

Negotiating Free Trade

by

Philippe Aghion, Pol Antràs and Elhanan Helpman*

December 18, 2006

Abstract

We develop a dynamic bargaining model in which a leading country endogenously decides whether to sequentially negotiate free trade agreements with subsets of countries or engage in simultaneous multilateral bargaining with all countries at once. We show how the structure of coalition externalities shapes the choice between sequential and multilateral bargaining, and we identify circumstances in which the grand coalition is the equilibrium outcome, leading to worldwide free trade. A model of international trade is then used to illustrate equilibrium outcomes and how they depend on the structure of trade and protection. Global free trade is not achieved when the political-economy motive for protection is sufficiently large. Furthermore, the model generates both “building bloc” and “stumbling bloc” effects of preferential trade agreements. In particular, we describe an equilibrium in which global free trade is attained only when preferential trade agreements are permitted to form (a building bloc effect), and an equilibrium in which global free trade is attained only when preferential trade agreements are forbidden (a stumbling bloc effect). The analysis identifies conditions under which each of these outcomes emerges.

JEL Classification: C78, F13

Keywords: bargaining, trade agreements, regionalism, multilateralism, coalition externalities.

*Aghion: Department of Economics, Harvard University, Littauer 231, Cambridge, MA 02138, National Bureau of Economic Research, and Centre for Economic Policy Research (e-mail: paghion@fas.harvard.edu); Antràs (corresponding author): Department of Economics, Harvard University, Littauer 319, Cambridge, MA 02138, National Bureau of Economic Research, and Centre for Economic Policy Research (e-mail: pantras@fas.harvard.edu), Tel.: +1 617 495 1236, Fax: +1 617 495 8570; Helpman: Department of Economics, Harvard University, Littauer 229, Cambridge, MA 02138, National Bureau of Economic Research, Centre for Economic Policy Research, and Canadian Institute for Advanced Research (e-mail: ehelman@harvard.edu). We thank Ran Spiegler for a useful discussion, Gene Grossman, Davin Chor, Bob Staiger and two referees for comments on an earlier draft, Jane Trahan for editorial assistance, and Helpman thanks the National Science Foundation for financial support.

1 Introduction

One of the most debated issues in international economics is whether regionalism or multilateralism is the most effective strategy for achieving global free trade. According to Bhagwati (1993), the first wave of regionalism took place in the 1960s, and it failed to spread because the U.S. supported a multilateral approach. But the U.S. changed positions, and — starting with the 1980s — has favored regional trade agreements. This led to a second wave of regionalism, which brought about a multitude of such agreements. The gradual enlargement of the European Union, the U.S.-Canada Free Trade Agreement, NAFTA and MERCOSUR, are examples of this trend. Between 1958 and March 2004, the GATT/WTO secretariat received notification of 203 agreements.¹ The recent stalling of the Doha round further suggests that multilateralism, which was the dominant force towards free trade in the first few decades after World War II, is falling out of fashion.

Economists disagree on whether preferential trade agreements are “building blocs” that facilitate the attainment of global free trade, or “stumbling blocs” that derail the process of trade liberalization. The latter view has been forcefully promoted by Bhagwati (1991, 1993), who coined these terms.² In this view, even when preferential trade agreements generate static welfare gains they reduce the incentives to seek further trade liberalization. The importance of this “dynamic path” question was clearly laid out by Bhagwati (1993) and Krugman (1993). The latter also showed that in some circumstances welfare reaches a minimum when the world consists of two or three customs unions.³ The welfare consequences of stalled multilateralism, caused by the rise of regionalism, could therefore be significant.

Other economists, such as Summers (1991), think that preferential trade agreements do not impede global free trade. They argue that partial trade liberalization is better than none, and that the consolidation of a large number of countries into a small number of trading blocs facilitates multilateral negotiations. And Baldwin (1996) argues that a deepening of integration between a subset of countries may raise the incentives of outside countries to seek accession to the free trade area. Under these circumstances preferential trade agreements encourage further trade liberalization and the expansion of the free trading blocs.⁴

¹Many preferential trade agreements are *not* regional. The U.S.-Israel free trade agreement is a notable example. Following Bhagwati (1993, p. 22), we use a terminology in which ‘regionalism’ is “...defined broadly as preferential trade agreements among a subset of nations.” That is, we downplay the geographical nature of preferential trade agreements and emphasize instead the fact that they constitute an agreement between a subset of countries.

²See Panagariya (2000) for a recent survey of this literature.

³See Deardorff and Stern (1994), Bond and Syropoulos (1996), and Frankel, Stein and Wei (1996) for an analysis of this issue in alternative economic frameworks.

⁴Furthermore, if an existing trading bloc follows a policy of “open” regionalism, by which accession is sequentially granted to all countries that demand it, this sequential process is likely to lead to worldwide free trade. See Yi (1996) for a discussion of “open” versus “closed” regionalism. In complementary work, Ethier (1998) presents a model in which multilateral liberalization among a subset of developed countries raises the incentives of outside (less developed) countries to seek preferential free trade agreements with particular

Another way to pose the question of regionalism versus multilateralism is to ask whether multilateral bargaining or sequential bargaining are more likely to lead to global free trade. In multilateral bargaining all countries simultaneously participate in a single round of trade negotiations. In sequential bargaining negotiators proceed through several rounds, with different subsets of countries participating at different stages of the process. In this paper we compare these two negotiation strategies. Since trade negotiations involve bargaining, we believe that it is important to address these issues in a framework that fully specifies the bargaining process.⁵

We develop a dynamic bargaining model of coalition formation, where a coalition consists of a preferential trade agreement. A leading country decides endogenously whether to negotiate sequentially with only a subset of countries or simultaneously with all countries. If the leading country chooses the sequential path, it also has to decide which follower countries to approach first, which second, and so on. We follow Bhagwati (1993) in adopting the view that the U.S. has been the leading country in the post-World War II period, and that it has disproportionately affected the process of trade liberalization. For this reason we model the bargaining game as a game in which one country, the leader, has special agenda-setting power.

In Section 2 we develop a simple transferable-utility game between three countries. One of the countries is the leader with agenda-setting power. In the first stage the leader decides to negotiate multilaterally or sequentially. If it chooses multilateral bargaining, the leader makes a simultaneous offer to form a coalition with the two follower countries. If it chooses the sequential path, the leader also decides which follower country to approach first. At each stage of the game the agenda-setter makes take-it-or-leave-it offers. This bargaining game allows us to identify the payoff of every coalition as a function of the *coalition structure*, i.e., the value function, and this mapping allows us then to characterize the solution to the bargaining game.

We first take the value function as given, and define two properties of this function that play a central role in the subsequent analysis: *coalition externalities* and *grand-coalition superadditivity*. A coalition is subject to coalition externalities when its payoff depends on which other coalitions form. In the simple three-country setup, this means that coalition externalities emerge whenever the size of a country's payoff depends on whether the other two countries form a coalition or not. Payoffs exhibit grand-coalition superadditivity when the payoff of the grand coalition is larger than the payoff of all countries combined in alternative coalition structures. This condition is satisfied in various models of international trade when free trade is Pareto-efficient and every country seeks to maximize its aggregate welfare. The

members of the initial group of liberalizing countries.

⁵Although the desirability of free trade can be questioned, for the purpose of this paper we assume that free trade is desirable, and we seek to identify negotiation strategies that lead to global free trade.

concepts of externalities and superadditivity have been used by Ray and Vohra (1997, 1999), Gomes (2005) and Maskin (2003) in various applications.

With these concepts in hand, we describe in Section 3 a benchmark result: if the payoffs are grand-coalition superadditive and no coalition externalities exist, then the agenda-setter is indifferent between multilateral and sequential bargaining and the grand coalition forms in equilibrium. Crucial for this result is the ability of countries to transfer utility within coalitions by means of side payments. This ensures that the leader country is able to internalize the welfare gains from the grand coalition. In the absence of such transfers global free trade may not be the equilibrium outcome, as Riezman (1985) showed in a cooperative game-theoretic model.⁶ We believe, however, that it is realistic to model trade negotiations as games with transferable utility, because the exchange of concessions on non-trade-related issues often serves the role of transfers that redistribute the gains from trade liberalization.⁷

The benchmark result relies on the assumption that there are no coalition externalities. As we show in Section 4, however, non-zero coalition externalities are the rule in the formation of free trade areas. Intuitively, if the reduction in trade barriers associated with a free trade area (FTA) affects world prices, the welfare of outside countries or trading blocs will be affected by the FTA.⁸ Importantly, we show that externalities can be positive or negative, depending on whether the FTA raises or lowers the world price of certain goods, and whether outside countries are net importers or exporters of these goods.

The generic presence of coalition externalities motivates our analysis in Section 5, in which we show that if the payoffs are grand-coalition superadditive and the coalition externalities are nonzero, then the leader is *not* indifferent between multilateral and sequential bargaining. In particular, the leader strictly prefers sequential bargaining when the coalition externalities are negative in at least one country, and it strictly prefers multilateral bargaining when the coalition externalities are positive in both follower countries. Furthermore, we show that — regardless of the sign and size of coalition externalities — the grand coalition forms in equilibrium, leading to global free trade. We label this result as our “Free Trade Proposition”.⁹

⁶Burbridge et al. (1997) develop an alternative coalition-formation game in which the grand coalition fails to form even with transfers within customs unions. Their result is, however, driven by the static nature of the game and special features of the coalition-formation process.

⁷Non-trade-related concessions include agreements on product and labor standards, and political reforms, such as the reforms that Mexico was expected to pursue in order to participate in NAFTA. Bagwell and Staiger (2004a) justify the modelling of lump-sum transfers in a similar manner. Their analysis is focused, however, on different issues of trade negotiations.

⁸Bagwell and Staiger (2002) argue that the WTO’s principles of reciprocity and nondiscrimination have been designed to neutralize such externalities. But preferential trade agreements rely on GATT’s Article XXIV, which specifies exceptions to the principle of nondiscrimination, and these exceptions create discriminatory rates of protection. See Bagwell and Staiger (2004b) for further details about this point. Chang and Winters (2002) provide evidence of externalities caused by MERCOSUR, and Winters (1997) reviews the evidence of such externalities in the process of European integration.

⁹Our efficiency result is distinct from the Kemp and Wan (1976) result about customs unions. In our general model, global free trade is attained for coalitions that can be customs unions, free trade areas, or

In Section 6 we use two examples to illustrate these results. In the first example sequential bargaining is the equilibrium outcome, and it leads to global free trade. In the second example multilateral bargaining is the equilibrium outcome, and it also leads to global free trade. We show how these equilibria depend on trade structure and the structure of protection.

A corollary of the results from Sections 5 and 6 is that, when grand-coalition superadditivity holds, preferential trade agreements are neither building blocs nor stumbling blocs on the way to worldwide free trade. Although, as in Levy (1997) and Krishna (1998), a preferential trade agreement may raise the reservation payoff of member countries in subsequent negotiations, grand-coalition superadditivity ensures that the leader has an incentive to strike deals that bring about global free trade. Similarly, although a preferential trade agreement may exert a negative externality on outside countries, as in Baldwin (1996), and make sequential negotiations more attractive for the agenda-setter, grand-coalition superadditivity ensures that multilateral negotiations also lead to global free trade. As a result, preferential trade agreements affect the distribution of payoffs but not the attainment of global free trade.

In Section 7 we explore implications of the failure of grand-coalition superadditivity. In particular, we derive conditions for the emergence of stumbling bloc and building bloc equilibria. A stumbling bloc equilibrium is one in which the agenda setter prefers sequential bargaining that does not lead to global free trade rather than multilateral bargaining that leads to global free trade. And a building bloc equilibrium is one in which the agenda setter prefers sequential bargaining that leads to global free trade rather than multilateral bargaining that preserves the status quo. We illustrate such equilibria with two examples in which negotiators maximize a function that describes a *political objective*, using an extreme version of the Grossman and Helpman (1994) model of politics with special interest groups in which this political-objective function coincides with aggregate profits.

In the first example, political pressure from special interests does not prevent multilateral bargaining from leading to free trade. Nevertheless, world profits are highest when the leader forms a free trade area with one country only, and the leader prefers this limited FTA to every other feasible outcome. Therefore the leader chooses sequential bargaining that does not lead to global free trade. In this case preferential trade agreements are stumbling blocs to worldwide free trade. If the WTO rules limited negotiations to multilateral bargaining, these constraints would ensure a free trade outcome in economies of this sort.¹⁰

economic unions. In particular, in the analysis of free trade areas that we use to illustrate the broader logic of these results the external tariffs of countries in a coalition do not change as a result of the formation of an FTA. Moreover, the impact of the coalition on outside countries is precisely what determines the choice between sequential and multilateral bargaining. Unlike Kemp and Wan's result, our efficiency result is driven by the transferability of utility, which ensures that one country fully internalizes the gains from trade liberalization.

Our result is also distinct from the main result in Furusawa and Konishi (2005), who consider a network of bilateral FTAs. Using network formation games, they show that in the presence of transferable utility the global free trade network is pairwise stable.

¹⁰In Krishna (1998) the stumbling bloc effect is derived in a model in which (i) markets are imperfectly

In the second example multilateral bargaining does not lead to global free trade, because the leader’s status-quo profits are higher than the residual profits it can get from an all-encompassing FTA that the follower countries are willing to join. Moreover, the leader prefers sequential bargaining, in which it gradually builds the grand coalition. In this case WTO rules that restrict trade negotiations to multilateral bargaining would harm the prospects of global free trade, whereas preferential trade agreements would encourage it.

In Section 8, we restore our assumption of GC superadditivity and explore alternative mechanisms that generate deviations from our global free trade result.¹¹ In particular, we show that an agenda setter who makes take-it-or-leave-it offers to follower countries may not be sufficient for global free trade when (i) following a rejection of its offer by a follower country in the first stage of the sequential subgame, the leader is allowed to make an offer to the second follower country in stage two; or (ii) the leader cannot commit to a payoff to the follower country in the first stage of the sequential bargaining subgame. These modifications in the bargaining protocols can enhance the bargaining power of the follower countries to a level at which the agenda-setter finds it too costly to attract both follower countries into a FTA. We also show in Section 8 that under these circumstances sequential bargaining can be a stumbling bloc for free trade but never a building bloc. Finally, we show that the structure of coalition externalities shapes the choice between multilateral and sequential bargaining in a way that is similar to the original model, as analyzed in Section 5. We offer a summary of the main results and suggestions for further research in Section 9.

2 The Bargaining Game

We consider a transferable-utility game between three countries: a , b , and c . We describe the game in partition form. We define a *coalition structure* as a partition Γ of $\{a, b, c\}$. That is, every country belongs to exactly one coalition. We interpret a coalition as a free trade area (FTA) in which member countries trade at zero tariffs.

For every partition Γ and every coalition $C \in \Gamma$ the value function $v(C; \Gamma)$ assigns a

competitive and internationally segmented; (ii) governments maximize domestic profits; and (iii) side payments between coalition members are not available. Our analysis suggests that the second of these features can produce a stumbling bloc effect. Furthermore, our analysis suggests that this feature can also produce a building bloc effect, which Krishna (1998) derives in footnote 20, but chooses to de-emphasize. Saggi and Yildiz (2006) also study the choice between bilateral and multilateral trade negotiations in an economic environment similar to Krishna (1998). We will discuss their contribution in Section 8.

¹¹At the end of Section 5, we show that our global free trade result holds in a variety of bargaining games: games with many countries in which a rejection of the leader’s offer ends the process of coalition formation, games with many countries in which a rejection of the leader’s offer transfers the agenda-setting power to a different country in a predetermined order, and games with many countries in which a rejection of the leader’s offer transfers stochastically the agenda-setting power to another country. We also show that it holds when the leading country is allowed to offer two subsequent FTAs to each of the two follower countries. Section 8 considers alternative bargaining protocols where our global free trade result may not hold.

payoff to C given the coalition structure Γ . This payoff is gross of lump-sum transfers. In this and the next section we treat these value functions as given, but later we will show how to construct them in specific models of international trade. The payoff functions have to be constructed from the objective functions that countries use to evaluate trade agreements.

The game is played as follows: One country is the leader, which means that it is the agenda-setter. Without loss of generality we assign this role to country c . In the first stage of the game the leader decides whether to enter multilateral or sequential bargaining, as shown in Figure 1.

[FIGURE 1 ABOUT HERE]

If c chooses multilateral bargaining, it makes a simultaneous take-it-or-leave-it offer to the follower countries a and b . The offer consists of an FTA that includes all countries and a system of lump-sum transfers. The transfers determine the payoffs $P(a)$ and $P(b)$ to countries a and b , respectively. If the offer is accepted by both countries, $\Gamma = \langle \{abc\} \rangle$ is the resulting coalition structure and the game ends. In this case Γ has a single element, consisting of the grand coalition, and the FTA leads to worldwide free trade. This sequence of events is described in the lower part of the game tree in Figure 1.

If one of the follower countries rejects c 's offer, then the coalition $\{abc\}$ does not form and the game ends with no agreement. In this event the coalition structure is $\Gamma = \langle \{a\}, \{b\}, \{c\} \rangle$.¹² This too is depicted in the lower part of the game tree in Figure 1.

Next consider the subgame in which c chooses sequential bargaining. In this event c has to decide whether to make the first take-it-or-leave-it offer to a or to b . If it makes the first offer to a , the offer consists of an FTA between a and c and lump-sum transfers between them that provide a with a payoff $P(a)$. If a accepts the offer, $P(a)$ is a 's payoff independently of whether the FTA is expanded to include country b .¹³ If a rejects the offer the game ends

¹²We believe that this is a reasonable specification for multilateral trade negotiations because the rules of the WTO (in particular, Article IX of the Marrakesh Agreement) state that the WTO should continue the GATT practice of decision-making by "consensus," where "the body concerned shall be deemed to have decided by consensus on a matter submitted for its consideration, if no Member, present at the meeting when the decision is taken, formally objects to the proposed decision." It is important to emphasize, however, that our assumption does not make full justice to the subtleties and nuances involved in the decision-making process within the WTO. First, the same Article IX mentioned above states that "when a decision cannot be arrived at by consensus, the matter at issue shall be decided by voting." Hence, it is not strictly true that a stubborn-enough country could single-handedly block a multilateral agreement. Second, even when decisions are reached by consensus, in reaching that consensus it is usually the case that a larger weight is given to views of particularly powerful countries in the trading system (see Jackson, 1997). Third, in the particular case of market access negotiations, the WTO allows countries to negotiate bilateral or small-numbers agreements, provided that that the provisions of these agreements are extended to the remaining members of the WTO on an MFN basis (see Bagwell and Staiger, 2004a, for a theoretical treatment). Because the focus of our paper is on a broad comparison of sequential trade liberalization versus multilateral trade liberalization, we have decided to abstract from these important details and assume the bargaining protocol in Figure 1.

¹³In the sequential bargaining subgame we assume that when c approaches country a first, it offers a a non-contingent payoff $P(a)$ for joining the FTA. See Section 8.2 for a discussion of contingent payoffs and the role of commitment in the leading country's offers.

and the coalition structure is $\Gamma = \langle \{a\}, \{b\}, \{c\} \rangle$.¹⁴

Whenever a accepts c 's offer, country c proceeds to make a take-it-or-leave-it offer to b , which consists of an expansion of the FTA to include all three countries and lump-sum transfers that provide b with a payoff $P(b)$. If b accepts the offer the coalition structure is $\Gamma = \langle \{abc\} \rangle$, and there is free trade. If b rejects the offer the coalition structure is $\Gamma = \langle \{ac\}, \{b\} \rangle$, i.e., a and c form a free trade area in which b is not included.

The subgame in which country c makes its first offer to b is symmetric and we omit its discussion. The upper part of Figure 1 describes the branches of the sequential bargaining subgame.

Note that in this game global free trade can emerge when the leader chooses either multi-lateral or sequential bargaining, and lack of free trade can also occur under both bargaining procedures. We seek a subgame perfect equilibrium. Country c chooses the bargaining method that maximizes its payoff.^{15,16}

In order to simplify the notation, we define the following functions, which describe gross payoffs (i.e., exclusive of lump-sum transfers):

$$W(j) \equiv v(j; \{a\}, \{b\}, \{c\}) \quad \text{for all } j = a, b, c,$$

$$W_F(j) \equiv v(j; \{j\}, \{k\ell\}) \quad \text{for all } j, k, \ell = a, b, c \text{ and } j \neq k, j \neq \ell, k \neq \ell,$$

$$W(k\ell) \equiv v(k\ell; \{j\}, \{k\ell\}) \quad \text{for all } j, k, \ell = a, b, c \text{ and } k \neq \ell, k \neq j, \ell \neq j,$$

$$W(abc) \equiv v(abc; \{a, b, c\}).$$

In this notation $W(j)$ is country j 's payoff when there are no free trade agreements; $W_F(j)$ is country j 's payoff when the other two countries form an FTA in which j is not included; $W(k\ell)$ is the joint payoff of countries k and ℓ when they form an FTA in which the third country is not included; and $W(abc)$ is the joint payoff of all three countries when they form an all-inclusive free trade agreement.¹⁷

¹⁴This assumption is relaxed in Sections 5 and 8.

¹⁵This formulation of the game shows clearly that if utility were nontransferable, the agenda setter would strictly prefer sequential bargaining only if it expected to form an agreement with only one follower country, excluding the other from the FTA. The reason is that when utility is not transferable, the agenda setter's payoff from a coalition structure is independent of the *path* through which this coalition structure is attained (and this is true for all coalition externalities). Thus, for example, the agenda setter obtains the same payoff from the grand coalition independently of whether it has been formed via multilateral or sequential bargaining. Under these circumstances preferential trade agreements cannot be building blocs of global free trade when global free trade is not the equilibrium outcome in the multilateral bargaining subgame.

¹⁶Our game is also relevant for European integration. See CEPR (1995), and especially Section 3.3 on principles of flexible integration strategies.

¹⁷The gross payoffs $W(\cdot)$ and $W_F(\cdot)$ will typically depend on the lump-sum transfers across countries, as is well known from the early work of Samuelson (1952, 1954), and the subsequent work of Jones (1970) and Bhagwati, Brecher and Hatta (1983), among others. This dependence stems from the fact that a transfer affects the terms of trade of the giving and receiving countries, as well as the terms of trade of other trading

A coalition C is not subject to coalition externalities when its payoff is independent of what other coalitions form. In our three-player game this suggests a simple definition:

Definition: Coalition Externalities *There are positive coalition externalities in country j when $W_F(j) > W(j)$, negative coalition externalities when $W_F(j) < W(j)$, and no coalition externalities when $W_F(j) = W(j)$.*

We also need a concept of superadditivity, which we define as follows:

Definition: Grand-Coalition (GC) Superadditivity *There is GC superadditivity if*

$$W(abc) > W(a) + W(b) + W(c), \text{ and}$$

$$W(abc) > W_F(j) + W(k\ell) \text{ for all } j \neq k, j \neq \ell.$$

In other words, grand-coalition superadditivity requires the joint payoffs of the three countries to be larger under global free trade than under no free trade agreements whatsoever or a free trade agreement between any two countries k and ℓ .

3 Benchmark

In this section we characterize equilibria for games with GC superadditivity and no coalition externalities in the follower countries. This helps in developing the intuition and provides a benchmark for more general games.

First consider the subgame with multilateral bargaining. Let c offer a free trade agreement between all countries, with payoffs $P(a)$ and $P(b)$. If $W(a) > P(a)$ country a rejects the offer, because a gets a higher payoff in the coalition structure $\Gamma = \langle \{a\}, \{b\}, \{c\} \rangle$. And if $W(b) > P(b)$ country b rejects the offer. When the offer is rejected by either a or b , the leader's payoff is

$$P(c) = W(c).$$

It is evident that under these circumstances c has to offer a at least $W(a)$ and it has to offer b at least $W(b)$ for the FTA to be accepted by both countries. Therefore c 's highest payoff from offers that are accepted by a and b is

$$P(c) = W(abc) - W(a) - W(b), \tag{1}$$

countries. Yet there are special cases, such as the economic model we develop in Section 6, in which transfers do not change international prices, and they therefore have no impact on the gross payoff functions. It is easiest to think about these special cases when interpreting our bargaining model, but we shall argue below that it is possible to reinterpret the payoff functions $W(\cdot)$ and $W_F(\cdot)$ in a way that makes the theory more generally applicable.

where $P(a) = W(a)$ and $P(b) = W(b)$ are c 's offers. GC superadditivity implies, however, that

$$W(abc) - W(a) - W(b) > W(c).$$

Therefore in the subgame of multilateral bargaining c prefers to make an offer that the follower countries accept, which leads to the formation of the grand coalition and to worldwide free trade.¹⁸

Next consider the subgame with sequential bargaining, and examine the case in which c approaches a first and offers it an FTA with a payoff $P(a)$. If $W(a) > P(a)$ the offer is rejected and c 's payoff is $P(c) = W(c)$. Therefore c has to offer a at least $W(a)$ for a to accept the offer, and it is in c 's interest to offer just $W(a)$. If c then proceeds to make b an offer that b rejects, the leader's payoff is

$$P(c) = W(ac) - W(a).$$

If, instead, c wants to make b an acceptable offer, c has to offer b a payoff of at least $W(b)$, and c has no interest in making a higher offer.¹⁹ Therefore (1) also describes c 's payoff from an offer that b accepts. But GC superadditivity implies that

$$W(abc) - W(a) - W(b) > W(ac) - W(a).$$

Therefore c prefers to make acceptable offers to both follower countries rather than only to a . Note also that under GC superadditivity

$$W(abc) - W(a) - W(b) > W(c),$$

which implies that country c prefers to make acceptable offers to the follower countries rather than an offer that a rejects.

Similar results obtain when country c makes its first offer to b . In fact, in the subgame with sequential bargaining the leader's payoff is the same independently of whether it approaches a or b first. In both cases c prefers to make offers that both follower countries accept, which

¹⁸Suppose that the payoff of the grand coalition depends on transfers, and let $\hat{W}[abc; P(a), P(b)]$ be this payoff as a function of the payoffs of countries a and b . But assume that $\hat{W}[abc; P(a), P(b)] - P(a) - P(b)$ is declining in $P(a)$ and in $P(b)$. This is the case if a transfer of income from country c to a or from c to b leads to a loss of welfare in c , after accounting for changes in the terms of trade. That is, this assumption excludes the possibility that, say, a transfer of income from c to a will improve c 's terms of trade to such an extent that it will gain on net from the transfer. Under this assumption the argument in the text applies when $W(abc)$ is interpreted to be $W(abc) \equiv \hat{W}[abc; W(a), W(b)]$. The gross payoff functions can be redefined in similar fashion for other arguments in the main text. For ease of exposition, however, we shall think about economic models of the type developed in Section 6, in which transfers do not affect the terms of trade.

¹⁹More accurately, c has to offer b at least $W_F(b)$, but $W_F(b) = W(b)$ because there are no coalition externalities in b .

leads to the formation of the grand coalition and to worldwide free trade.

We note that in both the multilateral and sequential bargaining subgames, the grand coalition forms, and (1) is country c 's payoff. We have therefore proved the

Benchmark Proposition *If there are no coalition externalities in the follower countries and there is GC superadditivity, then:*

- (i) *the leader is indifferent between multilateral and sequential bargaining; and*
- (ii) *the grand coalition forms and there is global free trade.*

This proposition establishes our benchmark. Deviations from this benchmark can result from coalition externalities or from the failure of GC superadditivity. We first show in the next section that coalition externalities are generic features of free trade agreements, and we characterize in the subsequent section equilibria with such externalities and GC superadditivity. There we argue that GC superadditivity is satisfied in a competitive environment in which the objective function of every country is to maximize the aggregate welfare of its residents.

4 Coalition Externalities

We show in this section that free trade agreements lead naturally to coalition externalities. We interpret an FTA as an agreement that removes tariffs on trade between members of the FTA whereas every country in the FTA maintains its original rates of protection vis à vis countries outside the FTA. This interpretation is consistent with GATT/WTO Article XXIV.

[FIGURE 2 ABOUT HERE]

Consider a particular industry whose goods are imported from b by countries a and c and in which the rate of protection is higher in a than in c . Figure 2 depicts the import demand function in country a , $C_a - X_a$, where C_a represents demand and X_a represents supply in a , as well as two possible supply functions in c , $X_c[1]$ and $X_c[2]$.²⁰ The international price of the product is p whereas τ_a and τ_c represent 1 plus the rate of protection in countries a and c , respectively. By assumption, $\tau_a > \tau_c$, and therefore the consumer and producer price in a , $\tau_a p$, exceeds the consumer and producer price in c , $\tau_c p$. We also assume that the export supply function of country b , not drawn in the figure, is upward sloping.

First suppose that $X_c[1]$ is the supply function in c and let us examine how the joint excess demand of countries a and c changes as a result of an FTA between them. It is evident from the figure that if the price in a were to decline to the price $\tau_c p$ in c , then country c

²⁰This discussion borrows from Grossman and Helpman (1995). See also Richardson (1993).

would be able to supply the entire import demand of a at this lower price. For this reason the price in a declines to $\tau_c p$ and a switches to import the product from c without violating the *rules of origin*, which are standard provisions of such agreements.²¹ This is a case of *reduced protection*, which leads to trade creation within the free trade area. Since prices do not change in c , c 's net import demand $C_c - X_c [1]$ does not change as well. It follows that, at the original international price p , the joint import demand of a and c rises. As a result, the world's excess demand for the product rises, leading to a higher international price p . The increase in the international price affects the payoff of country b . If, for example, the objective function of country b is to maximize the aggregate welfare of its residents, then the FTA between a and c imposes a positive coalition externality on b , because it improves b 's terms of trade. In this event $W_F(b) > W(b)$. Naturally, this discussion is confined to one industry only and a proper evaluation of coalition externalities requires an examination of the aggregate effects across all sectors. Yet the main message of this example is broad: we should expect nonzero coalition externalities when free trade areas form.²²

Coalition externalities can be positive or negative. We showed in the previous paragraph that they can be positive. Now we show that they can be negative.

Suppose that the supply function in country c is $X_c [2]$. In this event, suppliers in c do not offer enough output at the price $\tau_a p$ to satisfy country a 's import demand at this price, so even if country a were to purchase all of c 's output it would still need to import from b . As a result an FTA between a and c does not change the consumer and producer prices in country a , which remain $\tau_a p$, and it does not change the consumer price in c , which remains $\tau_c p$. However, it does change the producer price in c , which rises to $\tau_a p$, the price in a . The producer price in c rises because the FTA permits these producers to sell in a without the tariff impediments, and the price in a is higher than in c . As a consequence producers in c sell their entire output in a and consumers in c import their entire consumption from b . This is a case of *enhanced protection*; the FTA leads to higher (producer) prices. Since the consumer prices do not change while the producer price rises in c and does not change in a , the joint import demand of countries a and c declines. Therefore p declines, worsening b 's terms of trade. This worsening of the terms of trade generates a negative coalition externality on b if b 's objective is to maximize the joint welfare of its residents, i.e., $W_F(b) < W(b)$.

It is now clear that there are very good reasons for nonzero coalition externalities in free trade agreements.²³ We therefore proceed to discuss solutions to the bargaining game in the

²¹Note that the FTA reshuffles trade flows. Country a ceases to import from b despite a 's expansion of imports. But country c increases its imports from b in order to allow a to purchase goods in c . Yet standard rules of origin are not violated, because a can import from c only products that are produced in c . There is no need for products that c imports from b to be exported from c to a in order to meet a 's demand.

²²This example delivers precise answers about coalition externalities when the economic structure is of the type discussed in Section 6.

²³The empirical evidence points in the same direction. Reviewing the literature on European economic integration, Winters (1997) reports that most studies find what amounts to negative coalition externalities.

presence of such externalities.

5 Free Trade with Coalition Externalities

Consider payoffs $v(C; \Gamma)$ that exhibit coalition externalities, but which are GC superadditive. This specification deviates from the benchmark in Section 3 by allowing coalition externalities. Under these circumstances the payoff of c from multilateral bargaining is the same as in the benchmark case, i.e. (1), because the solution to the multilateral subgame depends only on GC superadditivity and not on coalition externalities. It follows that multilateral bargaining leads to the formation of the grand coalition and to free trade.

Next consider sequential bargaining. If c first offers a an FTA and a payoff $P(a) = W(a)$, then a accepts the offer.²⁴ In this event c has to offer b a payoff $P(b) = W_F(b)$ for b to join the FTA. Since GC superadditivity implies that

$$W(abc) - W(a) - W_F(b) > W(ac) - W(a),$$

country c gains from expanding the free trade area to include b once it has formed an FTA with a , because the left-hand-side of this inequality represents c 's payoff from an all-encompassing free trade area while the right-hand-side represents c 's payoff from a free trade area with a only. It follows that c 's payoff from making acceptable offers in a sequential bargaining subgame in which c approaches a first is

$$P^{a,b}(c) = W(abc) - W(a) - W_F(b).$$

By similar argument c 's payoff from making acceptable offers in a sequential bargaining subgame in which c approaches b first is

$$P^{b,a}(c) = W(abc) - W(b) - W_F(a),$$

as long as there is GC superadditivity. Comparing $P^{a,b}(c)$ with $P^{b,a}(c)$ we see that c is indifferent between which country it approaches first if and only if $W_F(a) - W(a) = W_F(b) - W(b)$, i.e., the coalition externalities are the same in the two follower countries. This holds in the benchmark case, in which the coalition externalities are zero. Moreover, a comparison of

Chang and Winters (2002) find that MERCOSUR has worsened the terms of trade of a number of non-member countries, including the U.S. and Japan. MERCOSUR is a customs union between Argentina, Brazil, Paraguay and Uruguay. Unlike a free trade area, a customs union imposes common external tariffs on non-member countries. Chang and Winters find that foreign prices charged to Brazil by non-MERCOSUR countries declined as a result of Brazil's lowering of tariffs on goods imported from Argentina.

²⁴Note that a accepts every offer that satisfies $P(a) \geq W(a)$, but it is in c 's interest to offer $W(a)$. In what follows we restrict c 's offers to the lowest payoffs $P(j)$ that the other parties accept, which is a condition for subgame perfection.

these payoffs shows that c strictly prefers to approach the country with the higher coalition externalities first, i.e., it prefers to approach a first if $W_F(a) - W(a) > W_F(b) - W(b)$ and it prefers to approach b first if $W_F(a) - W(a) < W_F(b) - W(b)$. The reason is that by approaching the country with the higher coalition externalities first the leader reduces the joint outside options of the follower countries. We conclude that c 's highest payoff from sequential bargaining with acceptable offers is

$$P_{accept}(c) = W(abc) - W(a) - W(b) - \min\{W_F(a) - W(a), W_F(b) - W(b)\}. \quad (2)$$

Now note that c has the option of making offers that are rejected by the first country. A rejection gives c the payoff $W(c)$. For this reason c does not proceed with acceptable offers unless $W(c) \leq P_{accept}(c)$.

It remains to compare the leader's payoffs from multilateral and sequential bargaining. Comparing (1) with (2) implies

$$P_{accept}(c) = P_{multi}(c) - \min\{W_F(a) - W(a), W_F(b) - W(b)\},$$

where $P_{multi}(c) = W(abc) - W(a) - W(b)$ is c 's payoff in the multilateral subgame. It follows immediately that c prefers sequential bargaining when

$$\min\{W_F(a) - W(a), W_F(b) - W(b)\} < 0$$

and multilateral bargaining when²⁵

$$\min\{W_F(a) - W(a), W_F(b) - W(b)\} > 0.$$

Moreover, whichever subgame c prefers leads to the formation of the grand coalition and to global free trade. We have thereby proved

Free Trade Proposition *If there is GC superadditivity, then:*

- (i) *the leader is indifferent between multilateral and sequential bargaining if and only if there are no coalition externalities in the follower countries;*
- (ii) *the leader strictly prefers sequential bargaining when there are negative coalition externalities in at least one of the follower countries;*

²⁵Note that $\min\{W_F(a) - W(a), W_F(b) - W(b)\} < 0$ implies $P_{accept}(c) > P_{multi}(c) > W(c)$, where the last inequality results from GC superadditivity. Therefore in this case c prefers sequential bargaining with acceptable offers to sequential bargaining in which the first offer is rejected. The only case in which c prefers sequential bargaining in which the first offer is rejected rather than accepted is when c also prefers multilateral to sequential bargaining.

(iii) the leader strictly prefers multilateral bargaining when there are positive coalition externalities in both follower countries; and

(iv) the grand coalition forms and there is worldwide free trade.

This proposition states that global free trade is the unique equilibrium outcome when payoffs are GC superadditive. It also identifies conditions under which sequential or multilateral bargaining is the equilibrium outcome. Sequential bargaining is the equilibrium outcome when negative coalition externalities exist in at least one follower country, while multilateral bargaining is the equilibrium outcome when positive coalition externalities exist in both follower countries.²⁶

Our Free Trade proposition has important implications.²⁷ Consider a neoclassical world in which production sets are convex, tariffs are the only distortions, and all markets are competitive. Also suppose that the payoff of every country is represented by the aggregate welfare of its residents and that the marginal utility of income is constant. Then GC superadditivity holds, because global free trade is Pareto-efficient. That is, in this sort of world, lump-sum transfers ensure that the joint welfare of all three countries combined is higher under free trade than under limited free trade agreements or no free trade agreement at all. Under these circumstances our free trade proposition applies, and trade negotiations lead to global free trade.

It follows from this proposition that in the presence of GC superadditivity countries need not be restricted to multilateral bargaining — as favored by Bhagwati (1991) — in order to safeguard free trade, because it is not in the interest of the leading country to choose sequential bargaining *unless* it leads to free trade. True, an institutional prohibition on sequential bargaining secures free trade. But, as we will show in Section 7, the potential advantage of multilateral bargaining disappears when payoffs are not GC superadditive. Finally, note that even with GC superadditivity, restrictions on bargaining have distributional implications. The leading country's payoff is higher when it is free to choose whether to bargain sequentially or multilaterally than when it is restricted to bargaining multilaterally, unless it prefers multilateral bargaining. Recall, however, that the leader prefers sequential bargaining if and only if at least one follower country has negative coalition externalities. In this case a switch from unrestricted bargaining to mandatory multilateral bargaining redistributes payoffs from the leading country to the follower country with the largest negative coalition externalities.²⁸

²⁶Note that country c has an incentive to pursue policies that hurt the outside option $W_F(j)$ of country j when c seeks to bring j into the FTA in the last stage of the sequential bargaining subgame. If this option is feasible, then such policies can be built into the payoff functions so that the payoffs reflect the implementation of these policies.

²⁷Although our focus is on free trade agreements, this type of analysis can be applied to other international agreements. In environmental agreements, for example, the coalition externalities are positive; see Carraro and Siniscalco (1993).

²⁸To illustrate, suppose that $W_F(b) - W(b) < W_F(a) - W(a)$ and $W_F(b) - W(b) < 0$. Then a 's payoff is

5.1 Generalizations

Our Free Trade Proposition can be generalized, and we offer four such generalizations.

First, consider the case in which the leader country is allowed to sign two subsequent bilateral FTAs with each one of the two follower countries. For this purpose we expand the definition of GC superadditivity to also cover the coalition structure $\Gamma = \langle \{ac\}, \{bc\} \rangle$, which consists of two FTAs, one between a and c , the other between b and c . To the definition of GC superadditivity we now add the requirement that $W(abc) > W(ac, bc)$, i.e., aggregate welfare is higher under the grand coalition than under $\Gamma = \langle \{ac\}, \{bc\} \rangle$.

To see that this modification does not affect our Free Trade Proposition, consider the branch of the sequential subgame on which c approaches a first and b second (the other case is analogous). Suppose c makes a an offer that a accepts and they form an FTA. Next c approaches b , and now allow c to not only offer global free trade to b , but also to offer b the possibility of a bilateral FTA. At this point a and c have formed a bilateral FTA, so regardless of the nature of c 's second-stage offer, country b 's reservation price is $W_F(b)$.

Now roll back and consider the negotiation between c and a in stage one. Since a rejection of c 's offer leaves a with the payoff $W(a)$, and this payoff is independent of what c might offer b in the second stage should a accept the offer, the agenda setter has to pay $W(a)$ in order to bring a into any coalition. It follows that c has to pay $W_F(b) + W(a)$ for the formation of the grand coalition as well as for the formation of $\Gamma = \langle \{ac\}, \{bc\} \rangle$. In other words, the two alternative coalition structures are equally costly to the agenda setter. But GC superadditivity implies that $W(abc) > W(ac, bc)$. Therefore c prefers to form the grand coalition. In sum, a pair of subsequent FTAs is a dominated strategy in our sequential subgame and therefore the Free Trade Proposition holds in this case too.

Three further generalizations and modifications, based on part (iv) of the proposition — which states that the grand coalition forms, leading to global free trade — are offered in Aghion, Antràs and Helpman (2006) (see Appendix A). The first generalization considers a world with many countries, but maintains the assumption of the simple model that the game ends when the leader's offer is rejected by one of the follower countries. As in the simple game, the leader can make a simultaneous offer to all follower countries, which we refer to as multilateral bargaining. Alternatively, it can engage in sequential bargaining, in which case it makes an offer to a subset of the follower countries. If this offer is rejected the game ends and there are no FTAs. If the offer is accepted, the leader can make a second offer to a subset of countries that are not yet included in the FTA. If the offer is rejected the game ends and the coalition structure consists of the FTA formed in round one. If the offer is accepted, the FTA is expanded and the leader country can make a new offer to a subset of countries that

$W(a)$ under sequential and multilateral bargaining, but b 's payoff is $W_F(b)$ under sequential bargaining and $W(b)$ under multilateral bargaining. Evidently, b prefers multilateral bargaining while c prefers sequential bargaining.

are still outside the FTA. This process continues until either an offer is rejected by one of the follower countries or all the countries are included in the free trade area.

The second extension also considers a world with many countries, but this time a rejection of the leader's offer does not end the game. Instead, the agenda-setting power is transferred to another country in a predetermined order. That is, if the original leader's offer is rejected by one of the follower countries, then the agenda-setting power is transferred to a follower country. The country chosen is the first in the queue for leadership among the countries that are not members of the first leader's FTA. The new leader can form a new FTA by offering membership to countries that are not yet members of an FTA. In this way a new FTA forms. When an offer of the new leader is rejected, the agenda-setting power is again transferred to a country that is not a member of an FTA, using the predetermined queue. And the process continues in the same way with additional leaders. The game ends when all countries are members of FTAs, some of which may consist of one country only, and there are no more leaders to whom the agenda-setting power can be transferred. Unlike the original game, this one allows for multiple FTAs.

In our final extension a rejection of the leader's offer again leads to the transfer of agenda-setting power to a follower country that is not a member of an FTA. This time, however, the next leader is chosen randomly from the eligible set of countries, defined as countries that are not members of an FTA and were not leaders in earlier rounds of negotiations. This setup is similar to Gomes (2005).²⁹

We conclude from these extensions and modifications that GC superadditivity is a powerful feature; it ensures the formation of the grand coalition and global free trade for a variety of bargaining protocols, independently of the structure of coalition externalities.³⁰ The coalition externalities affect the sequence in which the FTA expands, but not the equilibrium coalition structure. They also affect the equilibrium payoffs, i.e., the distribution of the gains from trade negotiations. Moreover, although multilateral negotiations ensure the formation of the grand coalition in all the above cases, the equilibrium bargaining process need not consist of an offer to all follower countries combined; the leader may gain more from sequential offers.

Finally, note that our results also apply to situations in which countries organize themselves into free trade areas prior to the beginning of the game. In this event a country can be interpreted as a bloc of countries that have formed an FTA, and all the arguments made above remain valid. This means that if there are regional trade agreements or other preferential trade agreements prior to the beginning of our negotiation process, then these limited agreements do not prevent the attainment of global free trade when payoffs are GC superadditive.³¹

²⁹We are grateful to Eddie Dekel and Attila Ambrus for referring us to this paper.

³⁰See, however, our discussion in Section 8 for different alternatives.

³¹Although, as we have shown, the bargaining protocol need not be restricted to sequential or multilateral

6 Illustrations of the Free Trade Proposition

In this section we construct an economic model that gives precise meaning to the payoff functions $W(\cdot)$ and $W_F(\cdot)$, and we use the model to illustrate the free trade proposition. In this model importing countries impose tariffs and a coalition consists of a free trade area. Standard rules of origin apply in the FTA and they prohibit a member country from importing goods from outside the FTA via another member country that has lower tariffs.

We assume that the utility function of the residents of country j is quasi-linear, given by

$$U_j = y_j + u_j(x_j), \quad (3)$$

where y_j is their consumption of good y and x_j is their consumption of good x . The function $u_j(\cdot)$ is increasing and concave. Good y is the numeraire; its price is one and it is not protected by tariffs. As is well known, if such consumers have enough income to consume both goods, which we assume to be the case, the demand for x depends only on its price, $x_j = C_j(q_j)$, and this demand function is downward sloping. Then country j 's indirect utility function is

$$V_j = I_j + S_j(q_j),$$

where I_j is its income and $S_j(q_j) \equiv u_j[C_j(q_j)] - q_j C_j(q_j)$ is the consumer surplus function. By standard arguments $S'_j(q_j) = -C_j(q_j)$.

Assume that good y is produced with 1 unit of labor per unit output, whereas x is produced with labor and a sector-specific input under constant returns to scale. Then the wage rate equals 1 as long as the country produces y , which we assume to be the case, and the income of the sector-specific input, which we shall identify with profits, is an increasing convex function $\Pi_j(q_j)$. By standard arguments $\Pi'_j(q_j) = X_j(q_j)$, where $X_j(q_j)$ is an upward sloping supply function of x .

Let τ_j be 1 plus the MFN tariff rate on imports of x .³² If x is exported by country j we set $\tau_j = 1$. That is, we assume that there are no export taxes or subsidies. Then $q_j = \tau_j p$ is the consumer and producer price in the absence of free trade agreements, where p is the

bargaining in order to achieve global free trade when payoffs are GC superadditive, restrictions on bargaining may be desirable for other reasons. Imagine, for example, a world of three countries in which payoffs are GC superadditive and countries are symmetric in the sense that (i) $W(j) = w$ for $j = a, b, c$ and $W_F(j) = w_F$ for $j = a, b, c$; and (ii) every country has an equal chance of becoming the agenda setter. Moreover, coalition externalities are negative, i.e., $w > w_F$. Under these circumstances the country that becomes the agenda setter chooses sequential bargaining, which gives it the payoff $W(abc) - w - w_F$, one of the follower countries gets the payoff w and the other follower country gets the payoff w_F . Then, behind a Rawlsian veil of ignorance these countries prefer to restrict the bargaining protocol to multilateral bargaining only, which gives the agenda setter the payoff $W(abc) - 2w$ and each one of the follower countries the payoff w . This restriction raises the payoff of the least fortunate country, meeting the Rawlsian criterion of seeking to minimize the downside risk.

³²An MFN tariff is one in which the same tariff rate applies to imports from all sources, according to the most favored nation (MFN) clause. We use this specification in the examples, but our free trade proposition applies also when the tariffs do not satisfy this requirement.

international price of good x . Tariff revenue is distributed to country j 's residents, who also own the country's labor and sector-specific input. Under these circumstances income I_j consists of labor income, profits and tariff revenue. Therefore in the absence of free trade agreements the indirect utility function is³³

$$V_j = L_j + \Pi_j(\tau_j p) + (\tau_j - 1)p[C_j(\tau_j p) - X_j(\tau_j p)] + S_j(\tau_j p), \quad (4)$$

where L_j is labor supply. In the absence of free trade agreements the international price p is determined by the market clearing condition

$$\sum_{j=a,b,c} [C_j(\tau_j p) - X_j(\tau_j p)] = 0.$$

Evidently, the international price depends on the rates of protection.

In the following examples we assume that $W(j)$ equals the indirect utility V_j that country j attains in the absence of free trade agreements; $W(jk)$ equals the sum of the indirect utilities $V_j + V_k$ that countries j and k attain when they form an FTA that does not include the third country; $W_F(j)$ equals the indirect utility V_j that j attains when it is not included in a free trade area formed by the other two countries; and $W(abc)$ equals the sum of the indirect utilities $V_a + V_b + V_c$ that the three countries attain under free trade. As is well known, under these circumstances free trade yields the highest sum of utilities and therefore the payoffs are GC superadditive. These payoffs are generated by governments that maximize the aggregate welfare of their residents.

Example 1: Equilibrium Sequential Bargaining

Suppose that in the equilibrium with no trade agreements country c exports x while countries a and b import x . Moreover, a and b impose tariffs $\tau_a > 1$ and $\tau_b > 1$. Also suppose that if c forms a free trade area with one of the follower countries the FTA leads to reduced protection (recall the discussion of reduced and enhanced protection in Section 4). Then the free trade area with one of the follower countries raises the international price of x .

This is illustrated in Figure 3 for an FTA between c and a . The supply of exports by c is represented by $X_c(p) - C_c(p)$ while the aggregate import demand function of the follower countries is $\sum_{j=a,b} [C_j(\tau_j p) - X_j(\tau_j p)]$. In the absence of free trade agreements the international price is determined by the intersection of these two curves, identifying p_n as the equilibrium price.

[FIGURE 3 ABOUT HERE]

³³As we have seen in Section 4, free trade agreements can produce a deviation of the consumer price from the producer price. The formulation of the indirect utility function has to be modified in an obvious way when this happens.

An FTA between c and a shifts rightward the aggregate import demand curve of countries a and b to the broken-line curve, because the import demand function of country a shifts from $C_a(\tau_a p) - X_a(\tau_a p)$ to $C_a(p) - X_a(p)$.³⁴ The international price rises, worsening b 's terms of trade. As a result V_b declines. In this event there are negative coalition externalities in b .³⁵ Therefore our proposition implies that in this example the equilibrium consists of sequential bargaining in which c makes an offer to the country with the larger coalition externalities first, and an offer to the country with the smaller coalition externalities second, which in this case is negative.³⁶

Example 2: Equilibrium Multilateral Bargaining

Suppose that in the equilibrium with no trade agreements country c imports x from each of the follower countries and c 's MFN tariff is $\tau_c > 1$. Figure 4 depicts c 's import demand function $C_c(\tau_c p) - X_c(\tau_c p)$ and the joint export supply function $\sum_{j=a,b} [X_j(p) - C_j(p)]$ of the follower countries. With no free trade agreements the equilibrium international price is p_n .

[FIGURE 4 ABOUT HERE]

Now suppose that c forms an FTA with a , and assume that the FTA leads to reduced protection.³⁷ Then the price in c declines to the price in a , which equals the international price. As a result the import demand function $C_c(\tau_c p) - X_c(\tau_c p)$ changes to $C_c(p) - X_c(p)$, which is depicted in the figure by a rightward shift of the $C_c - X_c$ curve to the broken-line curve, and the international price rises. Unlike the previous example, however, this time the price hike concerns b 's exports. Therefore b 's terms of trade improve and V_b rises. Evidently, b has positive coalition externalities.

We can establish in similar fashion positive coalition externalities in a . Under these circumstances our proposition implies that multilateral bargaining takes place in equilibrium.

These two examples together with the examples discussed in Section 4 suggest that the equilibrium bargaining method depends not only on the pattern of trade, but also on finer

³⁴Reduced protection requires $C_a(p_n) - X_a(p_n) < X_c(p_n)$.

³⁵Note that (4) implies

$$\frac{\partial V_j}{\partial p} = -(C_j - X_j) \tau_j + (\tau_j - 1) p (C'_j - X'_j) \tau_j .$$

This is negative if j imports x , and it is positive if j exports x and $\tau_j = 1$.

³⁶This example also illustrates the dangers of inferring the sign and size of coalition externalities from the available studies of regional trade agreements (see footnote 23 for references). Our model calls attention to the existence of an important selection bias, since we predict that these partial agreements will be signed *only* when they create a negative externality for the country that is temporarily left out. This suggests that positive externalities might well exist in the data, but we might never observe them in studies of partial agreements.

³⁷Namely, $C_c(p_n) - X_c(p_n) < X_a(p_n)$.

details of the supply and demand functions. To see why, reconsider Example 2. We assumed in the example that the formation of an FTA between c and either one of the follower countries leads to reduced protection. This resulted in positive coalition externalities in the follower countries. Now suppose instead that the formation of an FTA between c and a leads to enhanced protection. Namely, $C_c(\tau_c p_n) - X_c(\tau_c p_n) > X_a(\tau_c p_n)$. In this event b is subject to negative coalition externalities, because — by raising the supply price in a to the supply price in c — the FTA raises the export supply of country a , thereby reducing the international price. The lower international price of x is detrimental to b , which exports x . In this event our proposition implies sequential bargaining in equilibrium, because negative coalition externalities exist in one of the follower countries. Evidently, the same pattern of trade can lead to different equilibrium bargaining protocols.

7 No GC Superadditivity

GC superadditivity is central to our benchmark and free trade propositions. In particular, when GC superadditivity fails, free trade is not necessarily the unique equilibrium outcome. We argued in the previous section that GC superadditivity holds when the aggregate welfare of a country's residents is used as its payoff, no distortions exist other than tariffs, production sets are convex, and all markets are competitive. A modification of any one of these features can destroy GC superadditivity. If, for example, markets are not competitive, then free trade is not Pareto-efficient, and GC superadditivity may not hold.³⁸

Of particular interest is lack of GC superadditivity that stems from political economy considerations. Suppose that the payoff from a free trade agreement is not represented by the aggregate welfare of a country's residents, but rather by a political-objective function. Following Grossman and Helpman (1994), think about countries whose political activities lead to policies that maximize a function of aggregate welfare and political contributions. Then the interaction of the government with special interest groups leads to the maximization of a weighted average of aggregate welfare and the welfare of the special interests.³⁹ Under these circumstances GC superadditivity may fail, and equilibria without free trade can emerge.⁴⁰ When this happens, it is possible to evaluate in a meaningful way the desirability of restrictions on the bargaining protocol.

Our approach follows the logic of constitutional design, where one seeks restrictions on the

³⁸GC superadditivity may also fail when some sectors are excluded from the preferential trade agreements, or when tariffs are reduced, but not to zero, among the parties to an agreement. Exclusion of some sectors, and slow phasing in of other sectors, are common features of preferential trade agreements.

³⁹See Grossman and Helpman (2001) for a systematic discussion of such political objective functions and the ways in which they can arise in various polities.

⁴⁰Note that GC superadditivity holds when the weight on aggregate welfare is sufficiently higher than the weight on contributions. This follows from continuity and the fact that when this relative weight goes to infinity GC superadditivity holds.

executive that yield desirable outcomes. The desirable outcomes do not necessarily include objectives of the political players, but they do include objectives of the polity at large. Under these circumstances constitutional features seek to achieve the desirable outcomes despite the fact that the political players may have different objectives from the polity. In our context this means that we are interested in identifying situations in which free trade is attained even when policy makers do not seek to maximize aggregate welfare. Despite a possible lack of congruence between the objectives of policy makers and the polity at large, our Free Trade Proposition shows that with GC superadditivity in the *policy makers'* objective functions, free trade – which maximizes aggregate welfare under neoclassical conditions – is the unique equilibrium outcome.

As discussed in the introduction, economists disagree about the merits of restricting trade negotiations to multilateral bargaining. Some prominent economists, Bhagwati (1991) among them, hold the view that preferential trade agreements that do not include all countries are detrimental to the achievement of worldwide free trade. We interpret this position as an objection to sequential bargaining in the formation of trade agreements. Using Bhagwati's terminology, preferential trade agreements are “stumbling blocs” rather than “building blocs” on the way to global free trade. We argue in this section that this is not necessarily the case. Without GC superadditivity partial agreements may serve as stumbling blocs *or* building blocs of global free trade. Which case applies depends on identifiable features of the world economy.

We have developed two examples to illustrate these points (see Appendix B in Aghion, Antràs and Helpman, 2006, for details). In the first example we construct a world in which multilateral negotiations lead to free trade while sequential negotiations lead to an FTA between two countries only. Nevertheless, the leader prefers sequential negotiations. This is the sense in which the availability of partial agreements may prevent the attainment of free trade. In this sort of world a rule that prohibits partial agreements and forces the countries to engage in multilateral bargaining leads to global free trade. In the second example we construct a world in which multilateral negotiations are doomed to fail, i.e., they do not lead to global free trade. Yet sequential bargaining does lead to global free trade, as the leader offers an FTA first to one follower country and afterwards induces the second follower country to join. In this case, rules that restrict trade negotiations to multilateral bargaining harm the prospects for global free trade.

In both examples we use aggregate profits as a country's payoff. This political-objective function arises in the Grossman and Helpman (1994) framework when policy makers attach zero weight to aggregate welfare and the ownership of sector-specific inputs is highly concentrated.⁴¹ But we also provide a general analysis of stumbling bloc and building bloc equilibria,

⁴¹Many studies use political-objective functions that attach differential weights to producer and consumer surplus. See, for example, Laffont and Tirole's (1993) analysis of economic regulation.

which does not depend on the precise reasons for the failure of GC superadditivity. And this analysis identifies conditions under which stumbling bloc or building bloc equilibria are more likely to exist.

7.1 Stumbling Blocs

In Example 3 in Appendix B of Aghion, Antràs and Helpman (2006), there is an outside good y with constant marginal utility, which serves as numeraire, and a product x with diminishing marginal utility; both are traded internationally. The utility function of country j is given by equation (3). Countries a and c import x in the initial equilibrium; both impose import tariffs; and the tariff is higher in a . Goods in sector y are freely traded. Every country seeks to maximize its profits in sector x .⁴²

Under these circumstances global free trade leads to a higher international price of x , because the removal of tariffs raises import demand in countries a and c . But the hike in the international price is not enough to compensate producers in a and c for the removal of the tariffs. As a result, profits decline in countries a and c and rise in country b . But the increase in profits in b more than compensates for the decline in profits in a and c , leading to a rise in aggregate world profits. Therefore multilateral bargaining leads to the formation of the grand coalition and to global free trade.

When c chooses sequential bargaining, it finds that its payoff is highest when it approaches country a first. But an FTA between countries a and c leads to reduced protection, with the price of x in a declining to the international price times the rate of protection in c . This, in turn, leads to an increase in aggregate world demand, thereby bidding up the international price. As a result profits rise in b and c and decline in a , because the rise in the international price does not compensate producers in a for the fall in the rate of protection. The resulting aggregate world profits exceed aggregate world profits under global free trade. In this event country c 's payoff is higher from forming a free trade area with a only than from forming a free trade area with a and b . For this reason the equilibrium in the sequential bargaining subgame consists of an FTA between a and c only. Country c prefers the outcome of the sequential subgame to the outcome of the multilateral subgame. In this event sequential bargaining is a stumbling bloc to worldwide free trade

The key features of this example, which are general requirements for a stumbling bloc equilibrium in which c forms an FTA with a only, are the following:⁴³ Multilateral bargaining leads to the formation of the grand coalition, therefore $W(abc) - W(a) - W(b) > W(c)$. But, c prefers an FTA with a to multilateral bargaining, i.e., $W(ac) - W(a) > W(abc) -$

⁴²In particular, we assume preferences, technology and endowments such the following functional forms apply: $C_a(q) = 4 - q$, $C_b(q) = 3 - q$, $C_c(q) = 8 - q$, $X_a(q) = \frac{5}{4}$, $X_b(q) = \frac{5}{2} + q$, and $X_c(q) = \frac{17}{4} + q$. The tariff rates are $\tau_a = 2$, $\tau_b = 1$ and $\tau_c = 1.5$. See Aghion, Antràs, and Helpman (2006) for further details.

⁴³A symmetric set of conditions can be formulated for the case in which c forms an FTA with b .

$W(a) - W(b)$, and c has no incentive to attract b to the FTA with a , i.e., $W(ac) - W(a) > W(abc) - W(a) - W_F(b)$. These conditions hold if and only if

$$W(ac) + \min\{W(b), W_F(b)\} > W(abc) > \sum_{j=a,b,c} W(j).$$

Finally, c prefers to approach a first and b second in sequential bargaining, which requires

$$W(ac) > W(a) - W(b) + \max\{W(bc), W(abc) - W_F(a)\}.$$

It is evident from these inequalities that a large payoff $W(ac)$ and large and *positive* coalition externalities increase the likelihood that this sort of stumbling bloc equilibrium will emerge.

7.2 Building Blocs

In Example 4 in Appendix B of Aghion, Antràs and Helpman (2006), we also have two sectors, x and y , and preferences given by (3). Both goods are traded internationally and there are no impediments to trade in y . Countries a and b export x , and c has a tariff on imports of x . Every country seeks to maximize profits.⁴⁴

As in the previous example, free trade leads to an increase in c 's imports of x , thereby bidding up its international price. As a result, profits rise in countries a and b and decline in c , because the rise in the international price does not compensate producers in c for the removal of the tariff. In this case, however, the fall in c 's profits is larger than the rise in the joint profits of a and b . Therefore GC superadditivity fails, and c prefers the status quo to offers of an FTA that countries a and b will accept in the multilateral subgame. In this event multilateral bargaining does not lead to global free trade.

In the sequential subgame, c prefers to approach a first. An FTA between a and c leads to enhanced protection; namely, the producer price in a rises to the tariff rate in c times the international price. This raises the aggregate world supply of x and depresses its price. The lower international price hurts profits in b , but it raises profits in a , because the decline in the international price is smaller than the tariff rate in c . The new aggregate world profits are lower, however, than the profits under free trade. For this reason c has an incentive to bring b into the FTA. It follows that sequential bargaining leads to the formation of the grand coalition and to global free trade. Moreover, the equilibrium payoff to c exceeds c 's payoff from multilateral bargaining. Thus, in this example a restriction to multilateral bargaining does not lead to global free trade, yet the choice of sequential bargaining, which c prefers, leads to global free trade following a stepwise buildup of the FTA by adding a first and then

⁴⁴More specifically, in this case we assume $C_a(q) = 1 - 2q$, $C_b(q) = 15 - 2q$, $C_c(q) = 8 - q$, $X_a(q) = 2 + 2q$, $X_b(q) = 15 + q$, and $X_c(q) = 3$, as well as $\tau_a = 1$, $\tau_b = 1$, and $\tau_c = 1.5$ (see Aghion, Antràs, and Helpman, 2006, for details).

b. In this case the FTA between *a* and *c* is a building bloc to worldwide free trade.

The key features of this example, which are general requirements for a building bloc equilibrium in which *c* forms an FTA with *a* first, are the following:⁴⁵ First, country *c* prefers the status quo to the grand coalition in multilateral bargaining, i.e., $W(c) > W(abc) - W(a) - W(b)$. Second, *c* prefers *b* to join its FTA with *a*, which requires $W(abc) - W(a) - W_F(b) > W(ac) - W(a)$. Third, *c*'s payoff from sequential bargaining with *a* first and *b* second is higher than from multilateral bargaining, i.e., $W(abc) - W(a) - W_F(b) > W(c)$. Together these conditions hold if and only if

$$\sum_{j=a,b,c} W(j) > W(abc) > \max\{W(ac), W(a) + W(c)\} + W_F(b).$$

Finally, *c*'s payoff from approaching *a* first in the sequential subgame is higher than its payoff from approaching *b* first, or

$$W(abc) > W(a) - W(b) + \max\{W(bc), W(abc) - W_F(a)\} + W_F(b).$$

Evidently, these conditions are more likely to be satisfied the smaller are $W(ac)$ and $W(bc)$ and the more *negative* are the coalition externalities in *b*. In particular, *no such building bloc equilibrium exists when the coalition externalities are positive in the follower countries.*

We conclude from this discussion that stumbling bloc equilibria are more likely to exist the larger is the value of a bilateral FTA between *c* and one of the follower countries and the larger are the coalition externalities in the follower countries. Building bloc equilibria are more likely to exist the smaller are the values of bilateral FTAs between *c* and each one of the follower countries and the more negative are the coalition externalities in one follower country. Building bloc formation is furthered by an asymmetry in coalition externalities, i.e., a large negative externality in one follower country and a large positive externality in the other. *Without negative coalition externalities there are no building bloc equilibria.* Naturally, there can be equilibria that are neither stumbling nor building blocs of free trade.⁴⁶

8 On Alternative Features of the Bargaining Protocol

A major aim of this paper has been to show that free trade can be attained when countries may choose between multilateral and sequential bargaining, a choice that the WTO makes

⁴⁵A symmetric set of conditions can be formulated for the case in which *c* forms an FTA with *b*.

⁴⁶It has been suggested to us that a more realistic formulation of the bargaining game would allow for a reversion to sequential bargaining when multilateral negotiations fail. But in such a case the leader's payoff under multilateral bargaining would always coincide with that under sequential bargaining, and the choice between these two negotiation protocols would become indeterminate. Interestingly, however, stumbling and building bloc equilibria would still arise under the conditions discussed in the main text.

available to its member states. What our Free Trade Proposition provides is a set of sufficient conditions for the bargaining protocols and the objective functions of policy makers, i.e., GC superadditivity, which guarantee this outcome. Naturally, free trade may not emerge as an equilibrium outcome in other polities in which GC superadditivity fails, as discussed in the previous section, or when countries use different bargaining protocols.

As an example consider Saggi and Yildiz (2006). They construct an economy with reciprocal dumping a la Brander and Krugman (1983) in which there are two alternative bargaining protocols, which they term FTA and No FTA. As in our model there are three countries a , b and c , but no agenda setter and no transfers. In the FTA game, every country announces which of the other two countries it wants to form a free trade area with; the announcement can name one or both of the other countries. Then FTAs are formed among the countries that have announced each other. If, for example, a includes b in its announcement and b includes a , the two countries form an FTA. But if a includes b and b does not include a these two countries do not form an FTA. Alternatively, in the No FTA game, which represents multilateral bargaining, a country can only announce whether it is in favor or not of global free trade. If all three countries are in favor, there is free trade. If at least one country objects, the status quo prevails.

Evidently, these bargaining protocols differ from ours, and they lead to multiple equilibria even when the three countries are symmetric. Nevertheless, Saggi and Yildiz show that free trade is the unique stable (coalition proof) equilibrium in the No FTA game but not necessarily in the FTA game. In other words, there are conditions under which global free trade as well as a free trade agreement between two countries only are stable equilibria in the FTA game.⁴⁷

Interestingly, in the symmetric case (or small deviations from symmetry) their objective functions of the policy makers, which consist of aggregate welfare, are GC superadditive, despite the presence of markups and reciprocal dumping (see their Lemma 1). Moreover, their model exhibit *negative* coalition externalities (an FTA between any two countries hurts the third). Under these circumstances we can apply the Free Trade Proposition to our bargaining protocols. The implication is that when our bargaining protocols are applied to their symmetric world, then free trade is the unique equilibrium outcome and it is achieved via sequential bargaining.

As this discussion illustrates, details of the bargaining protocol can be important. For this reason we discuss in the rest of this section two modifications of our sequential bargaining protocol in order to explore how they affect the equilibrium outcomes. One modification concerns the continuation game after a rejection. We assumed that when c , the agenda

⁴⁷Saggi and Yildiz (2006) also analyze equilibria in asymmetric setups, in which two countries have identical marginal costs while the third country has higher marginal costs. Their model also delivers building and stumbling bloc equilibria.

setter, makes an offer to one of the follower countries which the follower country rejects, the game ends. In Section 8.1 we explore what happens if under these circumstances c can proceed to make an offer to the other follower country. In Section 8.2, we consider a different modification. We assumed that c can commit to a payoff when it makes an offer to a follower country, where this payoff is independent of the final outcome of the game. This hinges on two issues: whether transfers from c to a follower country can be contingent on the outcome of the game and whether c can make firm commitments. We shall examine the roles of these features in the bargaining protocol.

8.1 Rejection does not end the game

Consider a modification of the sequential bargaining game in which a rejection in the first stage does not end the game. Instead, after a rejection, country c can proceed to make an offer to the next follower country. We call it the S2 game. The rest of the game is the same as in the original formulation.

Does the Free Trade Proposition hold under these circumstances? The answer is No, yet a number of important elements of this proposition, such as the role of coalition externalities in the choice between sequential and multilateral bargaining, remain similar. A major implication of this modification is that GC superadditivity does not guarantee free trade as the equilibrium outcome. It therefore dispels a possible conclusion from our Free Trade Proposition that whenever the agenda setter can make take-it-or-leave-it offers to the other countries, then GC superadditivity is sufficient for free trade to be attained. In other words, this finding suggests that the Free Trade Proposition requires more than GC superadditivity and an agenda setter who can make take-it-or-leave-it offers.

To understand the source of differences between S2 and the original bargaining protocol, consider sequential bargaining in which the agenda setter approaches a first and b second. In the original specification a rejection by a gave c the payoff $W(c)$. Under S2, a 's rejection gives the agenda setter $\max\{W(c), W(bc) - W(b)\}$. Why? Because after the rejection c can make b a small offer that b rejects, in which case c gets $W(c)$. But c can also make b an offer that b accepts. Since $W(b)$ is the smallest offer that b accepts, this strategy gives c the payoff $W(bc) - W(b)$. If $W(bc) - W(b) > W(c)$ then c prefers to form a coalition with b .

Next note that if $W(bc) - W(b) > W(c)$ then the agenda setter has to make a an offer $W_F(a)$ in the first stage for a to accept, because if a rejects the offer then c will form a coalition with b . But if $W(bc) - W(b) < W(c)$ then c is expected not to form a coalition with b in case a rejects the offer, in which case c has to offer a the payoff $W(a)$ for a to accept. It follows from this reasoning that if

$$\max_{z \in \{a, b\}} W(zc) - W(z) < W(c) \tag{5}$$

then the agenda setter never forms an FTA in the second stage of the game, after its offer has been rejected in the first stage. Under these circumstances our Free Trade Proposition holds not only with our original sequential bargaining game but also with S2.

The question that arises is therefore whether there are circumstances in which sequential bargaining under S2 and GC superadditivity can tempt the agenda setter to form a coalition with one country only. We show in Appendix C of Aghion, Antràs, and Helpman (2006) that under S2 and GC superadditivity, free trade is *not* an equilibrium outcome if and only if the following two conditions are satisfied:

$$\max_{z \in \{a,b\}} W(zc) - W(z) > W(abc) - \min\{W(a) + W(b), W_F(a) + W_F(b)\}, \quad (6)$$

$$\min_{z \in \{a,b\}} W(zc) - W(z) > W(c). \quad (7)$$

These conditions have an important implication: they require the follower country z^* , where $z^* = \arg \max_{z \in \{a,b\}} W(zc) - W(z)$, to have positive coalition externalities and the other follower country to have negative coalition externalities. As a result free trade is the equilibrium outcome when either both follower countries have positive coalition externalities or both have negative coalition externalities.⁴⁸

We also show in Appendix C of Aghion, Antràs, and Helpman (2006) that the agenda setter chooses multilateral bargaining when both countries have positive coalition externalities and sequential bargaining when both countries have negative coalition externalities. In either one of these cases free trade results. Moreover, the agenda setter chooses sequential bargaining when the *aggregate* coalition externalities are negative, i.e., $W(a) + W(b) > W_F(a) + W_F(b)$. The last condition is obviously satisfied when the coalition externalities are negative in each one of these countries, but it also can be satisfied when they are positive in one country and negative in the other. It follows that negative coalition externalities favor sequential bargaining while positive coalition externalities favor multilateral bargaining. This is the sense in which coalition externalities play a similar role in the S2 bargaining protocol and in the original bargaining protocol.

What explains the failure of free trade under S2 and GC superadditivity when conditions (6) and (7) are satisfied? The intuition is as follows. The agenda setter's payoff after a rejection of an offer in the first round is now affected by what happens later in the game. If one follower country has sufficiently high positive coalition externalities and the other one has sufficiently high negative coalition externalities, then it may be too costly for c to induce them both to join the coalition. When first approaching the country with positive coalition

⁴⁸Note that in a world of reciprocal dumping we discussed above, where countries are symmetric, the equilibrium outcome is sequential bargaining with free trade because this economic environment satisfies GC superadditivity and the coalition externalities are negative. It follows that in this economic environment S2 and the original bargaining protocol yield the same outcomes.

externalities, say country a , this country demands a high payoff $W_F(a)$ in order to join the coalition, because it expects c to form a coalition with b when its offer is rejected (this is ensured by our necessary and sufficient conditions). On the other hand, when c approaches the negative-externality country first, which is country b , then once b accepts the offer it becomes expensive to attract country a to the coalition in the second stage because it has to be paid again $W_F(a)$. So either way the cost of forming the grand coalition via sequential bargaining is high, because the country with positive coalition externalities has to be paid $W_F(a)$. In the original protocol this could be avoided by approaching country a with positive externalities first paying it $W(a) < W_F(a)$, and then approaching country b with negative externalities paying it $W_F(b) < W(b)$. Under S2 this is not possible, which makes the formation of the grand coalition more costly.

It also follows that under S2, GC superadditivity and conditions (6)-(7), sequential bargaining is a stumbling bloc on the way to free trade. Moreover, when conditions (6)-(7) are not satisfied there is no equilibrium in this game in which sequential bargaining is a building block on the way to free trade. This contrasts with our findings in the previous section, where we examined lack of GC superadditivity; there sequential bargaining could be a stumbling or a building block of free trade.

8.2 Lack of commitment and non-contingent transfers

We now go back to the original bargaining protocol in which the rejection of an offer in the first stage of sequential bargaining ends the game. In this formulation we assumed that the agenda setter can make offers that secure fixed payoffs to the follower countries. This is without loss of generality in the last stage of the game, because at this stage the offer and outcome are both final. It raises, however, the question of what is required to justify this assumption in the first stage.

For concreteness consider the case of negative coalition externalities in b and positive coalition externalities in a , i.e., $W_F(b) < W(b)$ and $W_F(a) > W(a)$, and the sequential subgame in which c approaches a first and b second. Then, we have argued, the agenda setter offers a the payoff $W(a)$ which a accepts, followed by an offer $W_F(b)$ to b which b accepts. Under GC superadditivity this is an equilibrium of this subgame, giving c the payoff $W(abc) - W(a) - W_F(b)$. Evidently, it is assumed in this specification that c can commit in the first stage of the game to secure payoff $W(a)$ for country a at the end of the game.

To understand what is involved, let $W_z(abc)$ be the payoff of country z in the absence of transfers when the grand coalition forms, and let $W_z(jk)$ for $z = j, k$ be the payoff of country z in the absence of transfers when the coalition jk forms. In this event

$$\sum_{z=a,b,c} W_z(abc) = W(abc),$$

$$\sum_{z=j,k} W_z(jk) = W(jk).$$

Then there are two ways to support the equilibrium in the above described subgame. First, country c can commit in stage one to a transfer $T_a = W(a) - W_a(abc)$ to country a at the end of the game if a joins the coalition and a grand coalition forms, and a transfer $T_a = W(a) - W_a(ac)$ if country b does not join the coalition. With these commitments country a gets a payoff $W(a)$ when it agrees to join the coalition independently of whether b joins. In response a accepts the offer and c has the incentive to offer b a transfer $T_b = W_F(b) - W_b(abc)$ for joining the coalition, which b accepts.

A second way to support the above equilibrium is by having the agenda setter offer a a transfer $T_a = W(a) - W_a(abc)$ in the first stage of the game if a joins the coalition coupled with a commitment by c to make an acceptable offer to b in stage two. Under the circumstances, a accepts the offer and c has the incentive to offer b a transfer $T_b = W_F(b) - W_b(abc)$ for joining the coalition, which b accepts. Note that in both cases the strategies work because the agenda setter can make credible commitments; without credible commitments a promise of $W(a)$ in the first stage is not credible. In addition note that, in the first case, a is offered contingent transfers; one transfer when the grand coalition forms and another transfer when it does not. In the second case the transfer can be made upfront, because it is not contingent on the resulting coalition structure. But the commitment to include country b in the coalition has to be credible for a to receive the payoff it expects.

This discussion clarifies an important point: when the agenda setter can *commit* to a course of action it can use non-contingent transfers to support our equilibrium.⁴⁹ The remaining issue we discuss in this section is what happens when contingent transfers are not available and the agenda setter has no ability to commit to a course of action. As is well known, in such circumstances c may be tempted to behave opportunistically, and the question becomes what are the consequences of such opportunistic behavior.

To this end we now consider a modified version of the sequential subgame, call it S3. Under this protocol the agenda setter has no commitment power. As a result it offers non-contingent transfers and it cannot commit in stage one to a course of action in the second stage. The rest of the game is the same as before.

Let us reconsider the sequential subgame a, b under S3 when country b has negative coalition externalities and country a has positive coalition externalities. To support the equilibrium described above the agenda setter has to offer a the non-contingent transfer $T_a = W(a) - W_a(abc)$. If a accepts the offer then c has an incentive to expand the coalition

⁴⁹We did not specify the payoff to c when b rejects the offer to join the coalition after c committed to a to bring b into the coalition. The reason is that this payoff does not matter because c makes b a take-it-or-leave-it offer and b has the incentive to accept every offer $T_b \geq W_F(b) - W_b(abc)$.

to include b if and only if

$$W(abc) - W(ac) - W_F(b) \geq W_a(abc) - W_a(ac). \quad (8)$$

In other words, GC superadditivity is no longer sufficient to ensure c 's desire to expand the free trade area. Why? Because if c does not expand the FTA it gets a payoff $W_c(ac) - T_a$ and if it expands the FTA it gets a payoff $W_c(abc) - T_a - T_b$. Therefore c is willing to include b in the coalition if and only if $W_c(abc) - T_b \geq W_c(ac)$. But since $T_b = W_F(b) - W_b(abc)$, we obtain (8). In case (8) does not hold, which by GC superadditivity implies $W_a(abc) > W_a(ac)$, the agenda setter does not want to expand the free trade area. Then a 's payoff is

$$W_a(ac) + T_a = W(a) + W_c(ac) - W_c(abc) < W(a).$$

Anticipating this outcome a rejects the non-contingent transfer $T_a = W(a) - W_a(abc)$ and our equilibrium cannot be supported. In short, when $W_a(abc)$ is much larger than $W_a(ac)$ promises of c to a to bring country b into the coalition are not credible.

It follows from this discussion that when (8) holds and a similar condition holds for country b then our Free Trade Proposition remains valid under S3, and when these two conditions on preferences are not satisfied other equilibrium outcomes are possible. We show in Appendix C of Aghion, Antràs, and Helpman (2006) that when the coalition externalities are positive in both follower countries and GC superadditivity holds then multilateral bargaining is the equilibrium outcome under S3 and free trade prevails. This is similar to our Free Trade Proposition. It follows that in this case too negative coalition externalities are needed for sequential bargaining to be an equilibrium outcome. But unlike the original bargaining protocol, under S3 negative coalition externalities may result in sequential bargaining with free trade or sequential bargaining with a partial coalition.

9 Concluding Comments

We have developed a dynamic model of bargaining with transferable utility in order to evaluate the relative merits of multilateral and sequential trade negotiations. An evaluation of this sort is needed to assess the articles of agreement of the WTO. We believe that an explicit modelling of the bargaining process is necessary for this purpose. Although we recognize the limitations of our model, which may be too simple for the task at hand, we also feel that it provides valuable insights into these issues. In particular, it identifies superadditivity and coalition externalities in the structure of payoffs as important determinants of the relative performance of these bargaining protocols. True, the nature of these influences may vary with the bargaining procedure, but we believe that superadditivity and coalition externalities

are important in every realistic bargaining procedure.

Our main findings can be summarized as follows: First, in the absence of coalition externalities and under GC superadditivity global free trade is the unique equilibrium outcome, independently of whether preferential trade agreements are allowed or forbidden. This result extends to allowing for positive or negative coalition externalities under our assumptions that: (i) a rejection in the first stage under sequential bargaining prevents the leader country from making other offers in the second stage; (ii) the leader country can commit to a given course of action or to contingent transfers in advance. Therefore, although we have discussed only coalitions that consist of free trade areas, it is evident that our results also apply to customs unions, economic unions, and other forms of trade liberalization. This follows from the fact that the leading country is a residual claimant on the surplus from global free trade, and that it has the ability to compensate other countries for the abandoning of suboptimal agreements.

Second, under GC superadditivity the leading country strictly prefers sequential bargaining when coalition externalities are negative in at least one follower country. The reason is that in this case the leading country gains more from first forming an FTA with the country that has the higher coalition externalities and then expanding the free trade area to encompass all countries, than from making simultaneous offers to all follower countries. The key is that once an FTA exists, it is cheap to “buy” a country with negative coalition externalities. Conversely, the leading country prefers multilateral bargaining when the coalition externalities of the follower countries are positive, because then sequential bargaining makes it expensive to “buy” outside countries.

Third, when payoffs are not GC superadditive, global free trade may not occur in equilibrium. We showed that in this event preferential trade agreements facilitate the achievement of global free trade when coalition externalities are negative, and hamper it when coalition externalities are positive. Evidently, the structure of coalition externalities is an important determinant of which bargaining protocol secures global free trade.

We illustrated these conclusions with a simple competitive model of international trade in which global free trade is Pareto-efficient. The model clarifies the sources of coalition externalities. They are related to trade structure and the structure of protection. They also depend on features of demand and supply in each country. A free-trade agreement removes tariffs on trade between members of the FTA, whereas FTA members maintain their original rates of protection vis à vis outside countries. In this model, coalition externalities stem from the impact of FTAs on world prices, which affect the welfare or political objectives of the trading partners. Coalition externalities on welfare-maximizing nonmember countries tend to be negative when an FTA reduces the prices of their exportables, and positive when an FTA reduces the prices of their importables. If every country’s negotiators maximize aggregate welfare and trade taxes are the only distortions, then GC superadditivity holds and global free trade is attained in equilibrium. If, however, special interests induce country negotiators

to maximize a politically-motivated objective function, then GC superadditivity may fail to hold and preferential trade agreements can be either building blocs or stumbling blocs to free trade, as explained above.

Finally, we showed how stumbling blocs can also emerge when; (i) following a rejection in the first stage, the agenda setter can still make an offer in the second stage, or; (ii) the agenda setter cannot commit to a fixed payoff to the first country which is independent of what happens in the second stage. Even in those cases we found, however, that negative externalities tend to favor sequential bargaining while positive externalities tend to favor multilateral bargaining.

Our model of trade negotiations has special attributes, yet our methodology of using dynamic bargaining models of coalition formation to study trade negotiations can be explored and pursued in several interesting directions. First, we have examined cases in which the agenda setter is predetermined. Yet agenda setting power can be related to a country's characteristics, such as its economic size, the level of its technology, financial development and the like. Under these circumstances the characteristics of the leader will be correlated with the coalition externalities that its FTA imposes on nonmember countries. Second, it would be interesting to study the lack of GC superadditivity that emanates from different sources. Alternative sources of nonadditivity can be lack of competition, distortions in labor markets, or institutional constraints on economic transactions. It would be interesting to understand how these different reasons for lack of superadditivity impact the building and stumbling bloc effects of preferential trade agreements. Third, our model disregards geography. Many of the preferential trade agreements are regional, however. What features of geography drive this bias? And do these features affect the choice between sequential and multilateral bargaining? Fourth, we have assumed that the players' payoff functions are common knowledge. A more realistic formulation might involve the agenda setter facing some level of uncertainty about the follower countries' payoff functions. We believe that such asymmetric information might influence our results in interesting ways.⁵⁰

Finally, our analysis abstracts from issues related to bargaining costs. Typically, these costs should depend upon whether negotiations take place sequentially or multilaterally. Yet it is not obvious which of the two would be larger. One may argue that multilateral negotiations skip bargaining stages and thus reduce the delay in reaching free trade. On the other hand, bargaining costs may increase with the number of agents simultaneously involved in the bargaining, which plays against multilateral bargaining. Future research will clarify how these considerations interact with coalition externalities and properties of objective functions in shaping the outcomes of trade negotiations.

⁵⁰Indeed, we have worked out an example in which the leader views the follower countries' payoffs as uniformly distributed random variables. We found that, with this type of asymmetric information, negative externalities may promote the emergence of stumbling bloc equilibria.

References

- Aghion, P., Antràs, P., Helpman, E., 2006. Negotiating free trade. National Bureau of Economic Research Working Paper No. 10721.
- Bagwell, K., Staiger R.W., 2002. The Economics of the World Trading System. MIT Press, Cambridge, MA.
- Bagwell, K., Staiger R.W., 2004a. Backward stealing and forward manipulation in the WTO. National Bureau of Economic Research Working Paper No. 10420.
- Bagwell, K., Staiger R.W., 2004b. Multilateral trade negotiations, bilateral opportunism and the rules of GATT/WTO. *Journal of International Economics* 63, 1-29.
- Baldwin, R.E., 1996. A domino theory of regionalism, in: Baldwin, R.E., Haaparanta, P.J., Kiander, J. (Eds.), *Expanding the Membership of the EU*. Cambridge University Press, Cambridge, UK.
- Bhagwati, J.N., 1991. *The World Trading System at Risk*. Princeton University Press, Princeton.
- Bhagwati, J.N., 1993. Regionalism and multilateralism: An Overview, in: De Melo, J., Panagariya A. (Eds.), *New Dimensions in Regional Integration*. Cambridge University Press, Cambridge, UK.
- Bhagwati, J.N., Brecher, R.A., Hatta T., 1983. The generalized theory of transfer and welfare (I): bilateral transfers in a multilateral world. *American Economic Review* 73, 606-618.
- Burbidge, J.B., DePater, J.A., Myers, G.M., Sengupta, A., 1997. A coalition-formation approach to equilibrium federations and trading blocs. *American Economic Review* 87, 940-956.
- Bond, E.W., Syropoulos, C., 1996. The size of trading blocs market power and world welfare effects. *Journal of International Economics* 40, 411-437.
- Brander, J.A., Krugman, P.R., 1983. A 'reciprocal dumping' model of international trade. *Journal of International Economics* 15, 313-323.
- Carraro, C., Siniscalco, D., 1993. Strategies for the international protection of the environment. *Journal of Public Economics* 52, 309-328.
- CEPR, 1995. *Flexible Integration: Towards a More Effective and Democratic Europe, Monitoring European Integration 6*. Center for Economic Policy Research, London.

- Chang, W., Winters, L.A., 2002. How regional blocs affect excluded countries: the price effects of MERCOSUR. *American Economic Review* 92, 889-904.
- Deardoff, A.V., Stern, R.M., 1994. Multilateral trade negotiations and preferential trading arrangements, in: Deardoff, A.V., Stern, R.M. (Eds.), *Analytical and Negotiating Issues in the Global Trading System*. The University of Michigan Press, Ann Arbor.
- Ethier, W.J., 1998. Regionalism in a multilateral world. *Journal of Political Economy* 106, 1214-1244.
- Frankel, J.A., Stein, E.H., Wei, S-J., 1996. Regional trading agreements: natural or super-natural? *American Economic Review* 86, 52-56.
- Furusawa, T., Konishi, H., 2005. Free trade networks with transfers. *Japanese Economic Review* 56, 144-164.
- Gomes, A.R., 2005. Multilateral contracting with externalities. *Econometrica* 73, 1329-1350.
- Grossman, G.M., Helpman, E., 1994. Protection for sale. *American Economic Review* 84, 833-850.
- Grossman, G.M., Helpman, E., 1995. The politics of free trade agreements. *American Economic Review* 85, 667-690.
- Grossman, G.M., Helpman, E., 2001. *Special Interest Politics*. MIT Press, Cambridge, MA.
- Jackson, J.H., 1997. *The World Trading System: Law and Policy of International Economic Relations*, Second Edition. MIT Press, Cambridge, MA.
- Jones, R.W., 1970. The transfer problem revisited. *Economica* 37, 178-184.
- Kemp, M.C., Wan Jr., H.Y., 1976. An elementary proposition concerning the formation of customs unions. *Journal of International Economics* 6, 95-97.
- Krishna, P., 1998. Regionalism and multilateralism: a political economy approach. *Quarterly Journal of Economics* 113, 227-251.
- Krugman, P.R., 1993. Regionalism versus multilateralism: analytical notes, in: De Melo, J., Panagariya A. (Eds.), *New Dimensions in Regional Integration*. Cambridge University Press, Cambridge, UK.
- Laffont, J-J., Tirole J., 1993. *A Theory of Incentives in Procurement and Regulation*. MIT Press, Cambridge, MA.

- Levy, P.I., 1997. A political-economic analysis of free-trade agreements. *American Economic Review* 87, 506-519.
- Maskin, E.S., 2003. Bargaining, coalitions, and externalities. Mimeo Princeton University.
- Panagariya, A., 2000. Preferential trade liberalization: the traditional theory and new developments. *Journal of Economic Literature* 38, 287-331.
- Ray, D., Vohra R., 1997. Equilibrium binding agreements. *Journal of Economic Theory* 73, 30-78.
- Ray, D., Vohra R., 1999. A theory of endogenous coalition structure. *Games and Economic Behavior* 26, 286-336.
- Richardson, M.D., 1993. Endogenous protection and trade diversion. *Journal of International Economics* 34, 309-324.
- Riezman, R.G., 1985. Customs unions and the core. *Journal of International Economics* 19, 355-366.
- Saggi, K., Yildiz, H.M., 2006. Bilateral trade agreements and the feasibility of multilateral free trade. Mimeo Southern Methodist University.
- Samuelson, P.A., 1952. The transfer problem and transport costs, I: the terms of trade when impediments are absent. *Economic Journal* 62, 278-304.
- Samuelson, P.A., 1954. The transfer problem and transport costs, II: analysis of effects of trade impediments. *Economic Journal* 64, 266-289.
- Summers, L.H., 1991. Regionalism and the world trading system, in: *Policy Implications of Trade and Currency Zones*. Federal Reserve Bank, Kansas City.
- Winters, L.A., 1997. Regionalism and the rest of the world: theory and estimates of the effects of European integration. *Review of International Economics* 5, 134-147.
- Yi, S-S., 1996. Endogenous formation of customs unions under imperfect competition: open regionalism is good. *Journal of International Economics* 41, 153-177.

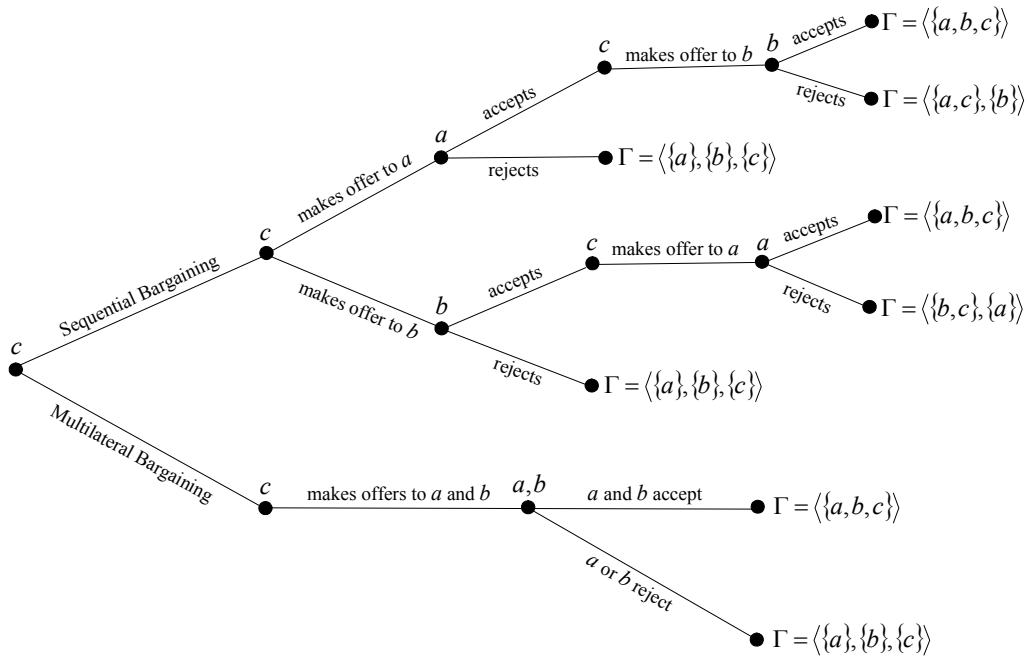


Figure 1: Game tree

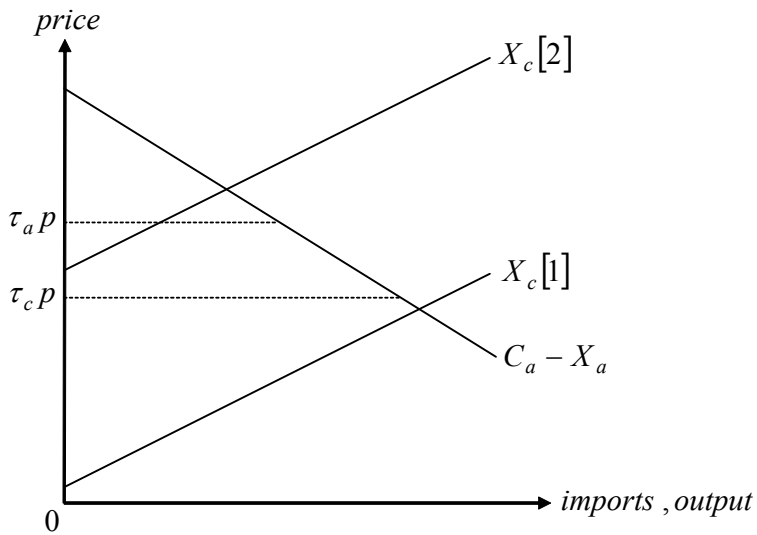


Figure 2: Coalition externalities

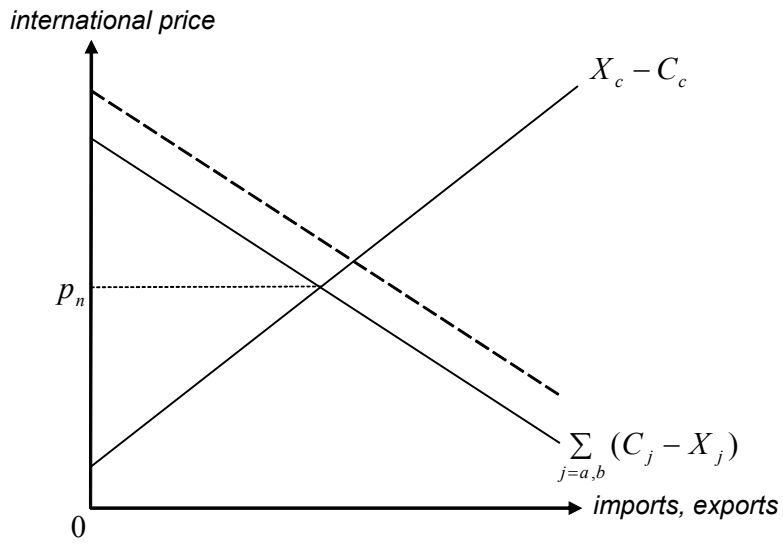


Figure 3: FTA raises b 's import price

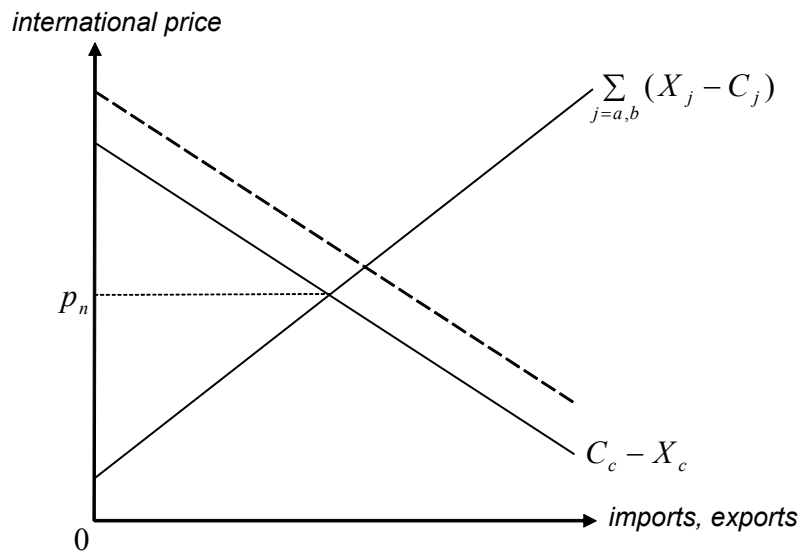


Figure 4: FTA raises b 's export price