviation is followed by infinite punishment.

Then, why does a player ever want to defect from the agreement despite future losses (i.e. the expected gains from cooperation)? As the name, *Prisoners' Dilemma*, strongly suggests, it is because each player has incentives to renege on agreements.<sup>10</sup> Then where do the incentives come from? In order to answer the question, let us consider the decision making process of the players. When players choose tariff levels simultaneously in each round, they take four

---

<sup>9</sup>States discount their future payoffs according to the discount factor  $0 < \delta < 1$ . The model allows the possibility that the benefits of cooperation may grow larger over time. Along with globalization, the multilateral trade regime of the WTO expands and the number of regional trade agreements continuously increases. Greater economic integration accelerates trade liberalization by lowering transaction costs, reallocating resources, and spreading the ever-developing technologies across the globe. Taken all together, the benefits of freer trade grow over time.

<sup>10</sup>The incentives to defection result in the failure of cooperation at the equilibrium.

components into account: welfare of consumers, profits of import and export-competing firms, and tariff revenues.

The higher the tariff the home government chooses, the greater the profit of import-competing firms.<sup>11</sup> By contrast, the profit of exporters is influenced by the foreign tariff. The higher the tariff of the foreign government on exporting goods from the home country, the lower the profit of the exporters at home. Changes in tariffs also affect consumers, too. An increase in tariffs raises the price of imported goods, thereby reducing consumer surplus in the home country.<sup>12</sup> Given that, government incentives to impose higher tariffs result from political pressures for protection that a government experiences. Political pressure varies in a stochastic manner over time and is captured by the weights a state puts on the interests of the import-competing sectors ( $a$  and  $\alpha$  for the home and foreign country, respectively).<sup>13</sup> Thus, the higher the political pressure, the greater the government gains from imposing high tariffs on imported products. After observing its level of political pressure, each player simultaneously chooses its tariff ( $t, \tau$ ) at the beginning of each round of the game. In other words, after observing incentives to defection, each player simultaneously chooses a strategy: imposing cooperative tariffs, defective tariffs with compensation, and defective tariffs without compensation.

Let  $C$  and  $C^*$  denote the payoffs of a strong and weak government, respectively, when both are cooperating by choosing a tariff that is lower than or equal to the agreed tariff level. Let  $N$  and  $N^*$  denote the payoffs of a strong and weak government, respectively, when both violate the tariff binding by choosing a tariff that is higher than the agreed tariff. Let  $D$  and  $D^*$  denote the payoffs of a strong and weak government, respectively, when it violates the tariff agreement it has with its trading partner. Let  $S$  and  $S^*$  denote the payoffs of a strong and a weak government, respectively, when it cooperates by choosing a tariff lower than the agreement while

---

<sup>11</sup>Firms' profits are assumed to be an increasing concave function.

<sup>12</sup>Changes in foreign tariffs affect not consumers at home but foreign consumers by raising prices of exported goods which foreign consumers face.

<sup>13</sup>The level of political pressure varies from round to round to represent changes in economic conditions and they are independently and identically distributed on  $(0, +\infty)$  with a cumulative distribution function (cdf) of  $\Phi$ . Each state knows its level of pressure but not the other's at the beginning of each period.



its trading partner violates the tariff binding. The following preference order creates a Prisoners' Dilemma game:  $D > C > N > S$ .<sup>14</sup> In this classic PD game, a Nash Equilibrium (NE) exists when both players defect from their agreed tariff by choosing higher tariffs in a one-shot game. In the infinitely repeated game, both players choose different tariffs which create an equilibrium alternative to defection from both sides.

The dispute settlement strategy (DSS) is defined as a strategy in which a player having drawn politics of type  $\hat{a}$ , plays *Defect* if the other player has played *Defect* in any period in the past; otherwise the state plays *Cooperate* if the politics type is smaller than its lower cut-off point  $\hat{a} < \underline{a}$ , plays *Settle* if the politics type is between the lower and the upper cut-off points  $\underline{a} \leq \hat{a} \leq \bar{a}$ , and plays *Defect* if the politics type is greater than its upper cut-off point  $\hat{a} > \bar{a}$ . (Rosendorff, 2005, 394).<sup>15</sup>

### 3.1 Extension

In the model extension, I differentiate players regarding their relative power. Power asymmetries in the model allow us to examine the different impacts of DSMs on trade strategies between strong and weak players. Without a power disparity, all members in the trade model have the same dispute settlement strategy if the given political pressure level is the same. In other words, the same level of political pressure leads to the same strategy of all players: Cooperation (*C*), Settlement (*S*), or Defection (*D*). If one state cooperates at a certain level of pressure, other states also choose to cooperate when facing that same pressure. By differentiating the players' relative power, the model allows the possibility that a certain level of political pressure could provide enough incentive for a player to defect but fail to incentivise the other player to defect from cooperation. Therefore, the extended version is more flexible in addressing heterogeneous players in the game.

---

<sup>14</sup>Players have only two actions in their action profiles: Cooperation (*C*) and Defection (*D*), which creates a classic 2 by 2 PD game. In the following section, players have an additional action, Settlement (*S*), which creates a 3 by 3 PD game.

<sup>15</sup>Rosendorff showed that a pair of the dispute settlement strategy (DSS) is a Nash Equilibrium. See Rosendorff (2005) for details.

Power disparity among the players in the model also helps incorporate attributes of RTAs to the trade game in which the Rosendorff's WTO model does not capture. In the extended model, power is associated with the likelihood that each player prevails in the case. Let  $\lambda$  ( $\lambda^*$ ) denote the probability that the panel finds in favor of the strong (weak) plaintiff such that  $0 < \lambda$  ( $\lambda^*$ )  $< 1$ . The WTO model does not distinguish between players so that all players are equally likely to win the case (denoted by  $\theta$ ). In the WTO model, the case is judged by its merits, while cases are biased by power in the RTA model.<sup>16</sup>

The extended RTA model also captures the burden of litigation costs. Disputing parties suffer from costs that international dispute settlement procedures incur. RTA DSMs entail disparity in the burden of costs attributable to the differences in financial and legal resources at their disposal. The burden of legal expenses in DSM processes for a member country is denoted by  $\kappa$  ( $\kappa^*$ )  $> 0$  for a stronger (weaker) player. I will come back to those terms later in subsection 4.1 with further details.

Recall that players must simultaneously choose tariffs in response to stochastic domestic political pressure. Both players have three action profiles. A state might cooperate ( $C$ ), settle a dispute through a DSM ( $S$ ), or defect by non-compliance ( $D$ ). In other words, a state plays  $C$  by setting a tariff lower than or equal to the agreed level,  $t \leq t^c$  or  $\tau \leq \tau^c$ . By playing  $S$ , a state violates the agreement by setting its tariff level higher than the tariff agreement,  $t > t^c$  or  $\tau > \tau^c$ , and pays a compensation if it loses a dispute. A state plays  $D$  if a state imposes a tariff such that  $t > t^c$  or  $\tau > \tau^c$  and does not pay compensation in the event of losing a case. Figure 1 is a  $3 \times 3$  normal form representation of the one-shot trade game in an extended model setting described above.

Each period, a state decides tariffs based on the cost and benefit calculation associated with these actions profiles. Suppose states exchange their goods and services in accordance with their choice of tariffs. Then, each state is always better off by imposing a high tariff. Both players

---

<sup>16</sup>If we assign the same probability of winning to all players in the game  $\lambda = \lambda^* = \theta$  (because they are differentiable by power), the game is equivalent to the original WTO model of Rosendorff.

**Figure 1:** Payoffs in the  $3 \times 3$  normal form regional trade game

		Weak		
		$C^*$	$S^*$	$D^*$
$C$	$C(a)$		$S(a, \hat{\alpha}) + \lambda L(\hat{\alpha}) - \kappa$	$S(a, \hat{\alpha}) - \kappa$
	$C^*(\alpha)$		$D^*(\hat{\alpha}) - \lambda L(\hat{\alpha}) - \kappa^*$	$D^*(\hat{\alpha}) - \kappa^*$
Strong $S$	$D(\hat{\alpha}) - \lambda^* L^*(\hat{\alpha}) - \kappa$	$N(\hat{\alpha}, \hat{\alpha}) + \lambda L(\hat{\alpha}) - \lambda^* L^*(\hat{\alpha}) - \kappa$	$N(\hat{\alpha}, \hat{\alpha}) - \lambda^* L^*(\hat{\alpha}) - \kappa$	
	$S^*(\hat{\alpha}, \alpha) + \lambda^* L^*(\hat{\alpha}) - \kappa^*$	$N^*(\hat{\alpha}, \hat{\alpha}) + \lambda^* L^*(\hat{\alpha}) - \lambda L(\hat{\alpha}) - \kappa^*$	$N^*(\hat{\alpha}, \hat{\alpha}) + \lambda^* L^*(\hat{\alpha}) - \kappa^*$	
$D$	$D(\hat{\alpha}) - \kappa$	$N(\hat{\alpha}, \hat{\alpha}) + \lambda L(\hat{\alpha}) - \kappa$	$N(\hat{\alpha}, \hat{\alpha}) - \kappa$	
	$S^*(\hat{\alpha}, \alpha) - \kappa^*$	$N^*(\hat{\alpha}, \hat{\alpha}) - \lambda L(\hat{\alpha}) - \kappa^*$	$N^*(\hat{\alpha}, \hat{\alpha}) - \kappa^*$	

have incentives to set tariff levels as high as possible, above the agreed level. The situation, therefore, portrays the issue of Prisoners' Dilemma. If the game repeats, however, players face expected future losses of gains from cooperation. In an infinitely repeated game with grim trigger punishments, a one-time violation without complying with panel rulings, i.e. playing  $D$ , leads the other side to play  $D$  forever. Therefore, this game shows a dynamic of states' choices in trade behavior in response to varying levels of domestic political pressure.

Before we move to the dispute settlement strategy (DSS) in an extended mode, I clarify a couple of terms:  $B$  and  $\Lambda_{DS}$ . First,  $B(a)$  (for a stronger player,  $B(\alpha)$  for a weaker player) represents the gains from defection, i.e.  $B(a) \equiv D(a) - C(a)$ .<sup>17</sup> An increase in political pressure to protect domestic industries raises gains from playing ( $D$ ) relative to playing  $C$ . Second,  $\Lambda_{DS}$  captures the difference in expected benefits between playing cooperation strategy - either  $C$  or  $S$  - and defective strategy  $D$ . The expected benefits are the aggregate payoffs from gains in the current period and continuation value of playing the strategy.<sup>18</sup> Intuitively, large gains from defection and high costs of litigation lead states to forgo cooperation forever.

The first cutpoint in the DSS above determines whether a player cooperates or defects without the intention of entirely withdrawing from the treaty. In order to do so, the state needs to know how much it will gain from a violation and how much it will lose from their payment of

<sup>17</sup>  $B'(a) > 0$  is assumed, which implies that the gains from defection increase as political pressure grows.

<sup>18</sup> I assume that  $(1-p)(\lambda - \lambda^*)L + p^2\kappa$  is small enough. For details, see the appendix.

compensation if they lose the resulting dispute. If a player gains more by temporary defection than by cooperation, it will defect once and stay in the treaty. This means that a player given a draw  $\hat{a}$  prefers (temporary) defection over full cooperation when the sum of gains from a violation, the expected loss from compensation to pay, and the costs involved in DSM-litigation outweighs its expected cooperation value. Compensation is the amount of loss the other - cooperative - player gets paid by the violator when the complainant wins the case with probability of  $\lambda^*$ .

Therefore, a player compares its expected cooperation value ( $\underline{C}$ ), with the suspension of their obligations ( $D(\hat{a}) - \lambda^* L(\hat{a}) - \kappa$ ). For example, a member country given a draw  $\hat{a}$  defects without exit from the treaty if  $D(\hat{a}) - \lambda^* L(\hat{a}) - \kappa > \underline{C}(\hat{a})$ , the expected value of cooperation. Now, we define the first cutpoint,  $a_1$ , such that  $D(a_1) - \lambda^* L(a_1) - \kappa = \underline{C}(a_1)$  as a lower bound for the player to temporarily defect followed by staying in the treaty forever, i.e., the settlement strategy (S). In other words,  $\lambda^* L^*(a_1) + \kappa = B(a_1)$ .

The second cutpoint in the DSS determines whether a violating member country pays compensation which is set by a panel in order to stay in the treaty forever or exits from the treaty by defying the panel ruling when the panel finds in favor of the plaintiff. Therefore, the second cutpoint  $a_2$  satisfies  $\lambda^* L^*(a_2) + \kappa = \wedge_{DS}$  where  $\wedge_{DS}$  is defined as “the level of the cost such that, if the government plays the cooperate strategy (either C or DS) into the indefinite future, the expected (net) benefits from doing so are equal to the expected benefits of defecting once and exiting the system forever.” (Rosendorff, 2005, 394).<sup>19</sup> In other words, a player given a draw  $\hat{a}$  will settle by paying compensation in order to stay in the system after defection rather than to exit the agreement, when  $\lambda^* L^*(a_2) + \kappa < \wedge_{DS}$ . The DSS in an extended regional trade model is defined as the following and described in Figure 4 from the stronger player perspective.

**Definition 1.** A Dispute Settlement Strategy (DSS) is a strategy in which a strong (weak) party

<sup>19</sup>  $\lambda^* L^*(a_2) + \kappa = \frac{\delta}{1-\delta} (p^2(N-S-D+C) + p(D-2N+S)) \equiv \wedge_{DS}$  (where  $p = Pr(a < a_1)$ , and  $I = \int_a \int_a I(a, \alpha) d\Phi d\Phi$  for  $I = D, N, S, C$ ).

with politics type  $\hat{a}$  ( $\hat{\alpha}$ ) drawn, plays **D** if **D** has been played by the other player in any period in the past; otherwise it plays **C** if  $\hat{a} < a_1$  ( $\hat{\alpha} < \alpha_1$ ) by setting  $t = t^c$  ( $\tau = \tau^c$ ), plays **S** if  $a_1 \leq \hat{a} \leq a_2$  ( $\alpha_1 \leq \hat{\alpha} \leq \alpha_2$ ) by setting  $t = t^D(\hat{a})$  ( $\tau = \tau^D(\hat{\alpha})$ ) and paying compensation if it loses, and plays **D** if  $\hat{a} > a_2$  ( $\hat{\alpha} > \alpha_2$ ) by setting  $t = t^D(\hat{a})$  ( $\tau = \tau^D(\hat{\alpha})$ ) and not paying compensation if it loses.

**Proposition 1.** *A pair of DSS strategies is an Nash Equilibrium (NE).*

The following section shows the dynamic of changes in the DSS from that of the WTO, made by the elements exclusive to RTAs: (1) the probability that a panel finds in favor of a plaintiff denoted by  $\lambda$  and  $\lambda^*$ , and (2) the burden of litigation cost denoted by  $\kappa$  and  $\kappa^*$ . Those terms represent two different ways by which power asymmetries between disputing parties influence the structure of the game. Adding one additional RTA component at a time to the WTO model allows us to examine the impact of each element of RTAs on strategies between the strong and the weak.

## 4 Power Asymmetries in RTA DSMs

### 4.1 Winning a Dispute

I analyze the effect of different chances of winning as a complainant between the strong and the weak, without taking the costs of DSM litigation into consideration. The extended regional trade model assumes that a member country's chance of winning in a dispute depends on the relative size of power among disputants, while the chance is equal to every member in the WTO model. Therefore, power asymmetries are transferred into the extended model through disparate chances of winning a dispute.

First, weaker members in disputes against stronger partners are more likely to be vulnerable to the economic and diplomatic consequences of being involved in dispute resolution. Consider a dispute between the U.S. and Mexico. The dispute will have a bigger impact on the economy of Mexico than the economy of the U.S. The deterioration of economic conditions along with aggravation of the relationship with the trading partner is more likely to bring an election

outcome against incumbents in Mexico than the U.S. Therefore, richer countries are more likely to be able to derive concession from weaker countries before a panel ruling is issued. Busch and Reinhardt (2003) find that in RTA DSMs, larger and richer countries disproportionately win disputes during consultations or panel deliberations prior to rulings at the expense of smaller and poorer countries.

Second, RTA DSMs are more power-based than the WTO DSM. Despite similarities in the design of DSMs, RTAs and the WTO are different in the level of legalism in their DSMs.<sup>20</sup> Most RTA DSMs are considered to be a medium level of legalism. They are less likely to hold impartial third-party review and require ad hoc panels.<sup>21</sup> The WTO DSM is considered to be a high level of legalism. It consists of a standing body, a so-called dispute settlement body (DSB), and the process of appellate body (AB) review. Thus, the system of dispute resolution under the WTO is considered to be more rule-based and more effective in offsetting power disparities across countries than RTAs (Davey, 2006). Under the power-based structure of RTA DSMs, panel rulings under RTAs tend to be more vulnerable to nationality bias and considered less legitimate than that of the WTO. Therefore, the WTO DSM mitigates asymmetries in bargaining power among members with different sized economies more effectively than more power-based RTA DSMs.

Given the reasons above, I expect that the design of dispute resolution in RTAs yields more disparity in the rate of winning disputes than in the WTO. Weaker and smaller countries disproportionately fail to win in disputes under the more power-based system of RTA DSMs than the more rule-based system of WTO DSM. The stronger country's chance of winning in a regional trade dispute is higher than its chance in a WTO dispute which is equal for every member country. Also, the weaker party has a lower chance of winning in disputes against a stronger party in RTAs than its chance in the WTO. The following assumption shows a clear-cut logic of such disparities

---

<sup>20</sup>Previous studies find that the design of dispute resolution in RTAs and that of the WTO are similar to each other due to emulation. Jo and Namgung (2012) show the effect of emulation on the design of DSMs in regional trade agreements by sharing memberships.

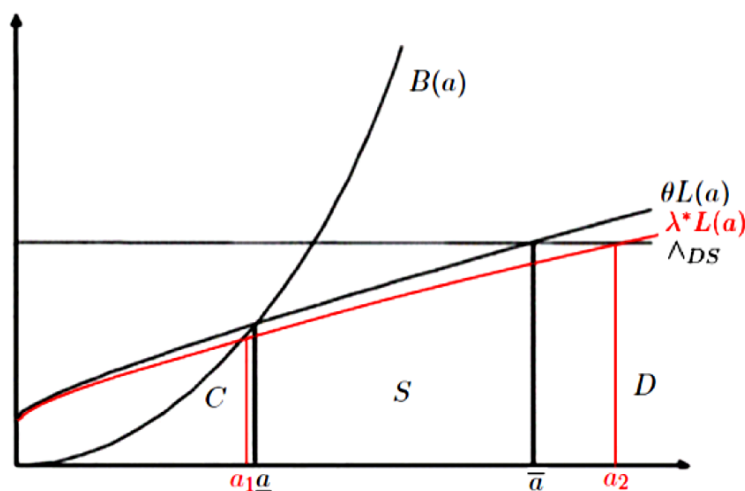
<sup>21</sup>The term, "high and medium level of legalism", follows the typology used in the study of Jo and Namgung (2012) which is associated with a more detailed classification carried out by Smith (2000). Chase et al. (2013) use different names for the level of legalism in RTAs, i.e., quasi-judicial and judicial for medium and high legalism, respectively.

between the stronger and weaker member states in the extended model.

**Assumption 1.**  $\lambda^* < \theta < \lambda$  where  $\theta$  is the common probability of winning in WTO disputes for all members and  $\lambda$  ( $\lambda^*$ ) is a strong (weak) complaining party's probability of winning in RTA disputes.

A complainant's chance of winning affects the expected costs of violation, thereby leading to changes in the players' cost and benefit calculation of strategies. A stronger player considers how much it needs to compensate a weaker player who becomes a sucker if it defects from an agreed tariff. The expected cost of compensation is the loss of its trading partner weighted by the probability the cooperative weak player prevails in the case, i.e.,  $\lambda^* L^*(a)$ .<sup>22</sup>  $\lambda^*$  influences the cut-point of the stronger player's dispute settlement strategy. For a strong player, the smaller chance of losing a dispute in RTA DSMs ( $\lambda^* < \theta$ ) pulls down its lower cutpoint  $a_1$  below that of the WTO  $\underline{a}$  and raises its upper cutpoint  $a_2$  above the WTO's  $\bar{a}$ .

**Figure 2:** Expansion of settlement strategy for a stronger party



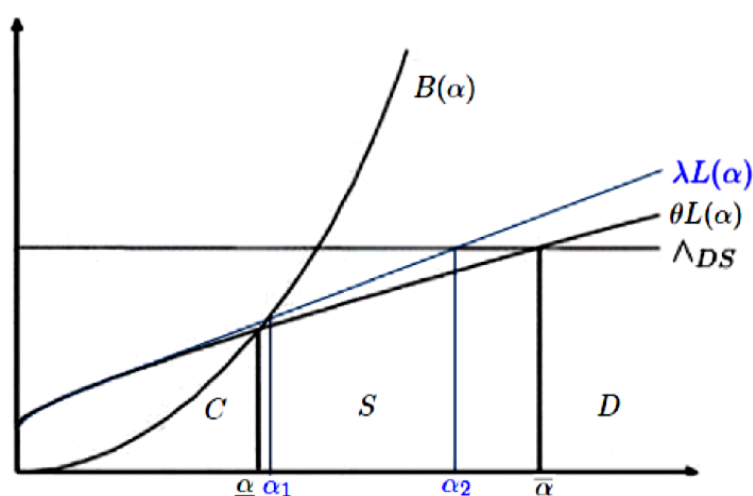
Note: a weaker party is a complainant whose probability to prevail in trial is  $\lambda^*$  s.t.  $\lambda^* < \theta$

Figure 2 represents these changes of smaller  $a_1$  and larger  $a_2$  compared to their counterparts in the WTO model. The changes widen the area of settlement  $S$  for a stronger party's DSS. The enlarged area of  $S$  also means a decrease in the areas of  $C$  and  $D$ , representing cooperation and

<sup>22</sup>By the same logic,  $\lambda L(a)$  is the weak player's expected costs of violation.

defection strategies, respectively. The larger area of temporary suspension of obligations through RTA DSMs represents an increased flexibility of the stronger party. Therefore, it implies that in the RTA trade model a stronger party is more capable of manipulating the trading system. The increased area of manipulation  $S$  results from a decrease in the area of full defection more than full cooperation. In the regional trading system, a stronger player is more likely to stay in the system not by cooperating more but by defecting temporarily.

**Figure 3:** Contraction of settlement strategy for a weaker party



Note: a stronger party is a complainant whose probability to prevail in trial is  $\lambda$  s.t.  $\lambda > \theta$

Then how does the disparity of winning disputes affect a weaker party? For a weaker party to violate, it considers its expected costs of violation which depend on the magnitude of defection and the probability the complaining party wins in the dispute,  $\lambda$ . A weaker party is more likely to lose a case against a stronger party in more power-based RTA DSMs than the WTO DSM,  $\lambda > \theta$ . The increased expected costs of violation from  $\theta L(\alpha)$  in the WTO model to  $\lambda L(\alpha)$  in the RTA model, therefore, pulls up the first cutpoint  $\alpha_1$  above  $\underline{\alpha}$  and drags down the second cutpoint  $\alpha_2$  below  $\bar{\alpha}$ . Such changes narrow down the area between the two cutpoints, the area of settlement strategy  $S$ , which is described in Figure 3. The narrower range of temporary suspension of their obligations represents the contraction of flexibility in DSS which is allocated to weaker states. Weak states become less capable of manipulating the trading system than they were in the



WTO. Also, the loss in the area of  $S$  results from an increased area of full defection  $D$  more than full cooperation  $C$ . Therefore, a weaker player is more likely to violate an agreement without following through RTA DSMs than the WTO DSM, leading to a breakdown of the trading system.

## 4.2 Litigation Costs

Now, let us consider the burdens of litigation costs through which power asymmetries are transferred to the extended regional trade model. Developing countries often fail to initiate litigation due to barriers of resources used for the legal process or concerns about political and economic relationships with other countries, especially developed country partners. Legalization of international trade institutions raises the costs of litigation, which requires participating countries to have greater financial capacity. Therefore, disputants suffer from costs that international dispute settlement procedures incur. Disparities in legal resources affect member countries' abilities to formally participate in litigation of the trade agreements. The amount of costs (e.g. efforts, time, and financial means) is more burdensome to developing countries than to developed countries. Previous studies show that larger and richer countries are more likely to become complainants participating in WTO litigation.

So far, only a small portion of developing countries are involved with legal disputes under international trade regimes as both plaintiffs as well as defendants. Those countries who have participated in DSM under the WTO are Argentina, Brazil, Chile, India, Indonesia, Korea, Mexico, the Philippines, Thailand, and Venezuela. Bown (2005) finds that a state's capacity to pay for legal services, measured by GDP as a proxy, is positively associated with the decision of the state to become a complainant or interested third party. Thus, developing countries have faced more barriers to initiate legal disputes due to a lack of legal resources. Consider the dispute between the U.S. and South Korea over the issue of safeguard, US-Safeguard on Circular Welded Pipe from Korea (DS202). South Africa, Turkey, and Venezuela were also adversely affected by the U.S. safeguard. Nevertheless, only South Korea formally participated in the dispute. Due to a lack of resources, those countries might not be able to bring a case to international arbitration even with external legal assistance provided by the WTO.

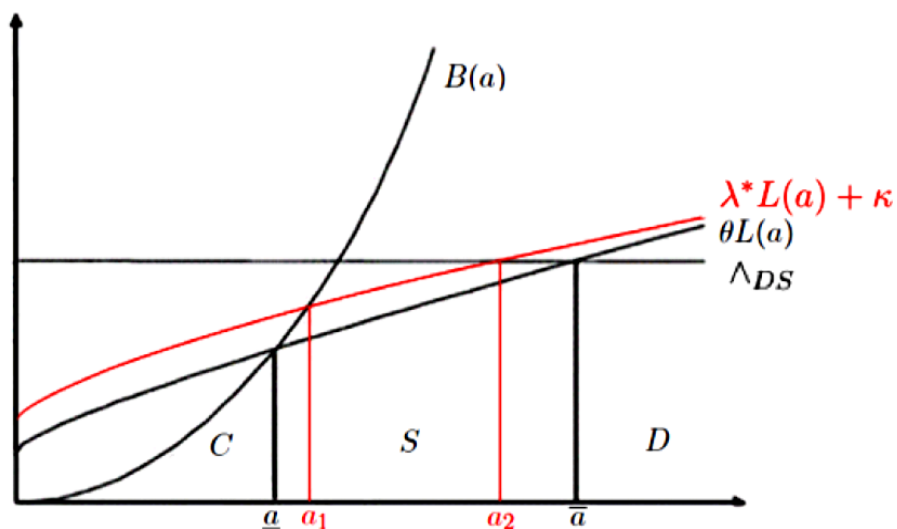
The issue of disproportionality becomes more serious in RTAs that involve a small number of memberships or partnerships with larger power asymmetries. Chase et al. (2013) show that the logic of economies of scale explains why the more multilateral an international agreement is the more likely it is to be institutionalized. Additionally, most RTAs do not provide member countries with access to professional legal assistance. In comparison, WTO members gain access to legal support from trained experts through the WTO secretariat, prior to and during the dispute settlement process. The existence of such an external support could alleviate disparities in legal resources between strong and weak states. In addition to the absence of legal assistance, many RTAs do not have a standing body, rather, they call an ad-hoc panel if necessary. Formation of an ad-hoc panel requires a more substantial amount of time and financial resources than using a standing body.<sup>23</sup>

Another issue of litigation costs in RTAs lies in the fact that in most cases the DSM costs of proceedings are shared equally by both disputants. The equal cost-sharing yields another layer of disparity in the financial burden between the strong and the weak states. This implies that a weak state even with a very strong case is afraid of paying high litigation costs without getting compensated in the event of losing the dispute. Taken all together, the design of RTA DSMs entails disparity in the burdens of its signatories attributable to the disparate allocations of financial and legal resources. The burden of legal expenses in the DSM process for a member country is denoted by  $\kappa > 0$ . Let  $\kappa$  denote the burden of litigation costs for the strong and  $\kappa^*$  for the weak. Given that the cost burden is inversely proportional to a state's power, we assume the following relationship:  $\kappa^* > \kappa$ .

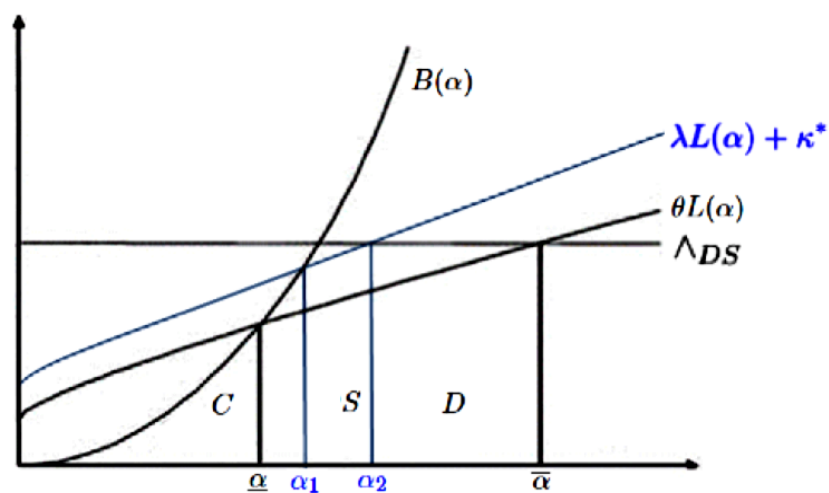
For a stronger party, those additional costs in the extended model pull up the lower bound where a player is indifferent between  $C$  and  $S$ , while pulling down the upper bound where a player is indifferent between  $S$  and  $D$ . Such changes widens both areas of  $C$  and  $D$  at the expense of the region for  $S$ , which is described in Figure 4. For a weaker party,  $\kappa^*$  narrows down the area

---

<sup>23</sup>The WTO also has the standing Appellate Body and its operating fund.

**Figure 4:** Contraction of settlement strategy

Stronger party with litigation cost  $\kappa > 0$

**Figure 5:** Greater contraction of settlement strategy

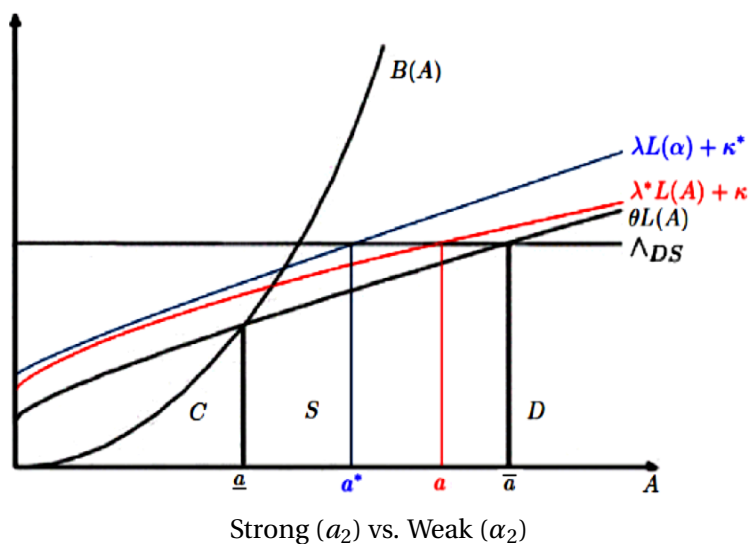
Weaker party with litigation cost  $\kappa^* > \kappa$

of settlement by dragging down the lower bound  $\alpha_1$  and raising the upper bound  $\alpha_2$  just as  $\kappa$  plays on the stronger party's DSS. For a weaker party, the costs burden results in a more vivid outcome, the contraction of a settlement strategy due to the shrinking effect of a large  $\lambda$  (Figure 5). Taken together, the second proposition summarizes the first central result that RTA DSMs reduce the flexibility of DSS for both players and more so for weak players.

**Proposition 2** (Flexibility). *RTAs are less flexible than the WTO. All members face contraction of the areas of  $S$ . The area of  $S$  for weaker members is narrower than that of stronger members.*

Differential costs  $\kappa$  and  $\kappa^*$  narrow down the area of  $S$  for the DSS of both stronger and weaker parties, respectively. The magnitudes of contraction for each player by which their litigation costs vary across the probability of winning a trade dispute. Recall that in more power-based RTAs a strong state is less likely to lose a case against a weak state. In such a condition, the downsizing effect of  $\kappa$  for a strong state is smaller than a weak state. On the contrary, a weaker state is more likely to lose a case in RTAs, which exacerbates contraction of the  $S$  region on its DSS.

**Figure 6:** Loss of stability between the WTO and RTAs



The comparison between Figure 4 and Figure 5 shows that the steeper slope  $\lambda$  of a weaker party's expected violation costs exacerbates the contraction of the  $S$  area by  $\kappa^*$ , while the lower slope  $\lambda^*$  alleviates such an effect for a stronger party. Therefore, a weaker state ends up with a much narrower range of  $S$  in its DSS than that of a stronger state. A decrease in the area of  $S$  mostly comes from an increase in the area of  $D$ . As Figure 6 describes, the defective areas for DSS of both players are expanded due to the lower level of upper cutpoints  $a_2$  and  $\alpha_2$  of the strong

and the weak player, respectively, than that of WTO  $\hat{a}$ . The disparity in the reduction of  $D$  exists across players, too.

Following the inequality  $a_2 > \alpha_2$ , a stronger player is less likely to defect without compliance than a weaker counterpart. Playing  $D$  comes with the breakdown of the trading system because the player leaves the system if it loses. Therefore, RTA DSMs raise the level of the per period probability of system breakdown: RTAs are less stable than the WTO. The next proposition establishes the second central result of this paper: RTA DSMs suffer from a reduced stability of the system than does the WTO DSM.

**Proposition 3 (Stability).** *RTAs are less stable than the WTO. All members face an increased per period probability of breakdown which is greater for weaker members relative to stronger members.*

## 5 The Case of NAFTA

Since the North American Free Trade Agreement (NAFTA) entered into force on January 1st, 1994, the FTA has listed more than 200 disputes. However, its success is far from clear given that most cases brought to the NAFTA DSM are investor-state disputes. Most cases are disputes between investors and states regarding Chapter 11 “Settlement of Disputes between a Party and an Investor of Another Party” or Chapter 19 “Review of Final Antidumping and Countervailing Duty Determinations.” The dispute settlement provisions of Chapter 20 “Institutional Arrangements and Dispute Settlement Procedures” are applicable to disputes among governments regarding the interpretation or application of NAFTA. There are only three state-to-state disputes regarding Chapter 20. The WTO to date has seventeen disputes among NAFTA members which proceeded to panel process. Compared with a total of three cases under NAFTA, it shows that the NAFTA DSM is less frequently used than the WTO DSM.

Out of the three cases under NAFTA, one case was brought by the U.S. against Canada.<sup>24</sup> However, the U.S. never initiated a dispute against Mexico under NAFTA; the U.S. brought a

---

<sup>24</sup>In 1996, the U.S. request consultations with Canada in the matter of “Tariffs Applied by Canada to Certain U.S. Origin Agricultural Products”.

total of six cases against Mexico to the WTO instead.<sup>25</sup> In comparison, in 1998 and 2001 Mexico requested consultations with the U.S. under the NAFTA DSM: “The U.S. Safeguard Action Taken on Broom Corn Brooms from Mexico” and “Cross-Border Mexican-owned Trucking Services” where Mexico prevailed. Despite a panel decision in favor of Mexico, Mexico brought eight cases against the U.S. to the WTO DSM since its experience with the NAFTA DSM.<sup>26</sup>

One might ask since RTA DSMs provide more flexibility to stronger countries at the expense of weaker countries, stronger countries are more likely to choose RTAs rather than the WTO as a forum for dispute resolution. Results of the model claim that weaker countries prefer the WTO to RTAs more than stronger countries, due to the disadvantages in flexibility and stability. The results, however, also point out that both strong and weak countries suffer from losses of strategic flexibility and systemic stability. It follows that if stronger countries fear the breakdown of the trading system, they would like to use the WTO DSM rather than RTA DSMs.

Previous studies suggest other reasons for the frequent use of the WTO DSM. For example, Mexico initiated a case against the U.S. in the WTO in 2001. It was the second case brought by Mexico against the United States. Mexico was not alone in this case. Canada was the other complainant and there were thirteen third-party participants: Argentina, Australia, Brazil, Costa Rica, European Communities, Hong Kong, India, Indonesia, Israel, Japan, Korea, Norway, and Thailand. Given that Mexico waited for nearly a decade before it imposed retaliatory tariffs in March 2009 for the NAFTA trucking dispute, this large number of third parties and the unsuccessful story of enforcement through NAFTA implies that enforcement power plays a significant role in the decision of forum for dispute resolution. Recall that Mexico brought eight other cases against the U.S. to the WTO and no case to NAFTA since 2001.

In addition to the losses in flexibility and stability in an absolute sense, strong countries

---

<sup>25</sup>The first case initiated in 1997 by the U.S. against Mexico: “Anti-Dumping Investigation of High-Fructose Corn Syrup (HFCS) from the United States” (DS101). Four of them came through the panel process.

<sup>26</sup>There are a total nine cases brought by Mexico to the WTO. The first case initiated by Mexico in 1996, “Anti-Dumping Investigation Regarding Imports of Fresh or Chilled Tomatoes from Mexico” (DS49), did not proceed to the panel process. Six out of the total nine cases went through the panel process.

have an incentive to use the WTO DSM over RTA DSMs. Studies also found that stronger states have other considerations in their choice of DSM. For example, Busch (2007) argues that RTAs allow “forum shopping” for member states who hold WTO membership; the complainants could submit the dispute to the DSM in regional or multilateral trade agreements, if they want litigation. His study of *Canada-Periodicals* and *Mexico-corn brooms* shows that the strategic choice of forum comes from the consideration of future usage of the ruling rather than the chance of winning in the current term.

Previous studies also consider different levels of enforcement and the costs of exit between the WTO and RTAs in deciding the forum for dispute resolution. From the perspective of a complainant who chooses the forum, less powerful players prefer the WTO due to its greater enforcement powers and higher exit costs compared to that of RTAs, especially against powerful opponents. These attributes help level the playing field in the WTO for less powerful players. Powerful players with greater retaliatory power against less powerful players also prefer the WTO because of not only greater enforcement from the system but also high exit costs which help prevent less powerful players from withdrawing from the system. In spite of the advantages of prevailing in cases under RTA DSMs, powerful players also have incentive to bring cases to the WTO because they consider the future usage of the ruling when deciding which forum to use. First, a case brought to the WTO is highly likely to be related to other potential disputes with other member countries. Second, a ruling is considered to be more legitimate due to its higher level of legalism as compared to RTAs.

In sum, the results of the extended model show that a weak state prefers the WTO to RTAs when it comes to a dispute against a powerful state, due to its losses in flexibility and stability. In addition, greater enforcement powers and high exit costs help less powerful members insulate themselves from power politics in the WTO more efficiently than RTAs. This implies that powerful states also seek the WTO DSM over RTA DSMs for cases against weak states. Despite advantages in the regional setting, powerful states lean towards the WTO DSM in fear of the breakdown of

the trading system. Given that powerful players consider the future usage of the ruling rather than the chance of winning in the current period, these powerful states prefer the WTO DSM to RTA DSMs because rulings under the WTO are considered to be more legitimate and related to other potential disputes. Taken together, a dispute between the strong and the weak in particular, is more likely to be brought to the WTO than RTAs.

## 6 Conclusion

The development of international trade cooperation alongside the consistent rise of global economic interdependence has driven the growth of multilateralism in the WTO. Nevertheless, the tradition of multilateral trade negotiation has more recently been replaced by RTAs since the early 1990s. To understand the dynamics of regional trade agreements and the implications for the world trade system, this paper analyzes a structural (dis)advantage of regional trade agreements to large and small economies by comparing the role of DSMs between regional and multilateral trading systems.

The extension of Rosendorff's game-theoretic model from the WTO to RTAs demonstrates that DSMs found in RTAs reinforce inequalities in the international trade system and increase systemic instability. The mechanisms of dispute resolution under RTAs grant more room for manipulating the trade agreement to stronger states. Nevertheless, weak states have less flexibility in RTA DSMs than they would have in the WTO DSM. Therefore, the WTO DSM is more efficient in leveling the playing field relative to RTA DSMs.

The strongly legalized dispute settlement mechanism increases the level of stability of the WTO by bringing flexibility into the strategies of member countries (Rosendorff, 2005). In RTAs, however, DSMs yield disparity in the extent of strategic flexibility allowed to asymmetric partners. This issue of inequality makes such positive effect of DSMs as in the WTO uncertain by reducing systemic stability of the regional trade agreements. Additionally, states face the loss of strategic flexibility regardless of power in RTA DSMs compared to the WTO DSM. This also indicates that RTAs are less stable than the WTO given that all members face the increased per



period probability of breakdown, which suggests that the superiority of the WTO DSM provides a compelling explanation for the more frequent use of the WTO DSM than RTA DSMs.

The results imply that if there is a high chance of political pressure for trade protection and consequent high odds of defection, small and weak countries should be more careful about joining RTAs, especially with larger and stronger partners. Given that legal capacity is one of the major obstacles for poor countries to use dispute settlement mechanisms, access to external legal support available for those countries in RTAs will improve the imbalance and the disuse of dispute settlement mechanisms in RTAs. The model demonstrates that small and poor countries face increased inequality in dispute settlement under RTAs than WTOs. Despite the apparent disadvantages, small and poor countries might be interested in joining RTAs in the fear of extremely high costs of being an innocent bystander in a natural trading region unless the WTO provides them with more incentives to lower the burdens of membership and increase gains from cooperation in the multilateral trading system.

## 7 Appendix

**Lemma 1.** Define  $a^*$  such that  $B(a^*) = \wedge_{DS}$ . Then  $\underline{a} \leq a^*$ .

**Proof.**  $B(a^*) = \wedge_{DS}$ ; then  $\theta L(\bar{a}) = \wedge_{DS} = B(a^*)$ . Since  $\bar{a} > \underline{a}$  by **Lemma 3** in Rosendorff (2005),  $\theta L(\bar{a}) > \theta L(\underline{a})$ . Therefore,  $a^* \geq \underline{a}$ . ■

**Lemma 2.** Define  $f(a) = B(a) - \theta L(a)$ . Then  $\bar{a} \geq a^*$ .

**Proof.**  $f(a) = B(a) - \theta L(a)$ ; then  $f''(a) = B''(a) - \theta L''(a) > 0$ . And  $f(\underline{a}) = B(\underline{a}) - \theta L(\underline{a}) = 0$ . Let  $a_1$  and  $a_2$  be  $a_2 > a_1 \geq \underline{a}$ . Then,  $\frac{f(a_2) - f(a_1)}{a_2 - a_1} > \frac{f(a) - f(0)}{a - 0} = \frac{\theta L(0)}{a} \geq 0$ .  $f(a)$  is convex because  $f(a_1) < f(a_2)$  and  $f''(a) > 0$ . Hence,  $f(a)$  is increasing over the range of  $a \geq \underline{a}$ . Now,  $f(a^*) = B(a^*) - \theta L(a^*) \geq f(\underline{a}) = 0$ ; then  $\wedge_{DS} \geq \theta L(a^*)$ .  $\theta L(\bar{a}) = \wedge_{DS} \geq \theta L(a^*)$ . Therefore,  $\bar{a} \geq a^*$ . ■

**Assumption.**  $\lambda^* \leq \frac{\wedge_{DS} - \kappa}{L(a^*)}$  from Rosendorff's assumption on  $\theta \leq \min\{\frac{\wedge_{DS}}{L(\bar{a})}, \frac{\wedge_{DS}}{L(\underline{a})}\}$ .

Note that  $\wedge_{DS} - \kappa > 0$ ; also  $\wedge_{DS} \leq \kappa + L(\bar{a})$ . The maximum level of cost to stay in a treaty is the cost of the dispute procedure with the maximum level of compensation. These costs for a state not to leave the treaty should be greater than the difference of the expected benefit from playing the “cooperate” strategy, including  $C$  and  $S$ , and that of defecting once with the entire exit. If the maximum costs of being in the treaty are smaller than the gain of the state from being a cooperative player, the state will never leave the treaty because it gains more when it cooperates and will even pay the maximum amount of penalty for its temporary defection than choose to defect once followed by exit.

Now, we divide our problem into two cases: (1)  $\frac{\wedge_{DS} - \kappa}{L(\bar{a})} \leq \lambda^* \leq \frac{\wedge_{DS} - \kappa}{L(a^*)}$ , and (2)  $\lambda^* \leq \frac{\wedge_{DS} - \kappa}{L(\bar{a})} \leq \frac{\wedge_{DS} - \kappa}{L(a^*)}$ .

Note that  $\wedge_{DS} - \kappa > 0$ . From the assumption  $\wedge_{DS} \leq \kappa + L(\bar{a})$ ; then  $\frac{\wedge_{DS} - \kappa}{L(\bar{a})} \leq 1$ . Also note that  $L(a^*) < L(\bar{a})$  by Lemma 2 which shows that  $a^* \leq \bar{a}$ .

**Lemma 3.**  $\lambda^* L(a^*) + \kappa \leq \wedge_{DS} \leq \lambda^* L(\bar{a}) + \kappa$ .

**Proof.** In the first case of  $\frac{\wedge_{DS} - \kappa}{L(\bar{a})} \leq \lambda^* \leq \frac{\wedge_{DS} - \kappa}{L(a^*)}$ . ■

**Lemma 4.**  $\lambda^* L(\bar{a}) + \kappa \leq \wedge_{DS}$  and  $\lambda^* L(a^*) + \kappa \leq \wedge_{DS}$ .

**Proof.** In the second case of  $\lambda^* \leq \frac{\wedge_{DS} - \kappa}{L(\bar{a})} \leq \frac{\wedge_{DS} - \kappa}{L(a^*)}$ . ■

**Assumption.**  $\lambda \leq \frac{\wedge_{DS} - \kappa^*}{L(a^*)}$ . Note that  $\wedge_{DS} - \kappa^* > 0$  from Rosendorff's assumption on  $\theta \leq \min\{\frac{\wedge_{DS}}{L(\bar{a})}, \frac{\wedge_{DS}}{L(\underline{a})}\}$ .

**Lemma 5.**  $\lambda L(a^*) + \kappa^* \leq B(a^*) = \wedge_{DS}$ .

**Proof of Proposition 1.** Given that the weaker party is playing a DSS, we must show that playing the DSS satisfies the no-defect condition for the stronger party. Given the current period draw  $\hat{a}$ , the expected current period return from defection for the stronger is  $D(\hat{a}) - \kappa$ , and hence the gain from defection for the stronger are  $D(\hat{a}) - \kappa - \max(C(\hat{a}), D(\hat{a}) - \lambda^* L(\hat{a}) - \kappa) = \min(B(\hat{a}) - \kappa, \lambda^* L(\hat{a})) = \min(B(\hat{a}), \lambda^* L(\hat{a}) + \kappa)$ . ■

Consider the event in which a deviation has been observed in some period. From then, the one-shot Nash strategies are played, yielding the Nash payoff (in expectation, because the draws in the future periods are unknown) forever. That is, the aggregate Nash is payoff  $V_D = \frac{1}{1-\delta}(N - \kappa)$ . What is the foregone cooperative aggregate payoff? If cooperation occurred in the last period, in the next each player has the option of cooperating again, or defection. Then the value of the game in a cooperative phase is the earning from the play in that period, plus the continuation value: for a stronger party,

$$V = p [ p (C + \delta V) + (1 - p)(S + \lambda L - \kappa + \delta V)] \\ + (1 - p)[ p (D - \lambda^* L - \kappa + \delta V) + (1 - p)(N + \lambda L - \lambda^* L - \kappa + \delta V)]$$

Solving, we have  $V = \frac{1}{1-\delta} [ p^2 (C - S - D + N) + p(D - 2N + S) + N - \kappa + (1 - p)(\lambda L - \lambda^* L) + p^2 \kappa]$ . Hence,  $V - V_D = \frac{1}{1-\delta} [ p^2 (C - S - D + N) + p(D - 2N + S) + (1 - p)(\lambda - \lambda^*)L + p^2 \kappa]$ . The no-defect condition in any period after  $\hat{a}$  is observed (and punishment starts in the next period) is  $\min(B(\hat{a}), \lambda^* L(\hat{a}) + \kappa) < \frac{\delta}{1-\delta} [ p^2 (C - S - D + N) + p(D - 2N + S) + (1 - p)(\lambda - \lambda^*)L + p^2 \kappa]$  or

$\hat{a} < a_2$ .

If  $\hat{a} < a_1 < a_2$ , then  $B(\hat{a}) < \lambda^*L(\hat{a}) + \kappa$  and the benefits of defection are too small to make either pure defection or use of the dispute settlement mechanism worthwhile: if  $a_1 < \hat{a} < a_2$ , the benefits of the DSM outweigh pure cooperation, but it is still intertemporally optimal to voluntarily pay the proportionality penalty to benefit from the possibility of cooperation in the next period. The no-defect condition is violated when  $\hat{a} > a_2$ ; then the gains from pure defection, and the Nash reversion play from then on are preferred to cooperation. Hence, a pair of DSSs is an equilibrium. ■

### **Proof of Proposition 2.**

We need to show  $\bar{a}_w - \underline{a}_w < \bar{a}_s - \underline{a}_s$  where  $\bar{a}_w = \alpha_2$ ,  $\underline{a}_w = \alpha_1$ ,  $\bar{a}_s = a_2$ , and  $\underline{a}_s = a_1$ . The following proof of Proposition 3 shows that  $\bar{a}_w < \bar{a}_s$ . Given  $B' > \wedge_{DS}'$ ,  $\underline{a}_s - \underline{a}_w < \bar{a}_s - \bar{a}_w$ .

### **Proof of Proposition 3.**

We need to show  $\bar{a} - a_s \leq \bar{a} - a_w$  where  $a_s = a_2$ , and  $a_w = \alpha_2$ .  $\lambda L(a_w) + \kappa^* = \wedge_{DS} = \lambda^* L(a_s) + \kappa$ . From the assumption  $\kappa^* > \kappa$ ,  $\lambda L(a_w) \leq \lambda^* L(a_s)$ ; then  $L(a_w) \leq L(a_s)$  by the assumption  $\lambda > \lambda^*$ . Now  $L' > 0$ , so  $a_w \leq a_s$ . ■

## References

- Aghion, P., P. Antràs, and E. Helpman (2007). Negotiating Free Trade. *Journal of International Economics* 73(1), 1–30.
- Bhagwati, J. and A. Krueger (1995). *The Dangerous Drift to Preferential Trade Agreements*. American Enterprise Institute for Public Policy Research.
- Bown, C. P. (2005). Participation in WTO Dispute Settlement: Complainants, Interested Parties, and Free Riders. *The World Bank Economic Review* 19(2), 287–310.
- Busch, M. L. (2007). Overlapping Institutions, Forum Shopping, and Dispute Settlement in International Trade. *International Organization* 61(4), 735–761.
- Busch, M. L. and E. Reinhardt (2003). Developing Countries and General Agreement on Tariffs and Trade/World Trade Organization Dispute Settlement. *Journal of World Trade* 37(4), 719–735.
- Chase, C., A. Yanovich, J.-A. Crawford, P. Ugaz, and WTO (2013). Mapping of Dispute Settlement Mechanisms in Regional Trade Agreements? Innovative or Variations on a Theme? *WTO Staff Working Paper-Economic Research and Statistics Division* 7, 1–58.
- Chaudoin, S. (2010). Information Transmission and the Strategic Timing of Trade Disputes. Presented at the 2010 International Political Economy Society.
- Davey, W. J. (2006). Dispute Settlement in the WTO and RTAs: a Comment. In L. Bartels and F. Ortino (Eds.), *Regional Trade Agreements and the WTO Legal System*, Chapter 14, pp. 343–357. Oxford University Press.
- Fang, S. (2008). The Informational Role of International Institutions and Domestic Politics. *American Journal of Political Science* 52(2), 304–321.
- Fearon, J. D. (1998). Bargaining, Enforcement, and International Cooperation. *International Organization* 52(2), 269–305.
- Goldstein, J., M. Kahler, R. Keohane, and A.-M. Slaughter (2000). Introduction: Legalization and World Politics. *International Organization* 54(3), 385–399.
- Hillman, J. (2009). Conflicts Between Dispute Settlement Mechanisms in Regional Trade Agreements and the WTO? What Should the WTO Do? *Cornell International Law Journal* 42(2), 193–208.

- Jo, H. and H. Namgung (2012). Dispute Settlement Mechanisms in Preferential Trade Agreements: Democracy, Boilerplates, and the Multilateral Trade Regime. *Journal of Conflict Resolution* 56(6), 1041–1068.
- Keohane, R. O., A. Moravcsik, and A.-M. Slaughter (2000). Legalized Dispute Resolution: Interstate and Transnational. *International Organization* 54, 457–488.
- Kucik, J. and K. J. Pelc (2016). Measuring the cost of privacy: A look at the distributional effects of private bargaining. *British Journal of Political Science* 46(4), 861–889.
- Ludema, R. D. (2001). Optimal International Trade Agreements and Dispute Settlement Procedures. *European Journal of Political Economy* 17(2), 355–376.
- Maggi, G. (1999). The Role of Multilateral Institutions in International Trade Cooperation. *American Economic Review* 89(1), 190–214.
- Marceau, G. and J. Wyatt (2010). Dispute Settlement Regimes Intermingled: Regional Trade Agreements and the WTO. *Journal of International Dispute Settlement* 1(1), 67–95.
- Park, J.-H. (2000). International Trade Agreements between Countries of Asymmetric Size. *Journal of International Economics* 50(2), 473–495.
- Riezman, R. (1985). Customs Unions and The Core. *Journal of International Economics* 19, 355–366.
- Rosendorff, B. P. (2005). Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure. *American Political Science Review* 99(3), 389–400.
- Rosendorff, B. P. and H. V. Milner (2001). The Optimal Design of International Trade Institutions: Uncertainty and Escape. *International Organization* 55(4), 829–857.
- Smith, J. M. (2000, 12). The Politics of Dispute Settlement Design: Explaining Legalism in Regional Trade Pacts. *International Organization* 54(1), 137–180.