A Theory of Trade Disagreement

Alan C. Spearot *
University of California - Santa Cruz

December 9, 2010

Abstract

This paper examines the scope for agreement in trade negotiations. By adopting a cooperative solution concept based on the core, and using a simple model of intraindustry trade between countries and regions, I prove that there exists a large range of relevant parameter values such that countries cannot agree on a mutually beneficial trading arrangement. This no-agreement outcome is more likely when cost differences between regions are high. These results highlight the critical role of transfers in facilitating a mutually beneficial trading equilibrium. Further, I show that political influence by firms has the effect of providing a focal point for negotiations based on producer market power, and thus, increases the scope for agreement in trade negotiations.

JEL Classifications: F13, F15
Keywords: Trade Agreements, Coalition Formation, Political Economy

1 Introduction

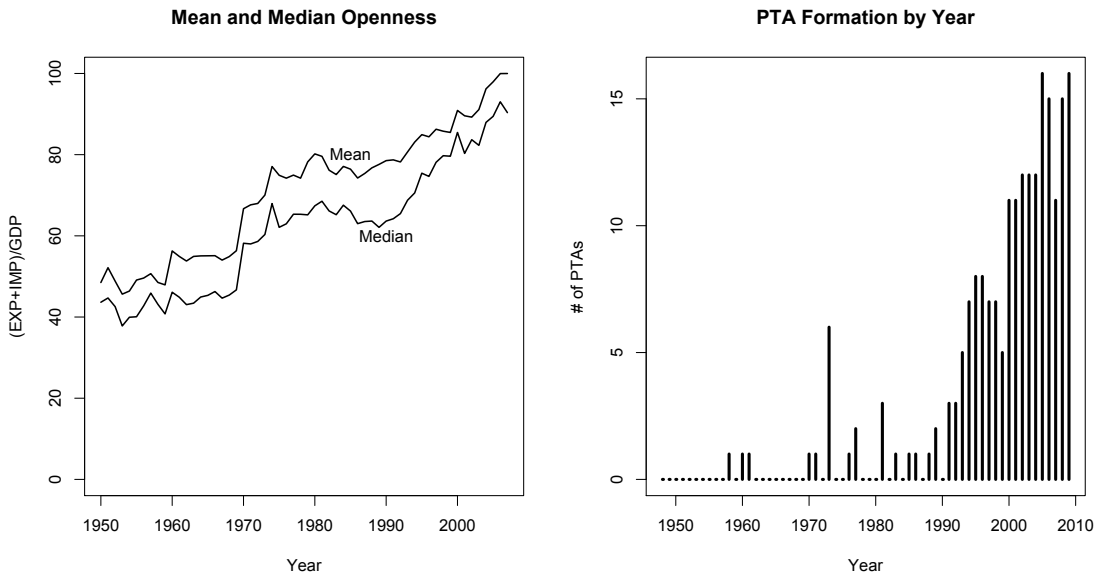
The debate over the optimal way to liberalize trade has a long and contentious history. In particular, the exists no clear preference among nations over a patchwork of regional agreements or a larger multilateral framework. However, despite ongoing academic and policy discussions, and unlike stalled multilateral negotiations, the recent growth in regional agreements has been astounding. The right panel of Figure 1 illustrates the growth in such agreements.[1] Clearly, the number of regional agreements reported to the GATT/WTO has exploded since the late 90s. Further, this expansion of regionalism is notable even in contrast with the steady growth of world openness over the same period.

While there is little argument over the practical significance of these regional agreements, the welfare consequences are less certain. In the trade literature, the key question is whether regional agreements form "building blocks" for larger multilateral agreements, or serve as "stumbling blocks"

*Email: acspearot@gmail.com. Address: Economics Department, 1156 High Street, Santa Cruz, CA, 95064. Tel.: +1 831 419 2813. The author would like to thank John Turner and Meredith Crowley for detailed comments, and participants at the 2008 Asia Pacific Trade Seminar and the 2008 Southern Economics Association Annual meeting. All errors are my own.

[1]The agreement data used to produce this graph was obtained from the WTO website at http://www.wto.org/english/tratop_e/region_e/region_e.htm. Openness data is calculated from the Penn World Tables

Figure 1: Openness and PTAs: 1950 to Present



which hinder the overall process of multilateral liberalization (Bhagwati, 1990).[2] Implicit in this question is a notion of equilibrium agreement formation. Until recently, however, a significant portion of the trade policy literature has evaluated the effects of trade agreements absent a full equilibrium model of the structure of global trade agreements.

In this paper, I use a framework of cooperative coalition formation to model the equilibrium structure of trade agreements. The use of cooperative coalitions, which relies on identifying the *core* of the agreement formation game, is not new in its own right (as detailed below). However, I make a number of modifications to the structure of the typical setup that yield insights regarding when countries can agree on a unique trading equilibrium, when they cannot, and what factors influence the scope for agreement in trade negotiations. Specifically, I apply the cooperative solution concept to a four-country, two-region model of intraindustry trade in the style of Brander and Krugman (1983). Assuming four countries rather than three is itself a departure from the majority of the existing literature. However, it is a critical departure in that it allows for increased flexibility in how countries respond to other agreements. In particular, in a three country model, if two countries decide to form an agreement, the third country is left unable to adjust trade partnerships in any meaningful way (aside from joining the FTA). This is remedied by adding a fourth country, with whom the third country can join to balance the regional agreement between countries one and two.

Further, the assumption of two regions allows for a fairly tractable treatment of regional cost

[2]The classic literature (Viner, 1950; Meade 1955) identifies trade creation and trade diversion as critical components of preferential agreements and their welfare effects. A preferential agreement could be welfare enhancing for all parties inside and outside of the agreement if trade from outsiders increases after the agreement is put in place. This is likely if the agreement reinforces traditional patterns of comparative advantage. Freund (2000), and Saggi (2006) are recent papers addressing the building or stumbling block nature of preferential liberalization. The former shows in a repeated game framework that multilateral tariff reduction facilitates further preferential reductions. The latter, also using a repeated game framework, shows that preferential agreements hinder multilateral cooperation when countries are symmetric. However, asymmetry can reverse this prediction.

differences and distance between regions, both being the key parameters over which I evaluate agreement formation. Two regions are separated by a per-unit trade cost, where each region has two countries which border one another. One region is labeled as the "North" with marginal costs of production weakly greater than the other region, the "South". Each country in each region may set a specific MFN tariff, but may also agree to free trade with some coalition of other countries. These coalitions are restricted to be one of four outcomes: (1) global free trade, (2) two intraregional agreements, (3) two extraregional "North-South" agreements, or (4) no agreements (unconstrained MFN).

Despite being fairly stylized, the results from the model are simple and provocative. For a majority of the parameter space (regional cost differences and trade costs) such that trade occurs between all countries, no equilibrium outcome is uniquely dominant for all countries. That is, countries cannot agree on a unique welfare maximizing equilibrium coalition. While absent political influence it is always beneficial to discard unconstrained MFN, regions may disagree whether it is best to do so via global free trade or intraregional agreements. In particular, countries tend to disagree over the first-best agreement equilibrium when cost differences are high. The intuition is that as cost differences between regions increase, northern countries wish to cooperate within a regional agreement to trade *some* market access, but also to act as a buffer from the more efficient southern region. Of course, the incentives of southern countries are exactly the opposite, as they can gain substantially from unrestricted market access to countries with less-competitive firms. Eventually, as cost-differences become large, this dichotomy becomes prohibitive, and there exist no parameter values such that the core of the agreements formation game consists of a unique trade agreement structure. In contrast, when cost differences are low, global free-trade is optimal when trade costs are relatively low, and regional agreements are optimal when trade costs are relatively high.[3]

The results from this particular cooperative framework are both enlightening and somewhat troubling. Under the baseline model, absent political influence, non-cooperative MFN is always dominated by *some* free trade agreement. However, for a majority of the relevant parameter space, countries cannot agree on the exact nature of this agreement. This suggests that either timing (as was the focus in Aghion, Antras, and Helpman, 2007), other policy instruments, or political influence may drive the formation of free trade areas. Indeed, I show that intra-agreement transfers partially solve the problem of trade disagreement, though not completely. Inter-agreement transfers completely solve the problem of trade disagreement, though are required for a majority of the relevant parameter space of production costs and trade costs.[4] This suggests that models of trade policy that measure welfare globally may be reliant on such transfers to reach global efficiency when countries are sufficiently asymmetric.[5] Regarding political influence, I also examine a restricted model in which producer profits are the only component of a country's welfare function (similar to Krishna, 1998). Within this stark political model, I show that for roughly half of the parameter space - more than with the standard welfare function - countries agree to unconstrained MFN

[3]Frankel, Stein, and Wei (1995) and Wonnacott and Wonnacott (1981) have also argued trade costs play an important role in the formation of trade agreements. Panagariya (2000) disagrees, stating "trade costs are not special". In theoretical work, Panagariya (1998) justifies this statement by showing that trade costs have no bearing on the equilibrium outcome when comparative advantage dictates that you trade with your extraregional partners. Indeed, given the assumption of intraindustry trade, my results are similar to Frankel, Stein, and Wei (1995).

[4]The particular approach I use allows me to analyze transfers in trade agreements beyond the original work of Riezman (1985) in characterizing the space of parameters such that transfers are required.

[5]One example is Bagwell and Staiger (2003). See section seven for a discussion.

(agreeing to have no agreement). Thus, political influence seems to enhance the possibility of a unique core by providing a focal point based on producer market power.

Related Literature

There are three distinct approaches to modeling the formation of trade agreements: extensive form negotiation, non-cooperative coalitions, and cooperative coalitions. To my knowledge, the lone example of the first is Aghion, Antras, and Helpman (2007), which examines how a lead country (the US, for example) chooses between negotiating regional agreements sequentially, or negotiating multilateral free trade all at once. In particular, they examine how externalities within preferential agreements dictate whether preferential agreements serve as "building blocks" or "stumbling blocks" in a move toward global free trade. Under certain conditions, especially when political pressures are significant, preferential agreements serve as stumbling blocks, and should be prohibited. In other cases, global free trade will not occur without using preferential agreements as building blocks.

A second approach is to model the formation of agreements as a one-shot game of non-cooperative coalition formation, where equilibrium agreements are determined by a Nash-style simultaneous announcement game. Saggi and Yildiz (2007) model agreement formation as such, where they derive non-cooperative coalition proof equilibria for a three country model. Under the assumption of symmetry, they show that free trade agreements produce a weak stumbling block effect. In contrast, free trade agreements can produce a strong building block effect when countries are asymmetric.

As in this paper, the third approach uses cooperative coalitions to model agreement formation. Earlier work using cooperative models includes Riezman (1985), Kennan and Riezman (1990), and later, Melatos and Woodland (2007). The cooperative approach is similar to the framework in Saggi and Yildiz (2007) in that it assumes no specific extensive form representation of the agreement formation game. However, it differs in how potential agreement equilibria are discarded. In a model of cooperative coalitions, a potential equilibrium dominates another if, for all countries whose agreement status changes between the two, each country is made better off. In equilibrium, any trade agreements which have not been dominated will comprise the *core* trading equilibria. When the core is comprised of multiple trade agreement structures, I conclude that for the given parameter values, we have *trade disagreement*. Generally, given its identification of the core of the trade agreement formation game, the cooperative framework provides a clear sense of the scope for agreement in trade negotiations as a function of model preliminaries.[6]

Outline

The paper is organized as follows. Section two presents a simple two country model that focuses on the effects of efficiency differences across countries, and the likelihood of a unique trade agreement in the core of the agreement formation game. In section three, I present the basics of a model with four countries organized into two regions. In sections four through six, I solve and apply this model to a variety of different settings of production cost heterogeneity and trade costs. Section seven discusses the stumbling block - building block debate, the role of transfers, and how political influence can provide a focal point for negotiations. In section eight, I briefly conclude.

[6]Other similar papers in this literature are Burbridge, DePater, Myers, and Sengupta (1997), who address tax-policy coordination, showing that with many states, trading blocs form in equilibrium. Furusawa and Koneshi (2007) take a slightly different approach, modeling agreements as a network formation game. In their work, they show that unless countries are symmetric, a global free-trade outcome is not stable.

2 A two country model

Before introducing the four country model described in the introduction, I will first discuss the basic intuition of the agreement framework within a simplified two country model. The timing of the model consists of three stages. In the first stage, countries X and Y decide to either form a free trade agreement (F), or remain in an environment of noncooperative tariff setting (M). In the second stage, if no agreement has been reached, countries choose specific tariffs. In the third and final stage, firms make output decisions for each market. I introduce further specifics in reverse order.

The consumer-producer setup is very similar to Brander and Krugman (1983). Consumers in each country demand one homogeneous good, where inverse demand is assumed to take the form $P = A - bQ$. I assume that each country has only one producer. The only difference between countries is the cost of production, where the firm in country X incurs a constant marginal cost c, and the firm in country Y incurs no marginal cost of production. Finally, the specific tariffs imposed on imports by X and Y are labeled τ_x and τ_y, respectively.

The profit maximization problem of the firm in country X is written as:

$$\Pi_x = \max_{q_{x,x}} \left\{ \left(A - b \left(q_{x,x} + q_{x,y} \right) - c \right) q_{x,x} \right\} + \max_{q_{y,x}} \left\{ \left(A - b \left(q_{y,x} + q_{y,y} \right) - c - \tau_y \right) q_{y,x} \right\} \tag{1}$$

Here, $q_{i,j}$ represents production by the firm in country j that is consumed in country i. Solving the maximization problem in (1), we have the following first order conditions:

$$\begin{aligned} A - 2bq_{x,x} - q_{x,y} &= c \\ A - 2bq_{y,x} - q_{y,y} &= c + \tau_y \end{aligned}$$

Analogously, for the firm in y, the maximization problem is

$$\Pi_y = \max_{q_{y,y}} \left\{ \left(A - b \left(q_{y,y} + q_{y,x} \right) \right) q_{y,y} \right\} + \max_{q_{x,y}} \left\{ \left(A - b \left(q_{x,y} + q_{x,x} \right) - \tau_x \right) q_{x,y} \right\} s \tag{2}$$

Here, I have imposed that firms in the south, as described above, incur no marginal production cost. Solving the maximization problem in (2), we have the following first order conditions:

$$\begin{aligned} A - 2bq_{x,y} - q_{x,x} &= \tau_x \\ A - 2bq_{y,y} - q_{y,x} &= 0 \end{aligned}$$

Solving the system of four first order conditions yields optimal quantities,

$$\begin{aligned} q_{x,x} &= \frac{A - 2c + \tau_x}{3b} \\ q_{y,x} &= \frac{A - 2(\tau_y + c)}{3b} \\ q_{y,y} &= \frac{A + \tau_y + c}{3b} \\ q_{x,y} &= \frac{A - 2\tau_x + c}{3b} \end{aligned}$$

and then profits:

$$\begin{aligned}
\Pi_{x,x} &= \frac{(A-2c+\tau_x)^2}{9b} \\
\Pi_{y,x} &= \frac{(A-2(\tau_y+c))^2}{9b} \\
\Pi_{y,y} &= \frac{(A+\tau_y+c)^2}{9b} \\
\Pi_{x,y} &= \frac{(A-2\tau_x+c)^2}{9b}
\end{aligned}$$

In the above profit functions, tariffs naturally protect home profits (eg, $\frac{\partial \Pi_{x,x}}{\partial \tau_x} > 0$), and hurt export profits (eg, $\frac{\partial \Pi_{y,x}}{\partial \tau_y} < 0$).

Rolling back to Stages 1 and 2, there are two possible outcomes: signing a free trade agreement, and not signing a free trade agreement. Under Free Trade (F), countries agree to levy no tariffs. For this case, welfare is simply the sum of producer surplus and consumer surplus setting $\tau_x = \tau_y = 0$. Hence, for each country, welfare is written as:

$$\begin{aligned}
W_x^F &= CS_x + \Pi_{x,x} + \Pi_{y,x} = \frac{(2A-c)^2}{18b} + \frac{(A-2c)^2}{9b} + \frac{(A-2c)^2}{9b} \\
W_y^F &= CS_y + \Pi_{y,y} + \Pi_{x,y} = \frac{(2A-c)^2}{18b} + \frac{(A+c)^2}{9b} + \frac{(A+c)^2}{9b}
\end{aligned}$$

For free trade, reciprocal exports occur as long as $c < 1/2$.

The second possible agreement outcome within the two country model is noncooperative tariff setting (M). For this case, welfare is written as the sum of consumer surplus, producer surplus, and tariff revenue. Countries maximize welfare by choosing an optimal value of their specific tariff. Hence:

$$\begin{aligned}
W_x^M &= \max_{\tau_x} \left\{ \frac{(A-2c+\tau_x)^2}{9b} + \frac{(A-2(\tau_y+c))^2}{9b} + \frac{\tau_x(A-2\tau_x+c)}{3b} + \frac{(2A-c-\tau_x)^2}{18b} \right\} \\
W_y^M &= \max_{\tau_y} \left\{ \frac{(A+\tau_y+c)^2}{9b} + \frac{(A-2\tau_x+c)^2}{9b} + \frac{\tau_y(A-2(\tau_y+c))}{3b} + \frac{(2A-\tau_y-c)^2}{18b} \right\}
\end{aligned}$$

In each welfare function, the first two terms are profits in each market, the third term is tariff revenue, and the last term consumer surplus. Solving each maximization problem, optimal tariffs are written as:

$$\begin{aligned}
\widehat{\tau}_x &= \frac{A}{3} \\
\widehat{\tau}_y &= \frac{A-c}{3}
\end{aligned}$$

Welfare evaluated at the optimal tariffs is written as:

$$\begin{aligned} W_x^M(\widehat{\tau}_x, \widehat{\tau}_y) &= \frac{(65A^2 - 124Ac + 113c^2)}{162b} \\ W_y^M(\widehat{\tau}_x, \widehat{\tau}_y) &= \frac{(65A^2 - 6Ac + 54c^2)}{162b} \end{aligned}$$

Finally, when there is no agreement, two-way trade is guaranteed as long as $c < A/4$.[7] For this section, I will restrict analysis to this region of c.

Agreement Choice

In this simple model, countries X and Y choose between noncooperative tariff setting and a free trade agreement. For the cooperative solution concept which I'll be using throughout the draft, the basic question is whether, when comparing two agreements, all countries that change agreement status prefer the same agreement. If this occurs, the inferior agreement is *dominated*, and discarded from the core of the agreement formation game. In the current two country framework, the question is whether welfare is highest for both X and Y for the same agreement, or whether over the given parameter space countries X and Y disagree over the welfare maximizing trade agreement.

For country X, the free trade agreement is preferred when $W_x^F > W_x^M(\widehat{\tau}_x, \widehat{\tau}_y)$. Precisely, this occurs when:

$$\Delta_x = W_x^F - W_x^M(\widehat{\tau}_x, \widehat{\tau}_y) = \frac{7A^2 - 56Ac + 40c^2}{162b} > 0$$

For country Y, this occurs when:

$$\Delta_y = W_y^F - W_y^M(\widehat{\tau}_x, \widehat{\tau}_y) = \frac{(7A^2 + 42Ac - 9c^2)}{162b} > 0$$

Clearly, both Δ_x and Δ_y are positive when $c = 0$. However, in the neighborhood of $c = 0$, Δ_x and Δ_y respond differently to changes in c. Precisely:

$$\begin{aligned} \frac{\partial \Delta_x}{\partial c} &= -\frac{56A - 80c}{162b} \\ \frac{\partial \Delta_y}{\partial c} &= \frac{42A - 18c}{162b} \end{aligned}$$

Here, when $c = 0$, $\frac{\partial \Delta_x}{\partial c} < 0$ and $\frac{\partial \Delta_y}{\partial c} > 0$. Interpreting $\frac{\partial \Delta_x}{\partial c} < 0$, the advantage of free trade diminishes when the firm in Y becomes relatively more efficient (c, the cost disadvantage of X, goes up). Intuitively, higher c shifts the competitive advantage toward the firm from Y in both markets. If this competitive advantage is large enough, the added consumer benefit of free access by firms from Y will be outweighed by the loss in profits by the firm in X and the loss in potential tariff revenue from levying a tariff on a large volume of imports from Y. The following proposition precisely characterizes when cost-differences between countries are large enough to yield disagreement over whether to sign a free trade agreement.

Proposition 1 *There exists a $\widehat{c} < A/4$ such that the core of the agreement formation game is as*

[7] To see this, plug $\widehat{\tau}_x$ and $\widehat{\tau}_y$ into quantities. Exports from X to Y will be positive only if $c < 1/4$.

follows:

$$\begin{array}{lll} \textit{For} & c \in [0, \widehat{c}) & \{F\} \\ & c \in (\widehat{c}, A/4) & \{F, M\} \end{array}$$

Proof. *See Appendix* ■

For an agreement outcome to be excluded from the core, both countries must prefer a common alternative. In Proposition 1, for $c < \widehat{c}$, both X and Y agree that free trade is optimal, where $\Delta_x > 0$ and $\Delta_y > 0$, and hence, M is dominated. However, for $c > \widehat{c}$, I find that neither free trade (F) or no agreement (M) are dominant. Precisely, over this region, $\Delta_x < 0$ and $\Delta_y > 0$. Again, while Y prefers the enhanced competitive advantage when c is relatively high, X is worse off given the enhanced advantage for Y. Overall, there exists a large enough value of c such that the dichotomy described above becomes prohibitive, and M is not dominated by F in equilibrium.

The overall point of this section is as follows. When evaluating the benefits of a free trade agreement between two countries that may differ in efficiency, one must account for the fact that tariffs may be critical in protecting the less efficient country from international competition. In some cases, as in Proposition 1, this results in a case in which the high cost country prefers noncooperative Nash tariff setting to a free trade agreement. I now examine this effect within a multi-country/region model in which agreement options are more varied, and later, when transfers and political influence are present.

3 A four country, two region model

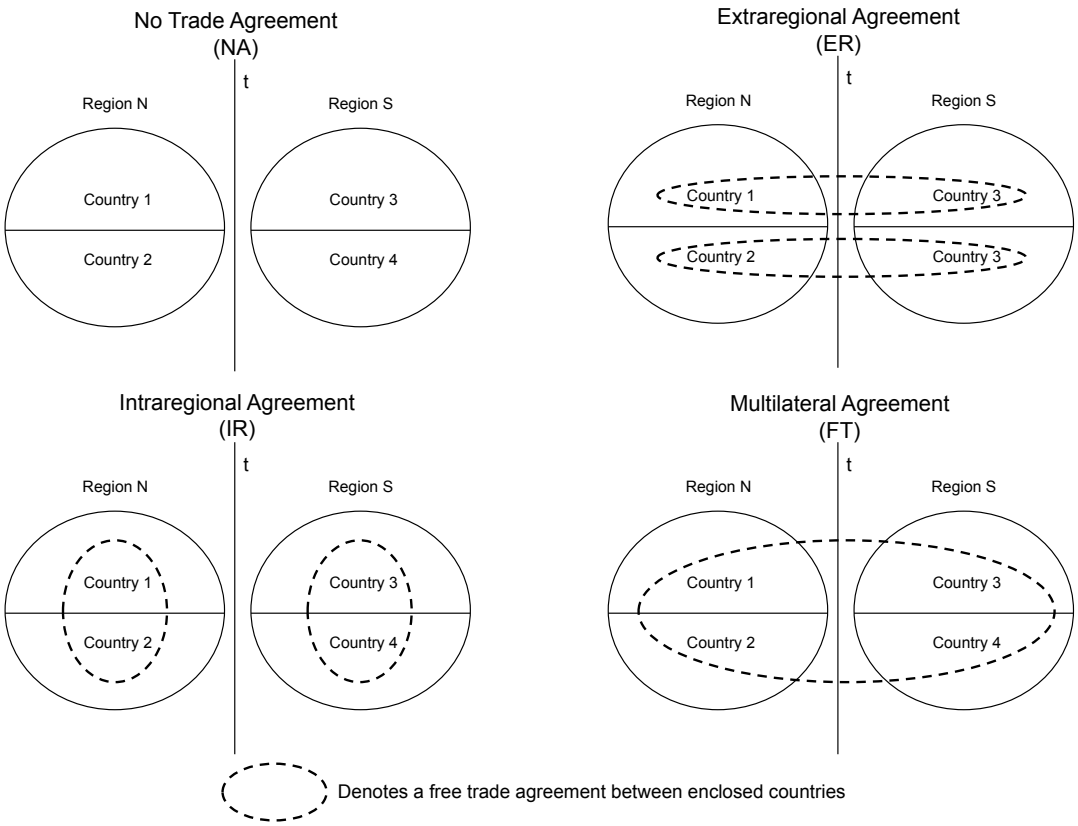
The setup of the expanded model is kept simple in order to focus on the key fundamentals of trade agreements, and how countries may agree and disagree over which trading arrangements are the best option. Different from the previous section, the world now consists of two regions, N and S, separated by a trade cost, t. The trade cost is per unit, and represents the costs of transportation. Within Region N, there are two countries, 1 and 2. Within Region S, there are also two countries, 3 and 4.

The timing of the model is again simple. There are three stages. In the first stage, countries form binding coalitions (trade agreements), discarding those coalitions which are mutually inefficient. The mechanism behind this stage will be detailed shortly. In the second stage, subject to these agreements, countries set their external MFN tariffs. Within coalitions, countries may set their own external tariff, and as such, this framework is better suited to address free trade agreement formation rather than currency union formation. In the third and final stage, subject to agreements and tariff choices, firms make output decisions for each market. I introduce further specifics in reverse order.

The consumer setup is very similar to above, where demand is assumed to take the form of $P = A - bQ$ in each country. The production side of the economy is also similar, where firms in countries one and two incur a constant marginal cost c, and firms in countries three and four incur no cost of production. I again assume that each country has only one producer.

In the tariff setting stage, each country is allowed to levy a common external tariff against imports from non-agreement countries. Tariffs against countries within an agreement are assumed to be zero. Countries maximize the sum of profits, consumer surplus and tariff revenue when choosing their external tariff. Later, I will drop this assumption and examine a restricted model in which profits are the only component of the welfare function.

Figure 2: Possible Equilibrium Trade Agreements



In the first stage, countries form binding trade agreements. A set of trade agreements will be called an *agreement structure*. A given agreement structure will be *dominated* if some other agreement structure provides higher welfare to all decisive countries. A *decisive country* is one which experiences a change in agreement status when comparing one structure to another. Assumptions of symmetry will guarantee that all countries are decisive countries. That is, I assume that countries choose between four possible agreement structures. These are illustrated in Figure 2.[8]

In Figure 2, the dashed circles represent a free trade agreement between the enclosed members. In the upper-left panel, there are no agreements (NA), and countries simply levy MFN tariffs. In the upper-right panel, each northern country forms an extraregional agreement (ER) with a low cost southern country. Here, each country levies a common MFN tariff on one low and one high cost country. In the bottom-left panel, countries form intraregional agreements (IR), where MFN

[8]Other asymmetric agreements have been analyzed, but add no further insight to the basic process of coalition formation. Specifically, asymmetric agreements only add to the number of agreements in the core when costs differences are high. Thus, they are omitted to reduce clutter.

tariffs are levied against the opposing region. Finally, in the bottom-right panel, the world forms a multilateral free trade agreement, and no MFN tariffs are imposed.

Similar to section two, $q_{i,j}$ represents production in country j sold in country i, and $\Pi_{i,j}$ represents profits earned by country j in country i. As mentioned above, let t be the per-unit transportation cost in shipping a good from region N to region S. Furthermore, let τ_i be the MFN tariff applied to all countries outside i. Finally, let $I_{i,j} = 1$ if there is a trade agreement between countries i and j, and 0 otherwise.

With this notation, without loss of generality, I will focus on production decisions for the markets in Countries 1 and 3. Countries 2 and 4 are identical to Countries 1 and 3, respectively, by the assumption symmetry. With this in mind, firms solve the following maximization problems in serving Country 1.

$$\begin{aligned}
\Pi_{1,1} &= \max_{q_{1,1}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}) - c)\, q_{1,1}\} \\
\Pi_{1,2} &= \max_{q_{1,2}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}) - (1 - I_{1,2})\tau_1 - c)\, q_{1,2}\} \\
\Pi_{1,3} &= \max_{q_{1,3}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}) - (1 - I_{1,3})\tau_1 - t)\, q_{1,3}\} \\
\Pi_{1,4} &= \max_{q_{1,4}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}) - (1 - I_{1,4})\tau_1 - t)\, q_{1,4}\}
\end{aligned}$$

Likewise the maximization problems in serving Country 3 are:

$$\begin{aligned}
\Pi_{3,1} &= \max_{q_{3,1}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}) - (1 - I_{3,1})\tau_3 - t - c)\, q_{1,1}\} \\
\Pi_{3,2} &= \max_{q_{3,2}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}) - (1 - I_{3,2})\tau_3 - t - c)\, q_{1,2}\} \\
\Pi_{3,3} &= \max_{q_{3,3}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}))\, q_{1,3}\} \\
\Pi_{3,4} &= \max_{q_{3,4}} \{(A - b(q_{1,1} + q_{1,2} + q_{1,3} + q_{1,4}) - (1 - I_{3,4})\tau_3)\, q_{1,4}\}
\end{aligned}$$

I will now solve each of these profit maximization problems for each particular case outlined in Figure 2. Then, I will present the relevant welfare function, and solve for the optimal MFN tariff for each country within each agreement structure. Welfare functions after substituting for the optimal value of tariffs are relegated to the appendix.

No Agreement (NA)

For this arrangement, $I_{i,j} = 0$ for all i and j. Quantities and profits in serving Country 1 are written as:

$$\begin{aligned}
\widehat{\Pi}_{1,1}^{NA} &= \frac{(A + 2t + 3\tau_1 - 3c)^2}{25b},\ \widehat{q}_{1,1}^{NA} = \frac{(A + 2t + 3\tau_1 - 3c)}{5b} \\
\widehat{\Pi}_{1,2}^{NA} &= \frac{(A + 2t - 2\tau_1 - 3c)^2}{25b},\ \widehat{q}_{1,2}^{NA} = \frac{(A + 2t - 2\tau_1 - 3c)}{5b} \\
\widehat{\Pi}_{1,3}^{NA} &= \frac{(A - 3t - 2\tau_1 + 2c)^2}{25b},\ \widehat{q}_{1,3}^{NA} = \frac{(A - 3t - 2\tau_1 + 2c)}{5b} \\
\widehat{\Pi}_{1,4}^{NA} &= \frac{(A - 3t - 2\tau_1 + 2c)^2}{25b},\ \widehat{q}_{1,4}^{NA} = \frac{(A - 3t - 2\tau_1 + 2c)}{5b}
\end{aligned}$$

The profits and quantities in serving Country 3 are written as:

$$\begin{aligned}
\widehat{\Pi}_{3,1}^{NA} &= \frac{(A-3t-2\tau_3-3c)^2}{25b},\ \widehat{q}_{3,1}^{NA} = \frac{(A-3t-2\tau_3-3c)}{5b} \\
\widehat{\Pi}_{3,2}^{NA} &= \frac{(A-3t-2\tau_3-3c)^2}{25b},\ \widehat{q}_{3,2}^{NA} = \frac{(A-3t-2\tau_3-3c)}{5b} \\
\widehat{\Pi}_{3,3}^{NA} &= \frac{(A+2t+3\tau_3+2c)^2}{25b},\ \widehat{q}_{3,3}^{NA} = \frac{(A+2t+3\tau_3+2c)}{5b} \\
\widehat{\Pi}_{3,4}^{NA} &= \frac{(A+2t-2\tau_3+2c)^2}{25b},\ \widehat{q}_{3,4}^{NA} = \frac{(A+2t-2\tau_3+2c)}{5b}
\end{aligned}$$

Again, note that $\widehat{\Pi}_{2,j}^{NA}$ and $\widehat{\Pi}_{4,j}^{NA}$ for all j are defined analogously, replacing τ_1 and τ_3 with τ_2 and τ_4, respectively. Naturally, trade barriers, whether they be physical or revenue collecting instruments, protect domestic profits (eg. $\frac{\partial \widehat{\Pi}_{1,1}^{NA}}{\partial t} > 0$ and $\frac{\partial \widehat{\Pi}_{1,1}^{NA}}{\partial \tau_1} > 0$) . Further, physical trade barriers protect export profits within regions, at the expense of export profits by firms in the other region (eg $\frac{\partial \widehat{\Pi}_{1,2}^{NA}}{\partial t} > 0$ and $\frac{\partial \widehat{\Pi}_{1,3}^{NA}}{\partial t} < 0$). However, before getting too far into the analysis of trade costs, one must endogenously set tariffs to examine the true effect of trade costs, both directly through production decisions, and indirectly through optimal tariffs.

For the moment, I assume no political pressures in setting the optimal tariff, where countries maximize the sum of consumer surplus, the profits of domestic firms (tariffs do not affect export profits in this framework), and tariff revenue. Focusing on firms in region N, the Country 1 government solves the following maximization problem when setting their optimal tariff:

$$W_1^{NA} = \max_{\tau_1} \left\{ CS_1^{NA}(\tau_1) + \widehat{\Pi}_{1,1}^{NA}(\tau_1) + +\tau_1 \left(\widehat{q}_{1,2}^{NA} + \widehat{q}_{1,3}^{NA} + \widehat{q}_{1,4}^{NA} \right) \right\} + \widehat{\Pi}_{2,1}^{NA}(\tau_2) + \widehat{\Pi}_{3,1}^{NA}(\tau_3) + \widehat{\Pi}_{4,1}^{NA}(\tau_4)$$

The optimal tariff derived from the above maximization problem is:

$$\widehat{\tau}_1^{NA} = \frac{9A-2t-7c}{33}$$

Moving on to the tariff chosen by countries in the south, I will focus on country 3. Welfare is written as:

$$W_3^{NA} = \max_{\tau_3} \left\{ CS_3^{NA}(\tau_3) + \widehat{\Pi}_{3,3}^{NA}(\tau_3) + \tau_3 \left(\widehat{q}_{3,1}^{NA} + \widehat{q}_{3,2}^{NA} + \widehat{q}_{3,4}^{NA} \right) \right\} + \widehat{\Pi}_{1,3}^{NA}(\tau_1) + \widehat{\Pi}_{2,3}^{NA}(\tau_2) + \widehat{\Pi}_{4,3}^{NA}(\tau_4)$$

The optimal $\widehat{\tau}_3$ tariff is:

$$\widehat{\tau}_3^{NA} = \frac{9A-2t-2c}{33}$$

Note that $\widehat{\tau}_3^{NA}$ is larger than $\widehat{\tau}_1^{NA}$ for positive c. The intuition is that the consumption distortion is lower for these countries, as firms in S are better suited than firms in N to offset the consumption distortion from a higher tariff at home. Thus, countries in S levy a higher tariff.

Prohibitive trade cost - marginal cost combinations

The NA structure will be the most competitive, and thus will generally require the most restrictive set of parameter values such that all countries trade with one another. I now derive the region of

parameters such that this is the case. First, substituting $\widehat{\tau}_1^{NA}$ and $\widehat{\tau}_3^{NA}$ into $\widehat{q}_{i,j}^{NA\prime}s$, I get:

$$\begin{aligned}
\widehat{q}_{1,1}^{NA} &= \frac{12(A+t-2c)}{33b},\ \widehat{q}_{1,2}^{NA} = \frac{3A+14t-17c}{33b},\ \widehat{q}_{1,3}^{NA} = \frac{3A-19t+16c}{33b},\ \widehat{q}_{1,4}^{NA} = \frac{3A-19t+16c}{33b} \\
\widehat{q}_{3,1}^{NA} &= \frac{3A-19t-19c}{33b},\ \widehat{q}_{3,2}^{NA} = \frac{3A-19t-19c}{33b},\ \widehat{q}_{3,3}^{NA} = \frac{12(A+t+c)}{33b},\ \widehat{q}_{3,4}^{NA} = \frac{3A+14t+14c}{33b}
\end{aligned}$$

Not surprisingly, $\widehat{q}_{3,1}^{NA}$ and $\widehat{q}_{3,2}^{NA}$ are the lowest production levels at any given pair of c and t. Specifically, these are firms in 1 and 2 exporting to 3. As firms in 1 and 2 are the highest cost suppliers, and incur the transportation cost t in serving the opposing region, these firms will be the first to stop trading when trade costs or production costs are too high. Thus, I adopt the following restriction on parameter values to ensure that bilateral trade remains positive between all countries:

$$3A - 19t - 19c > 0 \tag{3}$$

This condition will be used to identify whether points of indifference between agreement structures occur within the set of parameters in which trade occurs between all countries for all agreement structures.

Intraregional Agreement (IR)

For this agreement structure, each country enters an enforceable free trade agreement with the other country in its own region. This implies that $I_{1,2} = I_{2,1} = I_{4,3} = I_{3,4} = 1$, and $I_{i,j} = 0$ for all other i and j. With this assignment of $I_{i,j}$, quantities and profits in serving Country 1 are written as:

$$\begin{aligned}
\widehat{\Pi}_{1,1}^{IR} &= \frac{(A+2t+2\tau_1-3c)^2}{25b},\ \widehat{q}_{1,1}^{IR} = \frac{(A+2t+2\tau_1-3c)}{5b} \\
\widehat{\Pi}_{1,2}^{IR} &= \frac{(A+2t+2\tau_1-3c)^2}{25b},\ \widehat{q}_{1,2}^{IR} = \frac{(A+2t+2\tau_1-3c)}{5b} \\
\widehat{\Pi}_{1,3}^{IR} &= \frac{(A-3t-3\tau_1+2c)^2}{25b},\ \widehat{q}_{1,3}^{IR} = \frac{(A-3t-3\tau_1+2c)}{5b} \\
\widehat{\Pi}_{1,4}^{IR} &= \frac{(A-3t-3\tau_1+2c)^2}{25b},\ \widehat{q}_{1,4}^{IR} = \frac{(A-3t-3\tau_1+2c)}{5b}
\end{aligned}$$

The profits and quantities in serving Country 3 are written as:

$$\begin{aligned}
\widehat{\Pi}_{3,1}^{IR} &= \frac{(A-3\tau_3-3t-3c)^2}{25b},\ \widehat{q}_{3,1}^{IR} = \frac{(A-3\tau_3-3t-3c)}{5b} \\
\widehat{\Pi}_{3,2}^{IR} &= \frac{(A-3\tau_3-3t-3c)^2}{25b},\ \widehat{q}_{3,2}^{IR} = \frac{(A-3\tau_3-3t-3c)}{5b} \\
\widehat{\Pi}_{3,3}^{IR} &= \frac{(A+2\tau_3+2t+2c)^2}{25b},\ \widehat{q}_{3,3}^{IR} = \frac{(A+2\tau_3+2t+2c)}{5b} \\
\widehat{\Pi}_{3,4}^{IR} &= \frac{(A+2\tau_3+2t+2c)^2}{25b},\ \widehat{q}_{3,4}^{IR} = \frac{(A+2\tau_3+2t+2c)}{5b}
\end{aligned}$$

As before, Country 1 solves the following maximization problem when setting their optimal tariff:

$$W_1^{IR} = \max_{\tau_1} \left\{ CS_1^{IR}(\tau_1) + \widehat{\Pi}_{1,1}^{IR}(\tau_1) + \tau_1 \left(\widehat{q}_{1,3}^{IR} + \widehat{q}_{1,4}^{IR} \right) \right\} + \widehat{\Pi}_{2,1}^{IR}(\tau_2) + \widehat{\Pi}_{3,1}^{IR}(\tau_3) + \widehat{\Pi}_{4,1}^{IR}(\tau_4)$$

The optimal tariff derived from the above maximization problem is:

$$\widehat{\tau}_1^{IR} = \frac{A - 3t + 2c}{8}$$

For Country 3, welfare is written as:

$$W_3^{IR} = \max_{\tau_3} \left\{ CS_3^{IR}(\tau_3) + \widehat{\Pi}_{3,3}^{IR}(\tau_3) + \tau_3 \left(\widehat{q}_{3,1}^{IR} + \widehat{q}_{3,2}^{IR} \right) \right\} + \widehat{\Pi}_{1,3}^{IR}(\tau_1) + \widehat{\Pi}_{2,3}^{IR}(\tau_2) + \widehat{\Pi}_{4,3}^{IR}(\tau_4)$$

The optimal $\widehat{\tau}_3$ tariff is:

$$\widehat{\tau}_3^{IR} = \frac{A - 3t - 3c}{8}$$

Here, note that $\widehat{\tau}_3^{IR}$ is smaller than $\widehat{\tau}_1^{IR}$ for positive c. The intuition is that as the cost difference c increases, export supply decreases, which decreases the incentive of countries in S to collect tariff revenue.

Extraregional Agreement (ER)

For this arrangement, each country enters an enforceable free trade agreement with a country in the opposing region. Assuming without loss of generality that 1 joins with 3 and 2 joins with 4, this implies the following assignments of $I_{i,j}$:

$$I_{1,3} = I_{3,1} = I_{4,2} = I_{2,4} = 1$$

All other $I_{i,j} = 0$. Quantities and profits in serving Country 1 are written as:

$$\begin{aligned}
\widehat{\Pi}_{1,1}^{ER} &= \frac{(A + 2t + 2\tau_1 - 3c)^2}{25b}, \ \widehat{q}_{1,1}^{ER} = \frac{(A + 2t + 2\tau_1 - 3c)}{5b} \\
\widehat{\Pi}_{1,2}^{ER} &= \frac{(A + 2t - 3\tau_1 - 3c)^2}{25b}, \ \widehat{q}_{1,2}^{ER} = \frac{(A + 2t - 3\tau_1 - 3c)}{5b} \\
\widehat{\Pi}_{1,3}^{ER} &= \frac{(A - 3t + 2\tau_1 + 2c)^2}{25b}, \ \widehat{q}_{1,3}^{ER} = \frac{(A - 3t + 2\tau_1 + 2c)}{5b} \\
\widehat{\Pi}_{1,4}^{ER} &= \frac{(A - 3t - 3\tau_1 + 2c)^2}{25b}, \ \widehat{q}_{1,4}^{ER} = \frac{(A - 3t - 3\tau_1 + 2c)}{5b}
\end{aligned}$$

Profits and quantities in serving Country 3 are written as:

$$\begin{aligned}
\widehat{\Pi}_{3,1}^{ER} &= \frac{(A+2\tau_3-3t-3c)^2}{25b},\ \widehat{q}_{3,1}^{ER} = \frac{(A+2\tau_3-3t-3c)}{5b} \\
\widehat{\Pi}_{3,2}^{ER} &= \frac{(A-3\tau_3-3t-3c)^2}{25b},\ \widehat{q}_{3,2}^{ER} = \frac{(A-3\tau_3-3t-3c)}{5b} \\
\widehat{\Pi}_{3,3}^{ER} &= \frac{(A+2\tau_3+2t+2c)^2}{25b},\ \widehat{q}_{3,3}^{ER} = \frac{(A+2\tau_3+2t+2c)}{5b} \\
\widehat{\Pi}_{3,4}^{ER} &= \frac{(A-3\tau_3+2t+2c)^2}{25b},\ \widehat{q}_{3,4}^{ER} = \frac{(A-3\tau_3+2t+2c)}{5b}
\end{aligned}$$

As before, Country 1 solves the following maximization problem when setting their optimal tariff:

$$W_1^{ER} = \max_{\tau_1}\left\{CS_1^{ER}(\tau_1) + \widehat{\Pi}_{1,1}^{ER}(\tau_1) + \tau_1\left(\widehat{q}_{1,2}^{ER} + \widehat{q}_{1,4}^{ER}\right)\right\} + \widehat{\Pi}_{2,1}^{ER}(\tau_2) + \widehat{\Pi}_{3,1}^{ER}(\tau_3) + \widehat{\Pi}_{4,1}^{ER}(\tau_4)$$

The optimal tariff derived from the above maximization problem is:

$$\widehat{\tau}_1^{ER} = \frac{6A+7t-13c}{48}$$

Across the board, a higher value of c decreases the value of protection. In contrast, trade costs evidently increase the value of protection within the ER setup.

Moving on to the tariff chosen by Country 3, welfare is written as:

$$W_3^{IR} = \max_{\tau_3}\left\{CS_3^{ER}(\tau_3) + \widehat{\Pi}_{3,3}^{ER}(\tau_3) + \tau_3\left(\widehat{q}_{3,2}^{ER} + \widehat{q}_{3,4}^{ER}\right)\right\} + \widehat{\Pi}_{1,3}^{ER}(\tau_1) + \widehat{\Pi}_{2,3}^{ER}(\tau_2) + \widehat{\Pi}_{4,3}^{ER}(\tau_4)$$

The optimal $\widehat{\tau}_3$ tariff is:

$$\widehat{\tau}_3^{ER} = \frac{6A+7t+7c}{48}$$

Unlike $\widehat{\tau}_1^{ER}$, $\widehat{\tau}_3^{ER}$ is increasing in both the northern cost c and transportation cost t.

Multilateral Free Trade (FT)

For this arrangement, every country agrees to free trade with all countries. Thus, $I_{i,j} = 1$ for all i and j. Quantities and profits in serving Country 1 are written as:

$$\begin{aligned}
\widehat{\Pi}_{1,1}^{FT} &= \frac{(A+2t-3c)^2}{25b},\ \widehat{q}_{1,1}^{FT} = \frac{(A+2t-3c)}{5b} \\
\widehat{\Pi}_{1,2}^{FT} &= \frac{(A+2t-3c)^2}{25b},\ \widehat{q}_{1,2}^{FT} = \frac{(A+2t-3c)}{5b} \\
\widehat{\Pi}_{1,3}^{FT} &= \frac{(A-3t+2c)^2}{25b},\ \widehat{q}_{1,3}^{FT} = \frac{(A-3t+2c)}{5b} \\
\widehat{\Pi}_{1,4}^{FT} &= \frac{(A-3t+2c)^2}{25b},\ \widehat{q}_{1,4}^{FT} = \frac{(A-3t+2c)}{5b}
\end{aligned}$$

Profits and quantities in serving Country 3 are written as:

$$\begin{aligned}
\widehat{\Pi}_{3,1}^{FT} &= \frac{(A-3t-3c)^2}{25b},\ \widehat{q}_{3,1}^{FT} = \frac{(A-3t-3c)}{5b} \\
\widehat{\Pi}_{3,2}^{FT} &= \frac{(A-3t-3c)^2}{25b},\ \widehat{q}_{3,2}^{FT} = \frac{(A-3t-3c)}{5b} \\
\widehat{\Pi}_{3,3}^{FT} &= \frac{(A+2t+2c)^2}{25b},\ \widehat{q}_{3,3}^{FT} = \frac{(A+2t+2c)}{5b} \\
\widehat{\Pi}_{3,4}^{FT} &= \frac{(A+2t+2c)^2}{25b},\ \widehat{q}_{3,4}^{FT} = \frac{(A+2t+2c)}{5b}
\end{aligned}$$

With all welfare functions listed in the appendix, I can now roll back to stage 1 and derive which agreements may be mutually beneficial for countries in the North and South. I will first present results for the polar case of no cost differences between regions. Then, I will eliminate trade costs while addressing the case of asymmetric production costs. Finally, I will present results allowing for positive transportation costs and cost asymmetry between regions.

4 Equilibrium: Cost symmetry with costly trade

In this section, I first examine the polar case of no production cost differences between countries. Thus, the only costs which affect the distribution of production and consumption are the physical trade barriers embodied in t, and the optimal MFN tariffs. Finally, since there are no cost differences, and each trading arrangement is symmetric, welfare functions for all countries will be identical. Thus, it suffices to examine one country in assigning preference relationships over the four possible structures of trade agreements. By doing so, Proposition 2 details the equilibrium agreement structures which comprise the core of the agreement formation game.

Proposition 2 *When $c = 0$, over the non-prohibitive region of trade costs, NA and ER are always dominated. Further, if $t \in \left(0, \frac{9}{107}A\right)$, free trade ($FT$) is preferred to an intraregional agreement (IR). For $t \in \left(\frac{9}{107}A, \frac{3}{19}A\right)$, IR is preferred to FT. Thus, the equilibrium of the model is summarized as follows:*

$$\begin{array}{lll}
\textit{For} & t \in \left[0, \frac{9}{107}A\right) & FT \textit{ remains undominated} \\
 & t \in \left[\frac{9}{107}A, \frac{3}{19}A\right) & IR \textit{ remains undominated}
\end{array}$$

Proof. See Appendix ■

The result in Proposition 2 is very similar to a result in Frankel, Stein and Wei (1995). The intuition is explained as follows. Generally, in models of coalition formation, it is ideal to form a coalition with the strongest competitor(s). When trade costs are high, the strongest competitor for each country is its regional partner. In forming regional coalitions, countries trade market access while still collecting some tariff revenue. This more than compensates for a small consumer loss when compared with an outcome of free trade. Since countries outside of the agreement are distant, the amount of diverted trade from these partners is minimal. On the other hand, when trade costs are low, the trade diversion associated with tariffs is significant, and in a symmetric equilibrium, all countries are worse off under an intraregional agreement. Thus, countries choose instead to form a coalition with all countries (global free trade).

Of course, this logic fails to account for any cost-asymmetry between countries. If regions are asymmetric in terms of productivity, as in section two, high and low cost regions may prefer different agreements, and thus disagreement may arise in choosing a preferred equilibrium outcome. I now turn to addressing precisely this issue.

5 Equilibrium: Costless trade with cost asymmetry

Unlike Section 4, this section will allow for cost differences between regions, but with no trade costs between regions. Since regions are now asymmetric, welfare functions of countries in both regions must be analyzed to determine which equilibrium outcomes are dominated or undominated for a given value of c. Doing so in the appendix, I prove the following proposition:

Proposition 3 *For $t = 0$, there exists a c_1 and c_2 such that $0 < c_1 < c_2 < \frac{3}{19}A$, and that the core of the agreement formation game is summarized as follows:*

$$\begin{array}{lll} \textit{For} & c \in [0, c_1) & FT \textit{ remains undominated} \\ & c \in [c_1, c_2) & FT \textit{ and } IR \textit{ remain undominated} \\ & c \in \left[c_2, \frac{3}{19}A\right) & FT, IR \textit{ and } ER \textit{ remain undominated} \end{array}$$

Proof. See Appendix ■

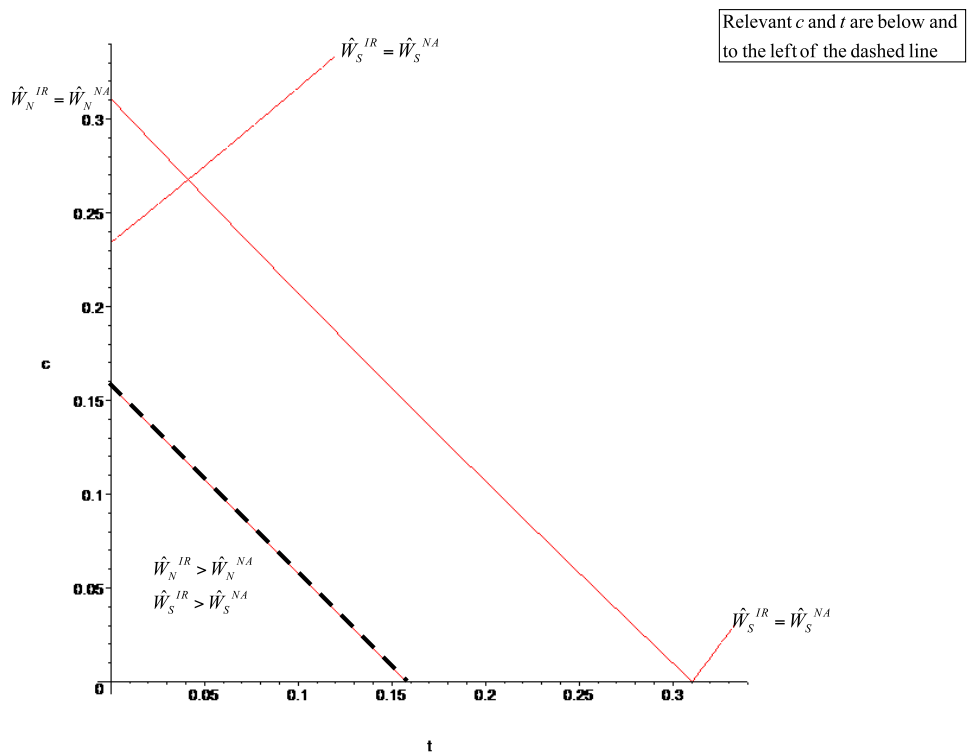
The results here are similar to section two. Countries will form a given trade agreement if all parties are better off relative to all other options. In the case of no trade costs, the critical feature is the cost advantage of southern countries. If this cost advantage is low, then all countries agree that the benefits of market access through free trade offset any loss in market power or tariff revenue. However, if the southern cost advantage is high, then northern firms wish to limit the amount of market access given to southern firms, usually by forming an intraregional trade agreement. This causes a disagreement regarding which agreement structure is best, as southern firms still desire full unrestricted market access to northern markets. In equilibrium, over a large region of relevant c, there exist multiple agreement structures which comprise the core of the agreement formation game.

6 Equilibrium: Cost asymmetry and costly trade

In this section, I allow for both production cost differences and trade costs between regions to analyze how both types of asymmetry jointly affect the core of the agreements formation game. Further, this will yield an estimate regarding the scope for agreement in trade negotiations. The bottom line of this section will be that there is little scope for agreement in the agreements formation game.

To clearly analyze the problem in two dimensions, I henceforth assume that $A = b = 1$. This will enable a graphical analysis of indifference loci between agreement outcomes. To begin, Figure 3 establishes the dominance relationship between NA and IR. In Figure 3, the heavy dashed line represents the largest values of c and t such that bilateral trade occurs between all countries for all agreement structures. Thus, the relevant range of parameters is below this dashed line and to the left. The other lines in Figure 3 represent indifference points between IR and NA for countries in N and S. Clearly, these loci of indifference are outside of the relevant range of c and t. Further,

Figure 3: No Agreement (NA) dominated by an Intraregional Agreement (IR)



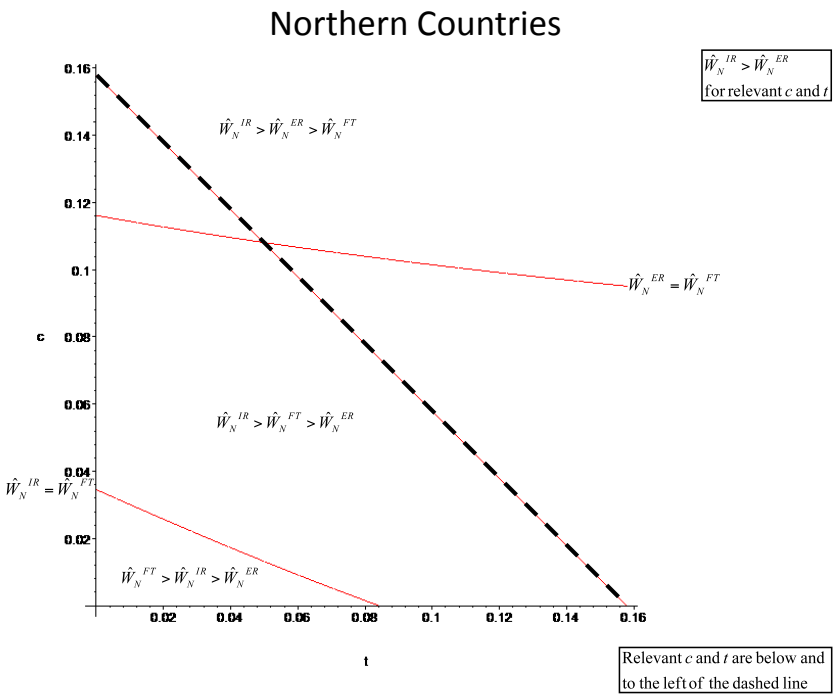
by comparing welfare functions at $c = t = 0$, it is easy to show that NA is dominated by IR for all relevant values of t and c.[9]

Moving forward, and henceforth excluding NA (since it is always dominated), I will focus on the preference conditions for Northern countries. These preference conditions are presented in the top panel of Figure 4. Northern countries prefer IR to all other arrangements, except when cost differences or trade costs are low (where they prefer FT). The intuition is that if Northern countries are willing to trade market access, they are wary of doing so with low cost countries, as it may disproportionately injure their domestic firms relative to an agreement with a regional partner. On the other hand, in the bottom panel of Figure 4, Southern countries prefer the enhanced market access to the northern region via FT and ER unless trade costs are high. The intuition is that their domestic sector suffers little from high cost import competition, and benefits greatly from increased market access to markets currently served by less competitive firms.

Putting the two panels of Figure 4 together, the trading arrangements that remain undominated in equilibrium are identified in Figure 5. Again, the heavy dashed line represents the outer locus of points such that bilateral trade between all countries occurs for all trading arrangements. Other lines

[9] Both ER and FT also dominate NA for a majority of the relevant parameter space. However, if cost asymmetries are high, firms in N will prefer NA to ER and FT, since NA offers more protection from low cost import competition. However, NA is still dominated by the IR as illustrated in Figure 3.

Figure 4: Preference conditions - North and South



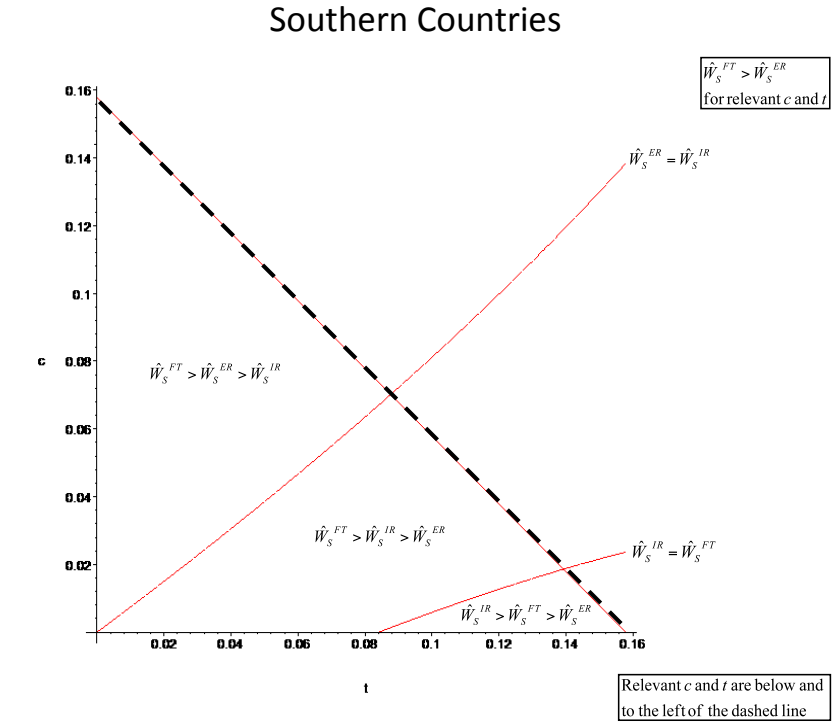
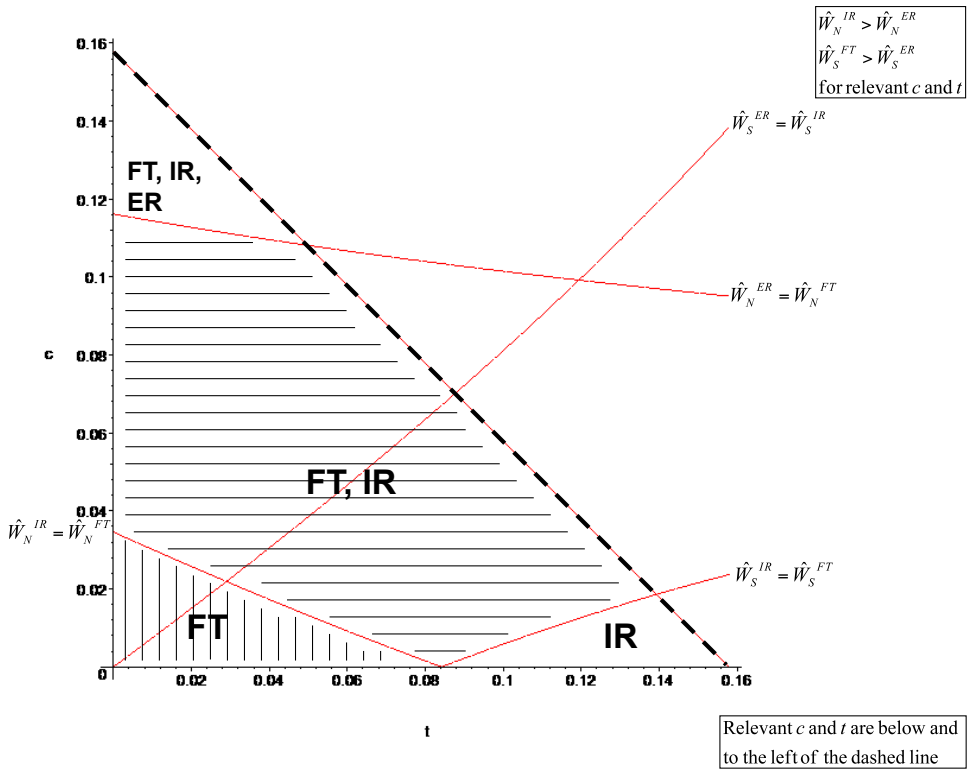


Figure 5: Undominated Trading Arrangements



represent the indifference points between the three remaining undominated trading arrangements, FT, IR, and ER. As an analytical benchmark, the result from Proposition 2 is on the horizontal axis, and the result from Proposition 3 is on the vertical axis.

For a majority of the range of tariffs and regional cost differences such that trade occurs for all equilibrium agreement structures, no one equilibrium outcome is dominant. That is, with multiple trading arrangements that remain undominated, countries cannot agree on a unique equilibrium coalition. This *disagreement* outcome is more likely when cost differences are high, where as cost differences increase, the number of undominated trading arrangements also tends to increase. In addition, the fraction of the parameter space which supports multiple undominated agreement structures also increases. When cost differences are low, as in Proposition 2, global free trade is optimal when trade costs are relatively low, and regional agreements are optimal when trade costs are relatively high.

The intuition for this result is that as cost differences between regions increase, high cost countries wish to join in a regional agreement to act as a buffer from low cost import competition. Of course, the incentive of low cost countries is exactly the opposite, as they can gain substantially from unrestricted market access to countries with less competitive firms. Eventually, as this cost difference becomes large, this dichotomy becomes prohibitive, and there exist multiple undominated agreement structures within the core. As there exists no unique preferred agreement structure, I

conclude that there exists "trade disagreement".

7 Extensions

The results in the previous section highlight the difficulty in forming trading coalitions which benefit all countries. In this section, I offer a number of extensions to the model to evaluate how regional agreement prohibition, transfers, and political influence affect the scope for agreement in tariff negotiations.

Building blocks or stumbling blocks?

A focal point in the debate over regionalism and multilateralism is whether preferential agreements are a help or a hindrance to the process of multilateral liberalization. As coined by Bhagwati, are regional agreements "stumbling blocks" or "building blocks" toward global free trade?

Many authors have provided competing viewpoints regarding how preferential agreements affect the incentives for multilateral cooperation. As stated earlier, Aghion, Antras, and Helpman (2007) have addressed this within an extensive-form model of agreement formation. A leader country chooses the multilateral or preferential track, and follower countries make decisions on whether to accept agreements or not. They show that the building or stumbling nature of preferential agreements is based on the form of coalition externalities. Saggi and Yildiz (2007) offer another approach, showing that the answer to this question depends on whether countries are symmetric. Interestingly, if countries are asymmetric, there can exist a strong building bloc effect of preferential agreements.

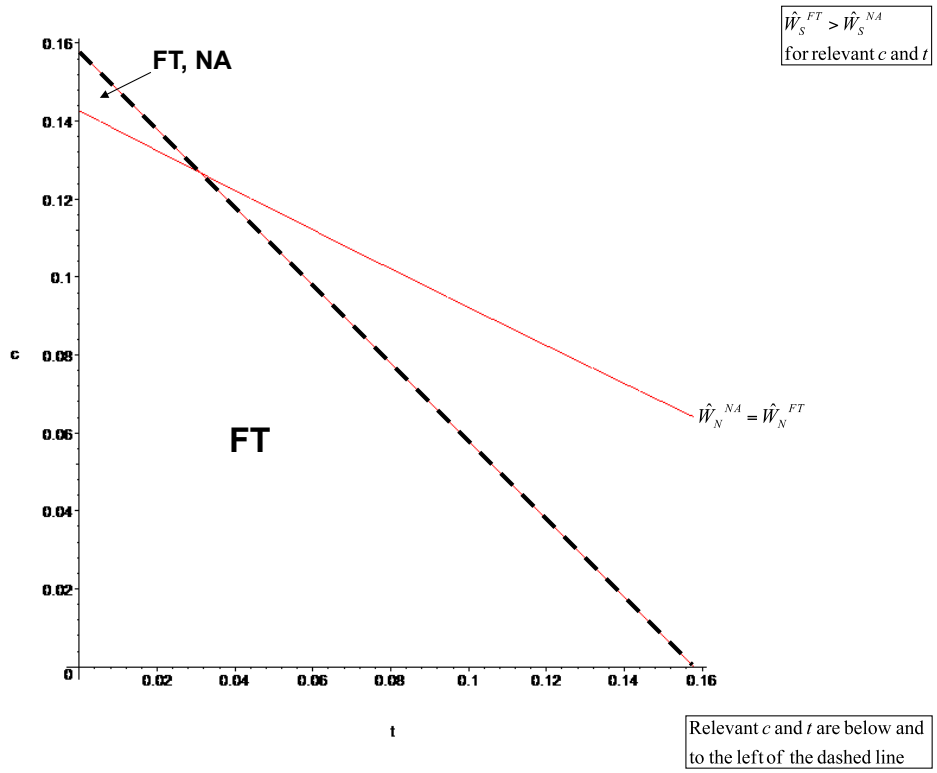
In contrast with Aghion, Antras, and Helpman (2007), I do not assume any timing or dynamic structure of how trade agreements are formed. Further, different from Saggi and Yildiz (2007), there are no non-cooperative deviations. Despite these differences, the model is still equipped for a simple examination of the building or stumbling bloc aspects of preferential agreements. The results from doing so are illustrated in Figure 6, where I have noted the equilibrium coalitions which remain undominated as a function of c and t when IR and ER are prohibited. Clearly, multilateral free trade, FT, is the dominant outcome for a majority of the relevant parameter space. For a small region in which trade costs are small and cost asymmetries are large, NA also remains undominated. The intuition is that for large cost asymmetries, high cost countries prefer no agreements, as otherwise their domestic firms will be disproportionately hurt at home and abroad by southern firms which are significantly more productive. Overall, the results in Figure 6 suggest that preferential agreements, and in particular those with regional trading partners, are a stumbling block in an effort to achieve multilateral free trade.[10]

Inter-agreement Transfers

Can transfers help facilitate cooperation and an equilibrium trading arrangement which maximizes world welfare? The short answer is yes, though the long answer depends on whether the transfers are intra-agreement or inter-agreement. This is not a new question, where Kemp and Wan (1976)

[10] However, this interpretation ignores an important point: free trade *should not* be the goal if intraregional agreements improve welfare for all parties, which is the case in Figure 5. Thus, within this framework, IR is a stumbling block when it should be.

Figure 6: Undominated Trade Agreements - Prohibited Preferential Agreements



show that transfers can always induce an enlargement of a given customs union. Riezman (1985) extends this question to a three-country model of customs union formation. Indeed, Riezman (1985) discusses how transfers might prevent countries from agreeing to free trade within a cooperative model. In this section, I build on his work using my four-country model, and the aforementioned two "types" of transfers.

Allowing for inter-agreement transfers greatly simplifies the cooperative solution of the model. To see this, suppose that over some parameter space, N countries prefer IR but S countries prefer FT. To convince N countries to agree to FT, S countries must be able to offer a transfer greater than or equal to the welfare loss in moving to free trade. Such a transfer will exist if:

$$\widehat{W}_S^{FT} - \widehat{W}_S^{IR} > \widehat{W}_N^{IR} - \widehat{W}_N^{FT}$$

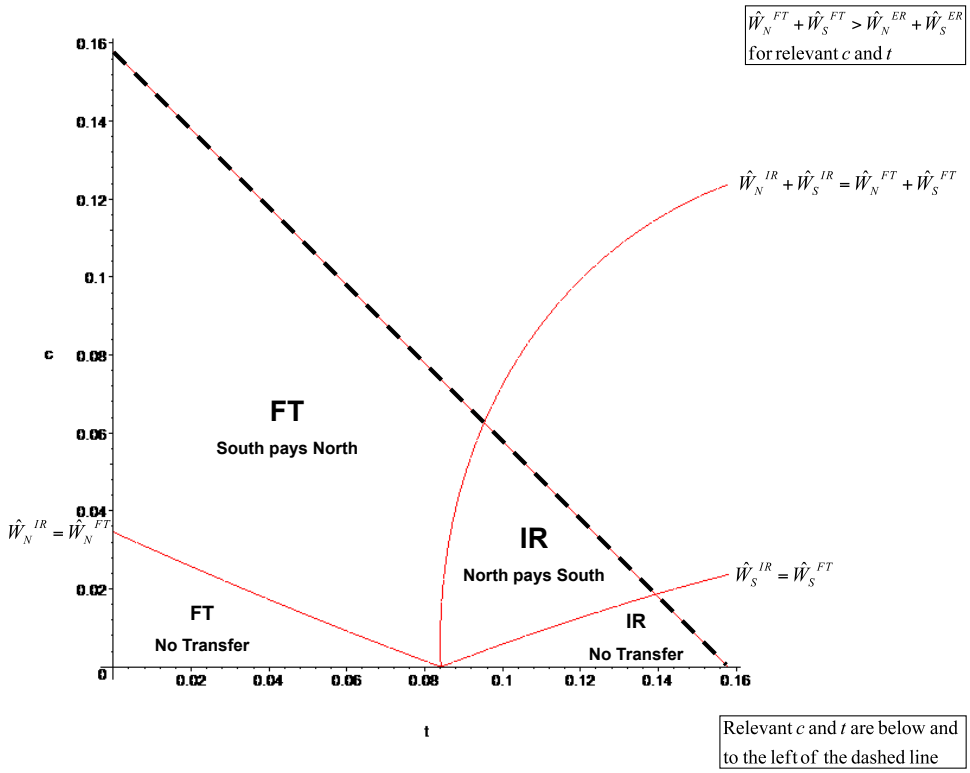
This is simplified as:

$$\widehat{W}_S^{FT} + \widehat{W}_N^{FT} > \widehat{W}_S^{IR} + \widehat{W}_N^{IR}$$

That is, total welfare under free trade must be higher than the total welfare under no agreements. This will be the metric by which the global optimum trading structure will be chosen, allowing for transfers to facilitate such a trading structure.

Notably, models of the GATT/WTO, such as Bagwell and Staiger (2003), measure welfare glob-

Figure 7: Undominated Trade Agreements with Transfers



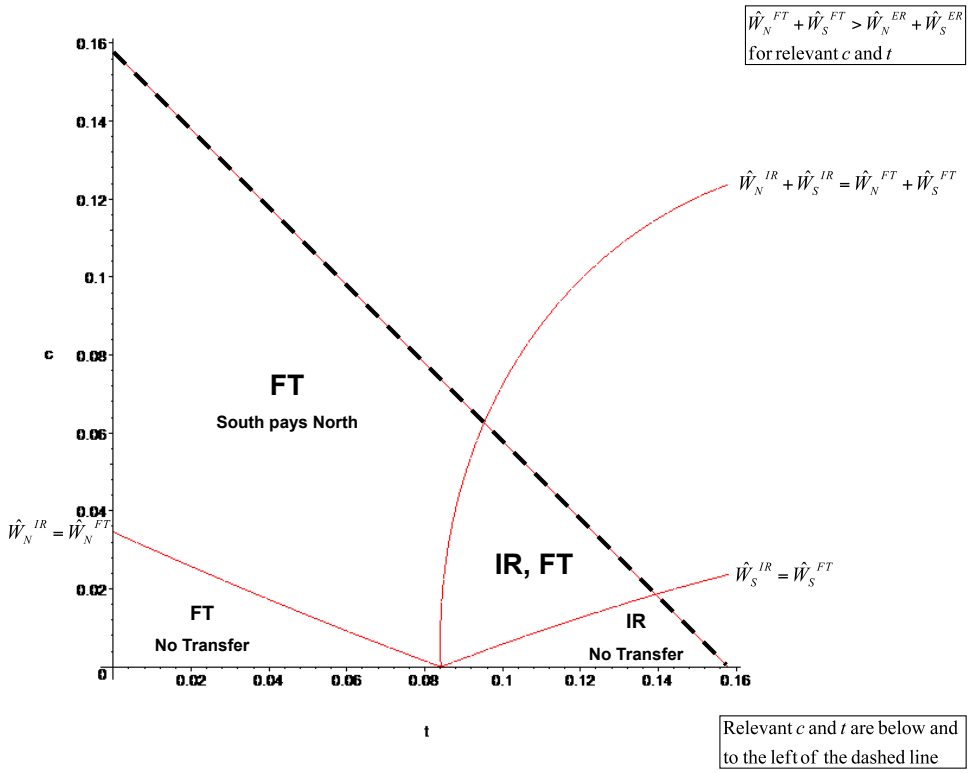
ally to evaluate whether optimal tariffs are efficient, and hence whether an efficient locus of policy choices can be achieved via GATT/WTO negotiations. However, in their model, countries only liberalize to the extent that they are willing, and transfers are not a part of this calculation. Despite this, the principle of reciprocity, which balances the value of concessions, neutralizes international cost shifting and yields optimal tariffs that induce global efficiency in an environment of symmetric countries. In contrast, when countries are sufficiently asymmetric, negotiated tariffs may fall short of those that are globally efficient.[11] Thus, in their model, an unanswered question is the extent to which transfers are necessary to facilitate global efficiency in an environment of asymmetric countries. Although I use a different modeling framework, my model is well suited to evaluate the extent to which transfers are necessary to induce global efficiency when countries are asymmetric.

To begin the analysis, note that since both parties prefer IR to NA in Figure 3, it must also be the case the the total welfare accrued within IR must be greater than the total welfare accrued in NA. The preference conditions for other arrangements are illustrated in Figure 7. For the complete set of relevant parameters, the welfare accrued under free trade, FT, is greater than the welfare under a extraregional agreements, ER. Thus, I have omitted any loci related to ER.

However, there is a relevant choice between IR and FT. In Figure 7, I have labeled the global

[11] See Bagwell and Staiger (2003), page 62, for a case where asymmetric countries fail to negotiate tariffs at the globally efficient level.

Figure 8: Undominated Trade Agreements - Within Agreement Transfers



preference locus, $\widehat{W}_N^{IR} + \widehat{W}_S^{IR} = \widehat{W}_N^{FT} + \widehat{W}_S^{FT}$, and those specific to each region, $\widehat{W}_N^{IR} = \widehat{W}_N^{FT}$ and $\widehat{W}_S^{IR} = \widehat{W}_S^{FT}$. Generally, for relatively high trade costs, IR is the preferred outcome. Conversely, for low trade costs, free trade is the preferred outcome. As cost asymmetries increase, these outcomes can only be reached via a transfer scheme. When IR occurs, the North must pay the South to compensate for their loss in market access relative to FT. In contrast, when FT occurs, the South must pay the North to compensate for the injured domestic sector relative to IR.

Intra-agreement Transfers

Next, I assume that transfers may occur, but only between countries that are party to the same trade agreement. For example, if FT provides higher world welfare than IR, transfers can occur within FT to prevent IR from occurring. However, if IR provides higher world welfare, it will only occur if both North and South countries are better off under IR, since transfers are prohibited between North and South when they are party to different agreements. With this slight modification, the equilibrium trade agreement structures are illustrated in Figure 8.

In Figure 8, there exists a region in which there is disagreement between IR and FT. Here, IR provides higher world welfare, but the South prefers FT. Since transfers may not occur between North and South when they are party to different agreements (IR includes two regional agreements),

the North cannot compensate the South via transfers. Further, since FT does not maximize world welfare, the South cannot compensate the North such that FT occurs. Thus, when allowing for transfers only within coalitions, we have equilibrium trade disagreement for high trade costs and relatively low cost differences.

Allowing for transfers within coalitions provides another comparison to the modeling strategy of Aghion, Antras and Helpman (2007), where equilibria are determined by comparing joint welfare within trade agreements (ie intra-agreement transfers are permissible). Given the extensive form structure of the game and complete information, there always exists a subgame perfect equilibrium. In contrast, I assume no extensive form representation of the agreement formation game, instead viewing trade agreements as cooperative coalitions with transfers within agreements. Without the extensive form representation, I find that disagreement occurs for a non-trivial portion of the parameter space even when transfers can occur within coalitions.

Political Influence

A focal point in the debate over regionalism is the degree to which political influence plays into the structure of trade agreements. As a final point of discussion, consider an extension of the framework described in section two in which countries only care about profits of their domestic firms, and not tariff revenue or consumer welfare. This is similar to the approach taken in Krishna (1998), where political influence is assumed to be absolute, with no consideration of consumer welfare or tariff revenue.

Within the framework outlined in section three, political influence has a very stark impact on tariff setting. That is, when profits are the only component of the welfare function, governments have incentive to increase tariffs to the level of non-prohibitive trade for all countries subject to the tariff. The reason is two-fold. First, governments don't care about the negative effect of tariffs on consumer surplus or any potential revenues generated. Further, as the model itself is not general equilibrium, there are no effects of import tariffs on export profits. Thus, when setting any tariff noncooperatively, governments have incentive to levy a prohibitive tariff outside of any agreement to which they are a party.

While this feature might seem limiting in terms of discussing the scope for agreement within the agreement formation game, it actually focuses the analysis on a key incentive for which many firms prefer high tariffs: domestic market power. Indeed, the forthcoming results will be similar in flavor to the results in Horn and Persson (2001), where trade costs and tariffs help support domestic monopolization through mergers. I will return to this point shortly.

First, I define profit functions for each firm under each agreement structure. For NA, profits (welfare) are written as:

$$P_N^{NA} = \frac{(A-c)^2}{4b}$$
$$P_S^{NA} = \frac{A^2}{4b}$$

Here, I have labeled profits from the political agreement formation game as P_N^k and P_S^k, where k represents agreement structure k. In NA, there are no agreements to facilitate free trade, and as such, countries raise tariffs to promote domestic monopolization. Hence, P_N^{NA} represents monopoly profits subject to cost, c, and P_S^{NA} monopoly profits subject to cost 0.

Next, the profits for the extraregional agreement, ER, are written as:

$$P_N^{ER} = \frac{(A+t-2c)^2}{9b} + \frac{(A-2c-2t)^2}{9b}$$
$$P_S^{ER} = \frac{(A+c+t)^2}{9b} + \frac{(A+c-2t)^2}{9b}$$

Here, since there is a trade agreement between one southern country and one northern country, firms can earn profits in two markets. In P_N^{ER}, the first term represents profits at home, and the second export profits in the southern market. In P_S^{ER}, the first term represents profits at home, and the second export profits in the northern market.

For IR, profits are written as:

$$P_N^{IR} = \frac{2(A-c)^2}{9b}$$
$$P_S^{IR} = \frac{2A^2}{9b}$$

Similar to NA, there is no trade between regions under IR. Hence, t does not factor into equilibrium firm-level profits. In the north, firms compete within-region as a duopoly, earning $\frac{(A-c)^2}{9b}$ in each northern market. In the south, firms earn $\frac{A^2}{9b}$ in each southern market.

Finally, for FT, profits are written as:

$$P_N^{FT} = \frac{2(A+2t-3c)^2}{25} + \frac{2(A-3c-3t)^2}{25}$$
$$P_S^{FT} = \frac{2(A+2c+2t)^2}{25} + \frac{2(A+2c-3t)^2}{25}$$

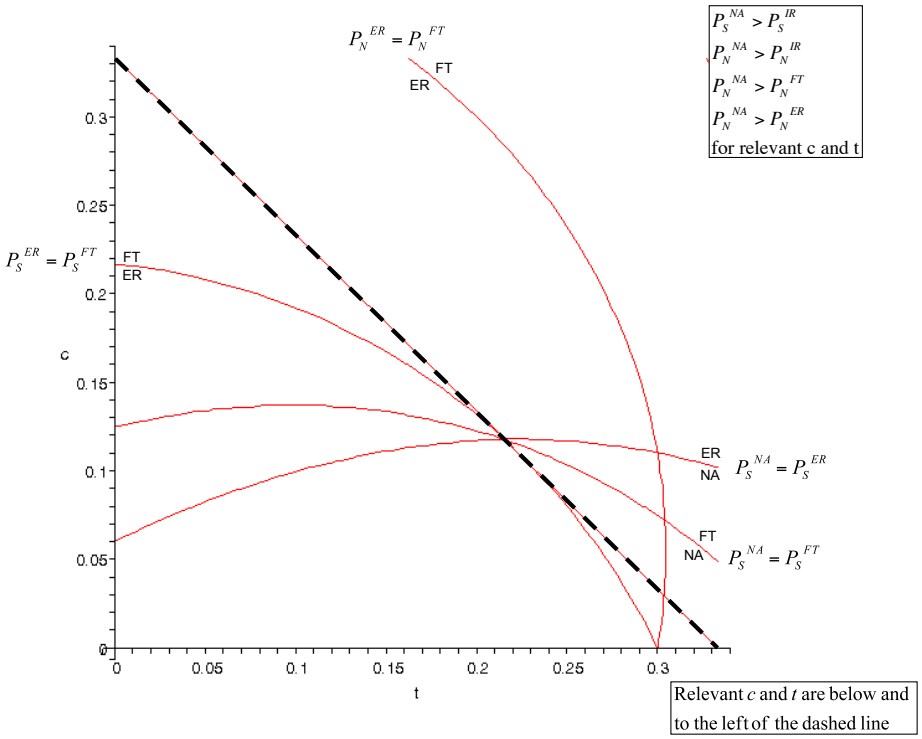
Under FT, no tariffs are set and firms compete as a quadropoly in each market. The first term in P_N^{FT} and P_S^{FT} are profits earned within their respective home region, and the second term profits earned in the opposing region.

The indifference loci between each agreement structure for north and south are illustrated in Figure 9, with the preferred agreement structure being labeled above and below each curve.[12] The first point to make is that IR is always dominated by NA. This is the exact opposite result to the case where governments care about consumer surplus, tariff revenue, and producer surplus. The intuition is that under IR firms earn duopoly profits in two identical markets, and under NA, firms earn monopoly profits in one of those markets. Obviously, the aggregate profits under the latter will be larger.

The rest of the indifference loci represent a trade-off between domestic monopolization through tariff setting vs enhanced market access through trade agreements. As in the previous section, the north is hesitant to sign a trade agreement with the south. Indeed, NA is always preferred to FT, and as such, FT will never be the dominant outcome. In contrast with the north, the south is willing to sign a trade agreement as their firms have a competitive advantage relative to firms in the north. However, when this competitive advantage is small (c is small), the south prefers NA to FT as profits will be higher under full monopolization via NA.

[12] For trade to occur under all agreement structures, I restrict the parameter space to $c + t < 1/3$.

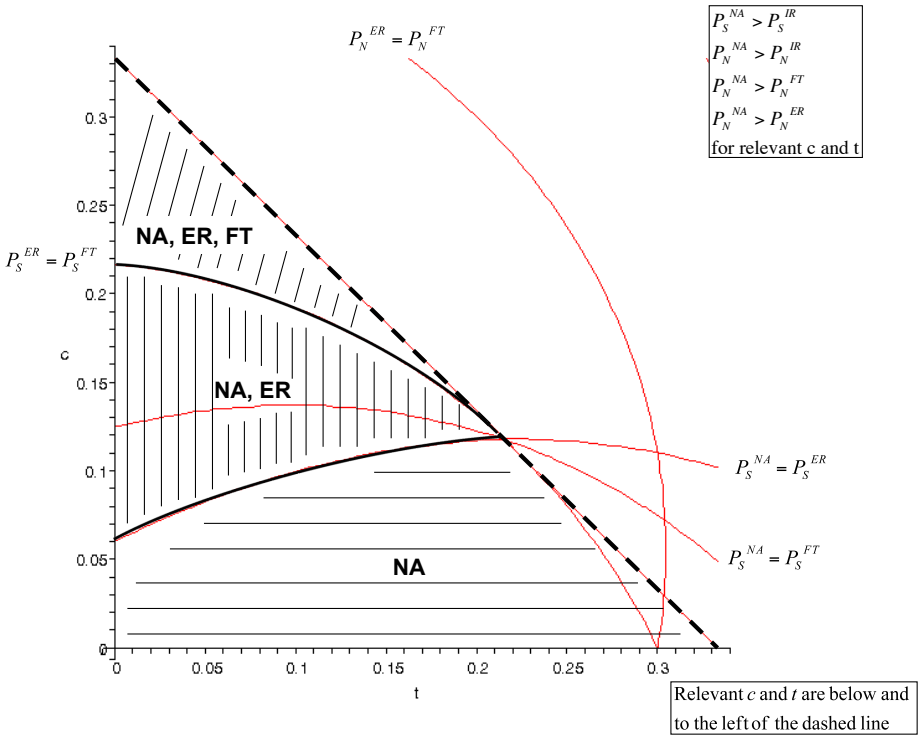
Figure 9: Preference Conditions - Profits as Welfare



When comparing all indifference loci, and evaluating preference conditions at $c = t = 0$, the equilibrium trade agreement structures for the political model are presented in Figure 10. Here, the results are in stark contrast with the baseline calculations when including tariff revenue and consumer surplus in the welfare function. Under a standard welfare function, countries usually prefer some sort of liberalization, but disagree over the optimal way in which liberalization should occur. In contrast, in Figure 10, for roughly half of the parameter space in which trade occurs for all outcomes, countries prefer "No Agreement" (NA) over all other options. For a given level of cost differences between regions, a unique dominant agreement structure is more likely when trade costs are high. However, again, there exists a threshold level of cost differences above which the core is non-unique, and hence countries cannot agree regarding the equilibrium agreement structure.

As mentioned earlier, the results in Figure 10 are very similar to those in Horn and Persson (2001), where trade costs create an incentive for domestic monopolization rather than foreign investment. Indeed, as trade costs become relatively large, this provides additional market power for each firm in their respective region. Thus, when welfare is only a function of firm-level profits, and profits are increasing in market power, any trade agreements become dominated, as they would serve to reduce the market power of each domestic monopolist. However, once again, as cost differences increase, the low cost region is willing to forego some of its own market power for a large improvement in market access in the opposing region. Since they are relatively efficient, the lost

Figure 10: Undominated Agreements - Profits as Welfare



market power is relatively small compared with the large gains in market access. In equilibrium, when trade costs are low and cost differences high, there is disagreement over whether to liberalize at all.

8 Conclusion

Overall, the results of model show that when characterizing a trade agreements game between heterogeneous and distant regions using the core solution concept, there exists a large and persistent parameter space in which countries cannot agree on a unique, welfare maximizing agreement equilibrium. Intra-agreement transfers, regional agreement prohibition, and political influence each can serve to alleviate, but not eliminate, the regions of trade disagreement. Inter-agreement transfers eliminate trade disagreement, though as evident in sections six and seven, these transfers are required for a majority of the parameter space in which trade occurs.

One interesting extension of the model would be to examine how the WTO allows FTAs to proceed, and whether that process promotes an efficient structure of trade agreements. Viewing the results of this paper with those in Aghion, Antras, and Helpman (2007), it is clear that even with intra-agreement transfers, the equilibrium of the agreement formation stage depends crucially on timing and/or initial conditions. While the permissibility of regional agreements is an issue

that likely has little policy flexibility, perhaps the WTO can/should design a scheme which allows countries to lobby for or against such agreements. This is normative issue which may be fruitful for future research.

References

[1] Aghion, Philippe, Pol Antras and Elhanan Helpman, "Negotiating free trade", *Journal of International Economics,* Vol. 73, No. 1, (Sep. 2007), pp. 1-30.

[2] Bhagwati, Jagdish; "Regionalism versus Multilateralism", *The World Economy*, Vol. 15, No. 5 (1992) , pp. 535–556

[3] Bagwell, Kyle and Robert Staiger (2003), *The Economics of the World Trading System*, MIT Press.

[4] Brander, James A. and Paul Krugman. "A 'Reciprocal Dumping' Model of International Trade." *Journal of International Economics*, Vol. 15, No. 3/4, (Nov. 1983), pp. 313-321.

[5] Burbidge, John B.; James A. DePater; Gordon M. Myers; Abhijit Sengupta, "A Coalition-Formation Approach to Equilibrium Federations and Trading Blocs", *The American Economic Review*, Vol. 87, No. 5. (Dec., 1997), pp. 940-956.

[6] Frankel, Jeffrey; Ernesto, Stein and Shang-jin Wei, "Trading blocs and the Americas: The natural, the unnatural, and the super-natural", *Journal of Development Economics*, Vol. 47 (1995), pp. 61-95.

[7] Freund, Caroline, "Multilateralsim and the endogenous formation of preferential agreements", *Journal of International Economics*, Vol. 52 (2000), pp. 310-335.

[8] Furusawa, Taiji, and Hideo Konishi, "Free trade networks", *Journal of International Economics*, Vol. 72 (2007), pp. 310-335.

[9] Horn, Henrik and Lars Persson (2001), "The equilibrium ownership of an international oligopoly", *Journal of International Economics*, Vol 53-2, pp. 307-333

[10] Kemp, Murray and Henry Wan (1976), "An Elementary Proposition Concerning the Formation of Customs Unions", *Journal of International Economics*, Vol 6-1, pp 95-97.

[11] Kennan, John and Raymond Riezmann (1990), "Optimal Tariff Equilibria with Customs Unions", *Canadian Journal of Economics*, Vol 23-1, pp. 70–83.

[12] Krishna, Pravin, "Regionalism And Multilateralism: A Political Economy Approach", *The Quarterly Journal of Economics* (Feb. 1998), pp. 227-251

[13] Meade, James E., *The Theory of Customs Unions*, Amsterdam, 1955.

[14] Melatos, Mark and Alan Woodland (2007), "Endogenous Trade Bloc Formation in an Asymmetric World", *European Economic Review*, Vol. 51-4, pp. 901-924.

[15] Panagariya, Arvind, "Preferential Trade Liberalization: The Traditional Theory and New Developments", *Journal of Economic Literature*, Vol. 38, No. 2 (Jun., 2000), pp. 287-331

[16] Panagariya, Arvind, "Do Transportation Costs Justify Regional Preferential Trading Arrangements?: No", *Weltwritschaftliches*, Vol. 134, No. 2 (Jun., 1998), pp. 280-301

[17] Riezmann, Raymond (1985), "Customs Unions and the Core", *Journal of International Economics*, Vol 19, pp. 355-365.

[18] Saggi, Kamil, "Preferential trade Agreements and Multilateral Tariff Cooperation", International Economic Review, Vol. 47, No. 1, (Feb. 2006), pp, 29-57.

[19] Saggi, Kamil and Halis Murat Yildiz (2007), "Bilateral trade agreements and the feasibility of multilateral free trade", mimeo Southern Methodist University.

[20] Viner, Jacob, *The Customs Union Issue*, New York 1950.

[21] Wonnacott, Paul, and Ronald Wonnacott, "is Unilateral Tariff Reduction Preferable to a Customs Union? The Curious Case of the Missing Foreign Tariffs", *American Economic Review*, Vol. 71, No. 4 (Sept. 1981), pp. 704-714

A Proposition 1

Recall the formulas for Δ_x and Δ_y:

$$\begin{aligned}\Delta_x &= W_x^F - W_x^N(\widehat{\tau}_x, \widehat{\tau}_y) = \frac{7A^2 - 56Ac + 40c^2}{162b} \\ \Delta_y &= W_y^F - W_y^N(\widehat{\tau}_x, \widehat{\tau}_y) = \frac{7A^2 + 42Ac - 9c^2}{162b}\end{aligned}$$

Using the quadratic formula, $\Delta_x = 0$ when $c = \{-0.06456A, 0.6733A\}$. Hence, since $\Delta_x(c = 0) > 0$, the north prefers F when $c \in [0, A/4]$ - all relevant values of c. As for the south, $\Delta_y = 0$ when $c = \{-2.06066A, 0.06066A\}$. Hence, the south prefers F when $c \in [0, 0.06066A)$, and N when $c \in [0.06066A, A/4]$. Overall, for $c \in [0, 0.06066A)$ F remains undominated. If $c \in [0.06066A, A/4]$, F and N remain undominated.

B Welfare functions

Substituting $\widehat{\tau}_1^{NA}$ into W_1^{NA}, equilibrium welfare is written as:

$$\widehat{W}_N^{NA} \equiv \widehat{W}_1^{NA} = \widehat{W}_2^{NA} = \frac{945A^2 - 684At - 2046Ac + 2364t^2 + 1276tc + 3045c^2}{2178b} \tag{4}$$

Substituting $\widehat{\tau}_3$ into W_3^{NA}, equilibrium welfare is written as:

$$\widehat{W}_S^{NA} \equiv \widehat{W}_3^{NA} = \widehat{W}_4^{NA} = \frac{945A^2 - 684At + 156Ac + 2364t^2 - 592tc + 1944c^2}{2178b} \tag{5}$$

Substituting $\widehat{\tau}_1^{IR}$into W_1^{IR}, equilibrium welfare is written as:

$$\widehat{W}_N^{IR} \equiv \widehat{W}_1^{IR} = \widehat{W}_2^{IR} = \frac{15A^2 - 10At - 30Ac + 23t^2 - 6tc + 33c^2}{32b} \tag{6}$$

Substituting $\widehat{\tau}_3$ into W_3^{IR}, equilibrium welfare is written as:

$$\widehat{W}_S^{IR} \equiv \widehat{W}_3^{IR} = \widehat{W}_4^{IR} = \frac{15A^2 - 10At + 23t^2 + 16tc + 18c^2}{32b} \tag{7}$$

Substituting $\widehat{\tau}_1^{ER}$ into W_1^{ER}, equilibrium welfare is written as:

$$\widehat{W}_N^{ER} \equiv \widehat{W}_1^{ER} = \widehat{W}_2^{ER} = \frac{540A^2 - 564At + 1295t^2 - 1356Ac + 994tc + 1691c^2}{1152b} \tag{8}$$

Substituting $\widehat{\tau}_3^{ER}$ into W_3^{ER}, equilibrium welfare is written as:

$$\widehat{W}_S^{ER} \equiv \widehat{W}_3^{ER} = \widehat{W}_4^{ER} = \frac{540A^2 - 564At + 1295t^2 + 276Ac - 430tc + 875c^2}{1152b} \tag{9}$$

Finally, welfare under free trade is written as:

$$\widehat{W}_N^{FT} \equiv \widehat{W}_1^{FT} = \widehat{W}_2^{FT} = \frac{2\left(6A^2 - 6At + 14t^2 - 16Ac + 8tc + 19c^2\right)}{25b} \tag{10}$$

and

$$\widehat{W}_S^{FT} \equiv \widehat{W}_3^{FT} = \widehat{W}_4^{FT} = \frac{2\left(6A^2 - 6At + 14t^2 + 4Ac - 2tc + 9c^2\right)}{25b} \tag{11}$$

C Four country model: Proofs of Lemmas and Propositions

Proposition 2

To prove proposition 2, I will first establish that FT is preferred to ER and NA for all t. Comparing the welfare functions in (4) and (10) at $c = 0$, countries prefer free trade to no agreement if:

$$\widehat{W}_N^{FT} - \widehat{W}_N^{NA} = \widehat{W}_S^{FT} - \widehat{W}_S^{NA} = \frac{837A^2 - 3012At + 628t^2}{18150b} > 0$$

Clearly, at $t = 0$, countries prefer FT to NA. For larger values of t, note that $837A^2 - 3012At + 628t^2 = 0$ if $t = \{\frac{93}{314}A, \frac{9}{2}A\}$. Using (3), the values must satisfy $t < \frac{3}{19}A$, which they do not. Thus, countries always prefer FT to NA.

Next, using (8) and (10), FT is preferred to ER if:

$$\widehat{W}_N^{FT} - \widehat{W}_N^{ER} = \widehat{W}_S^{FT} - \widehat{W}_S^{ER} = \frac{(6A + 7t)(54A - 17t)}{28800b} > 0$$

This will clearly be satisfied for $t < \frac{3}{19}A\left(< \frac{54}{17}A\right)$. Thus, FT is preferred to ER over all relevant values of t

Finally, Proposition two is proven by evaluating welfare between FT and IR. Using (6) and (10) at $c = 0$, FT is preferred to IR if:

$$\widehat{W}_N^{FT} - \widehat{W}_N^{IR} = \widehat{W}_S^{FT} - \widehat{W}_S^{IR} = \frac{(A - 3t)(9A - 107t)}{800b} > 0$$

This is positive if $t < \frac{9}{107}A$, which is clearly within the region specified by $t < \frac{3}{19}A$. The other zero, $t = \frac{1}{3}$, is not. Thus, FT is preferred to IR if $t < \frac{9}{107}A$, and IR preferred to FT otherwise.

Proposition 3

To prove Proposition 3, I will first establish preference relationships between NA and IR for countries in the north and south. Precisely, IR is preferred to NA for countries in N if:

$$\widehat{W}_N^{IR} - \widehat{W}_N^{NA} = \frac{1215A^2 + 66Ac - 12783c^2}{34848b} > 0$$

By using the quadratic formula, the LHS is greater than zero if $c \in (-0.30573A, 0.31089A)$. The relevant range of parameter values, $c \in \left(0, \frac{3}{19}A\right)$ from (3), is within this region. Region S countries

prefer IR to NA if the following condition holds:

$$\widehat{W}_S^{IR} - \widehat{W}_S^{NA} = \frac{1215A^2 - 2496Ac - 11502c^2}{54450b} > 0$$

By using the quadratic formula, the LHS is greater than zero if $c \in (-0.45115A, 0.23414A)$. The relevant range of parameter values is also within this region. Hence, IR dominates NA for all relevant values of c when $t = 0$.

Next, I show that there exists no dominance relationship between IR and ER. To see this, note that IR is preferred to ER for countries in N if:

$$\widehat{W}_N^{IR} - \widehat{W}_N^{ER} = \frac{(276A - 503c)\,c}{1152b} > 0$$

This is satisfied for $c \in \left(0, \frac{3}{19}A\right)$. For firms in S, IR is preferred to ER if:

$$\widehat{W}_S^{IR} - \widehat{W}_S^{ER} = \frac{-\,(227c + 276A)\,c}{1152b} > 0$$

This will clearly not be satisfied. Thus, there is no dominance relationship between IR and ER.

Next, I compare FT and ER, where FT is preferred to ER for countries in N if:

$$\widehat{W}_N^{FT} - \widehat{W}_N^{ER} = \frac{324A^2 - 2964Ac + 1501c^2}{28800b} > 0$$

The RHS is greater than zero if $c < 0.11614A$ or $c > 1.85854A$. The first is the only relevant condition within the range $c \in \left(0, \frac{3}{19}A\right)$. For countries in S, FT is preferred to ER if:

$$\widehat{W}_S^{FT} - \widehat{W}_S^{ER} = \frac{324A^2 - 2964Ac + 1501c^2}{28800b} > 0$$

The RHS is greater than zero if $c \in (-0.1314A, 2.1648A)$. This is clearly within $c \in \left(0, \frac{3}{19}A\right)$, and as such, countries in S always prefer FT to ER.

Overall, for $t = 0$, over the region $c \in (0, 0.11614A)$, FT dominates ER. For $c \in \left(0.11614A, \frac{3}{19}A\right)$, there is no dominance relationship between FT and ER.

Finally, I establish the preference relationships between FT and IR. To begin, note that for countries in N, FT is preferred to IR if:

$$\widehat{W}_N^{FT} - \widehat{W}_N^{IR} = \frac{9A^2 - 274Ac + 391c^2}{800b} > 0$$

The RHS is greater than zero if $c < 0.03455A$ or $c > 0.66622A$. The first condition is clearly in the region of relevant c, and the second is not.

For countries in S, FT is preferred to IR if:

$$\widehat{W}_S^{FT} - \widehat{W}_S^{IR} = \frac{9A^2 + 256Ac + 126c^2}{800b} > 0$$

This is always satisfied. Thus, overall, FT dominates IR if $c < 0.03455A$. Otherwise, there is no dominance relationship between FT and IR.

Putting these preference relationship together yields Proposition 3. Precisely, labeling $c_1 = 0.03455A$ and $c_2 = 0.11614A$, we have the following characterization of the core of the agreement formation game:

For	$c \in [0, c_1)$	FT remains undominated
	$c \in [c_1, c_2)$	FT and IR remain undominated
	$c \in \left[c_2, \frac{3}{19}A\right)$	FT, IR and ER remain undominated