The New Regionalism and the Threat of Protectionism

Andrew Hughes Hallett
Carlos A. Primo Braga

A multilateral trade system inhibits noncooperative behavior among trading blocs. The successful conclusion of the Uruguay Round extended and deepened the network of variables covered by multilateral rules. For developing countries, a working (even imperfect) multilateral trade system remains the best hope against excesses by those with market power.
Summary findings

Drawing on game theory concepts, Hallett and Primo Braga discuss why countries form themselves into trading blocs and what the relations between these blocs are likely to be.

They identify three types of trade regime:

- Unilateral trade policies — which are noncooperative.
- Multilateral agreements (such as the GATT) — which are cooperative.
- Coalitions (regional integration arrangements or minilateral agreements) — which are mixed (cooperative internally and noncooperative externally).

They argue that regional integration arrangements can work better than global rules as precommitment devices for internally cooperative policies because they create a denser network or interlinked policy targets. The losses for a participant ostracized (or disciplined) by his bloc are immediate and tangible.

Crucial to the results of analysis is the external policy stance adopted by each bloc after it has formed. External relations will determine whether regional blocs are welfare-improving, consistent with the aims of the GATT, and a vehicle for securing commitments to the regime; or whether they will become a vehicle for spreading “political economy biases.”

Should higher or lower external barriers be expected for nonmembers? That depends on how large the benefits or costs, in trade and investment creation (or diversion), would be to members if the move to free trade within the bloc is not accompanied by any increase in the bloc’s external barriers (an “open” bloc).

Widening tends to be easier the more open a bloc is, since insiders are less concerned with the erosion of their preferences.

In the alternative scenario, lower intra-bloc trade and investment barriers are accompanied by an increase in the external barriers, giving any specific set of potential participants strong incentives to join (a “closed” bloc). “Deepening” by expanding the list of variables covered by the trade agreement also tends to make the bloc more cohesive. In both cases — a closed bloc or deep integration — greater cohesion is obtained at the cost of increasing the costs of entry for nonmembers.

The hope that regional integration arrangements can pave the way for global free trade is unrealistic. As regional integration arrangements enlarge, they may be better off exerting market power against outsiders rather than following a globally cooperative path. Inter-bloc trade relations will ultimately depend on how effective special interest groups are at distorting bloc-wide trade policies that suit their interests. A multilateral trade system inhibits noncooperative behavior among trading blocs. The successful conclusion of the Uruguay Round extended and deepened the network of variables covered by multilateral rules. For developing countries, a working (even imperfect) multilateral trade system remains the best hope against excesses by those with market power.

This paper — a product of the International Trade Division, International Economics Department — is part of a larger effort in the department to understand new regionalism in trade policy. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Anna Kim, room R2-042, extension 33715 (36 pages). August 1994.
THE NEW REGIONALISM AND
THE THREAT OF PROTECTIONISM

by

Andrew Hughes Hallett
and
Carlos A. Primo Braga
THE NEW REGIONALISM AND THE THREAT OF PROTECTIONISM

Andrew Hughes Hallett

Carlos A. Primo Braga

1. INTRODUCTION

It has become fashionable to assert that regionalism is not only back, but that this time it is here to stay (Bhagwati, 1992). The fact that preferential trade agreements affect a significant share of global trade flows is not a new development. Recent regional integration arrangements (RIAs), however, have attracted a great deal of interest because they are broader in scope than previous experiments in integration and because they encompass North-South reciprocal arrangements. Moreover, there is a growing perception that RIAs are becoming the policy regime with the greatest influence on trade flows.

What are the incentives which lead countries to form themselves into trading blocs? Once we have discussed these incentives, we can say something about the welfare implications (and likely outcomes) of a world of regional trading blocs, as compared to a world of global (the GATT regime) or non-cooperative (unilateral) trading arrangements. We build this analysis up from a game theory basis and introduce the new trade theory as a particular version of this kind of argument. That makes it easier to see what the new trade theory has to say about the benefits of liberalizing trade on a regional basis.

1 A. Hughes Hallett is a professor at Strathclyde University and Visiting Professor at Princeton University. Carlos A. Primo Braga is a senior economist with the International Trade Division, International Economics Department, World Bank. Comments by K. Anderson, J.M. Finger, P. Low and W. Martin are gratefully acknowledged. The findings, interpretations and conclusions are the authors' own. They should not be attributed to the World Bank, its Board of Directors, its management, or any of its member countries.

2 Trading blocs proliferated between the wars even though the grouping was rather different then (being based on empire or spheres of influence) and the "contractual" arrangements were less formal than at present (Eichengreen and Irwin, 1993). The demonstration effect of the "Treaty of Rome" and the pursuit of import-substitution industrialization led to numerous regional initiatives among developing countries in the 1960s. By 1988, more than 40 per cent of world trade was occurring within regions under preferential conditions (Primo Braga and Yeats 1992).

3 By "new trade theory" we refer to the body of literature that over the last two decades has brought imperfect competition to the forefront of discussions about trade policy. For a survey of this literature see Helpman and Krugman (1989).
If we can say that there are obvious incentives for forming regional blocs, compared to a regime of unilateral trade initiatives or an incomplete and imperfectly functioning GATT system, we still have to ask what the trading arrangements between the blocs are likely to be. If they are non-cooperative, then we have to balance the gains from freer trade internally against any potential losses from less trade between blocs, in order to determine whether this regime is a Pareto improvement over the current situation and also to determine whether the losses against the Pareto optimum of free trade everywhere (which participants may not wish to risk committing themselves to, if they think others may be tempted to default) are acceptably small. The incentives to commit may be higher within regional blocs because of the opportunity to "deepen" the agreement by linking a number of targets into one cooperative regime.

But if the process of forming blocs is more likely either to leave the existing barriers as they are or to lower them (i.e., promote a more cooperative trade regime), then we must ask whether allowing countries to form trade blocs and then lower the barriers between blocs — a series of mini-GATTs in which nonmembers are always welcome to join — might not be an easier way of achieving the ideal of free trade than negotiating multilateral arrangements (and sanctions) within a global GATT framework.

In section II, we discuss what the new trade theory has to say about regionalism. Special attention is given to the incentive structure that could explain the recent drift towards regionalism. The paper goes on to discuss what are the implications of the new regionalism for the GATT system (section III) and the role of special interests in shaping trading blocs (section IV). It ends with a brief analysis of the potential implications of the new regionalism for developing countries.

II. THE NEW TRADE THEORY AND RIAs

In a recent survey, Gunter (1989) identified three main phases in the literature on RIAs. The first phase, inaugurated by Jacob Viner's seminal contribution (Viner, 1950), focused on the welfare effects.

---

4 We assume that governments try to maximize their political support, as well as social welfare. Voters (as consumers) face a free-rider problem in supporting free-trade. Producer interests, in turn, tend to be more concentrated and more effective in organizing coalitions to foster their objectives (e.g., protection in the case of import-substituting industries). Governments involved in negotiations to form RIAs are more likely to respond to concerns about the trade-creation aspects of these arrangements (and their implications for vocal special interests in the member countries such as producers and labor) than to requests to minimize their trade-diverting consequences (reflecting the interest of producers in non-member countries). For further details see Hirschman (1981) and Krueger (1992).
of preferential trade arrangements.\(^5\) The second phase addressed the more fundamental question of why customs unions are formed, since it can be shown that unilateral non-preferential liberalization is a superior alternative in terms of resource allocation under certain assumptions.\(^6\) The contemporary phase of research, in turn, is identified as a composite of efforts trying to advance our understanding of the effects of RIAs on economic development and terms of trade, as well as the role of economies of scale in RIAs. A summary of the different types of RIAs is presented in Box 1. In this section, we re-visit the two first phases of research on RIAs – to use Gunter’s taxonomy – relying on a game-theoretic version of the new trade theory.

---

**Box 1: The Different Meanings of Regionalism**

The term "regionalism" is used to characterize a broad array of different economic and political processes. Typically, it is used to identify the formation of a supra-national economic region (i.e., a region in which governmental policies at the national level play a diminishing role in constraining international exchange among participating countries). In other circumstances, however, the term is used as a synonym of national disintegration (e.g., the splintering of the former Socialist Federal Republic of Yugoslavia).

Economic regionalism can be either market-driven or policy-driven. In the first case, the process is led by private enterprises, pursuing the gains from trade via the integration of productive networks, independently of preferential policies. In other words, market-driven regionalism (or regionalization) is the outcome of a natural locational phenomenon that promotes stronger economic ties within a region (Lorenz, 1992). Policy-driven regionalism (or minilateralism) is built around trade preferences and often complemented by investment preferences, efforts towards regulatory harmonization, and cooperation in the development of joint infrastructure projects.

RIAs are characterized by the fact that they entail reciprocal negotiations of trade preferences. There are, however, many modalities of RIAs. Participants can, for example, levy lower tariffs on each other's imports vis-a-vis those imposed on imports from third countries, forming a preferential trade area (PTA). Or they can adopt zero tariffs for the trade between the signatories of the arrangement, forming a free trade area (FTA). If countries in an FTA also adopt a common external trade policy against nonmembers, then the resulting RIA is characterized as a customs union (CU). Common markets (CMs), in turn, correspond to an agreement in which free trade for goods and services between member countries comes together with free movement of factors of production. Finally, a RIA may take the form of an economic union (EU), where in parallel with the formation of a common market, member countries explicitly pursue the harmonization of their micro and macroeconomic policies.

---

\(^5\) For a review of this phase see Lipsey (1960).

\(^6\) The literature associated with this phase is surveyed in Krauss (1972).
IL.1 The Terms of the Debate

(a) Trade Regimes. We distinguish three trade regime types: unilateral trade policies, global or multilateral agreements (i.e., the GATT) and coalitions (RIAs or minilateral agreements). They are non-cooperative, cooperative, and mixed (cooperative internally, noncooperative externally) respectively. Welfare analyses typically focus on the gains in performance, and their distribution, from pursuing liberalization through each one of these trade regimes. In the absence of any legal or constitutional requirement for countries to follow free-trade (or mechanisms for supranational income transfers), agreements involving cooperation must be self-enforcing. That means that each participant must enjoy gains compared to alternative trade regimes — not just that participants must gain on average. The issue then is, do regional solutions produce gains, either on average or for all, over unilateralism; and are they inferior (and by how much) to global regimes, or do the coalition members benefit at the expense of non-members? That allows us to evaluate the incentives for, and costs and benefits of, different trading regimes.

(b) Political Economy: Conventional trade-policy analyses assume decision makers to be immune to domestic political pressures. As Grossman and Helpman (1993a, b) point out, the case for free trade, and the impact of different tariff structures, have usually been examined on strictly economic criteria with non-economic distortions not being taken into account. To correct for that we have to build in a component which shows how domestic political-economy pressures may change the outcomes of different regimes. In that world private sector lobbies become explicit players with their own objectives for the decisions to be made. They create strategic interactions (cooperative or noncooperative) between the behavior of governments and lobbies — in which case we get a political economy model of the move towards regionalism or global arrangements — or between different governments, conditional on their lobbies' reactions — i.e., a political economy evaluation of "trade wars" vs. "trade talks." That allows us to analyze how the trade outcomes vary from the case were there is no pressure from special interest groups and whether RIAs reduce those distortions.

(c) Strategic Behavior: The question then is how to combine the political economy interactions with the strategic behavior of governments under the three regimes identified earlier? Because decision makers

7 See Magee, Brock and Young (1989) for a political economy analysis of how trade policy is influenced by special interests.
are now forced to play in two (possibly three) different areas of activity simultaneously — national economic performance, regional/global interactions, and an internal political influence game — we propose simultaneous or overlapping games as a method for trading-off outcomes in different arenas (as well as between players) and show the advantages and disadvantages of regionalism in particular. The overlapping vs. simultaneous game formulation corresponds to the "widening" vs. "deepening" distinction, which is (as we shall see) crucial to evaluating the outcomes of the different types of trade policies. These ideas are formalized in the appendix to this paper, with some illustrations of how the "deepening" vs. "widening" strategies fit into a formal game theory framework. The point made there is that the use of simultaneous/overlapping games allows these kinds of interactions to be analyzed within one framework. In particular they show that no new solution techniques are required. Instead everything turns on evaluating the links between the various players and their theaters of activity; that is, on identifying the structure of those links. The analysis therefore depends on how you structure the problem, rather than on how you choose to solve it.

(d) Institutions: To the extent that this trend towards regionalism is concerned with attempting to capture the benefits of greater cooperation, and because we are necessarily dealing with simultaneous interactions in different arenas, we must ask if the world’s existing economic institutions are appropriate. These institutions were founded to overcome macroeconomic adjustment difficulties, to relieve supply side constraints in individual countries, or to establish disciplines for narrowly defined trade policies. But the main challenges to international cooperation are now presented by the externalities which economic events or policy actions in one country impose on another. That is a very different thing to easing adjustment and supply-side problems, and leads naturally to the formation of "managed" partnerships. That in turn may require quite different institutions to manage commitment at the regional level, and perhaps, to regulate behavior between large blocs that have market power at a global level. At the regional level, the only example currently available is the European Community (EC, now the European Union) which explicitly deals with an interlinked network of different policy targets. In section III, we argue that some kind of coordinating supranational agency is needed to trade-off those targets while preserving the commitment to cooperation.

Moreover, the existing institutions are single target institutions; they have no objectives or competence outside their designated arena. Yet any new institutions must be concerned with facilitating cooperation and trading-off interactions between regimes and between arenas. The broad agenda of the Uruguay Round and the proposed World Trade Organization are to a certain extent explicit attempts to
address this problem. But as with any cooperation problem, there is the difficulty of demonstrating credible commitments by all players. We argue that the ability to precommit while pursuing a network of objectives is easier to accomplish at a regional level (e.g., by making the different countries' objectives interdependent through one regulating — and hence sanctions imposing — institution) than at multilateral level. These features can be analyzed through simultaneous or overlapping games as discussed below.

II.2 The Distinction Between "Open" and "Closed" Blocs

The crucial factor for the results of our analysis is the external policy stance adopted by each bloc after it has been formed. External relations will determine whether regional blocs will be welfare improving, consistent with the aims of GATT, and a vehicle for securing commitments to the regime; or whether they will become a vehicle for spreading "political economy biases." At a more mundane level, we have to ask whether higher or lower external barriers vis-a-vis nonmembers should be expected. That must turn on how large the benefits or costs would be to members in terms of trade and investment creation (or diversion) if the move to free trade within the bloc is not accompanied by any increase in the bloc's external barriers (an "open" bloc). Widening tends to be easier the more open a bloc is, since insiders will be less concerned with the erosion of their preferences. The alternative scenario is the case where lower trade and investment intra-bloc barriers are accompanied by an increase in the external barriers providing stronger incentives for any specified set of members to join (a "closed" bloc). Note that "deepening" by expanding the list of variables covered by the trading agreement also tends to make the bloc more cohesive. In both cases (a closed bloc or deep integration), greater cohesion is obtained at the cost of increasing the costs of entry for nonmembers.

Theoretical models show that forming a bloc may benefit members at the cost of nonmembers. But it is more likely to turn out that some participants will lose in some industries, or in the performance of some variables, but gain in others; and for some of them the cost may exceed the gains. For them, the bloc may have to be constructed so that those costs are reduced. And if that is done, it will certainly be at the expense of reducing any benefits which might have accrued to the nonmembers. Participants have to be sure that they all gain before they can commit themselves. Theory also suggests that the time

* At this stage we are abstracting of the possibility of retaliation by outsiders in response to the formation of a closed trading bloc. The issue of retaliation is addressed in section IV.

See, for example, Bond, Syropoulos and Winters (1993).
profile of these gains and losses may be different in the open and closed cases: the cost and benefits do not all arrive together. That could make it difficult to persuade participants at their bloc's potential sustainability.

There is a trade-off then; more openness generates fewer direct gains for producers in the member countries, but a larger share of the gains for the rest of the world (which may imply some feed back, increasing the benefits to member countries). That implies elements of cooperation. A more closed version, by construction, generates greater direct gains for the producers in member countries at the cost of less for nonmembers (which may in turn adversely affect the gains of the members). That is the coalition solution. Typically therefore import-substituting producers will favor a more closed bloc; export-oriented producers and consumers a more open one. These kind of conflicts show that special interest groups within a country or bloc will play an important complicating role and that we will need a simultaneous game framework to handle that.

From our perspective, the central questions are: which type of arrangement is most likely, which would end up benefiting its members most, and would the rest of the world be better or worse off? There are two schools of thought here. Some maintain the difficulties experienced in bringing the GATT negotiations to an end and the recent proliferation of regional trading blocs signal the breakdown of the cooperative trading regime.¹⁰ According to this interpretation, countries are now attempting to increase the (large, they hope) gains from trade by forming coalitions with their main trading partners at the cost of losing the small (they hope) benefits of freer trade between blocs of more distant trading partners. That might make GATT unravel to leave a number of free trading blocs with strong barriers between them. That would obviously be inferior to a system in which everyone has access to one world market with no barriers. But compared to the system we have today, we can still ask if members of any given bloc would actually gain more from reducing their internal barriers than they would lose from raising the external barriers to maintain the bloc? If so, they will be moving from a third-best to a second-best solution even though they cannot get to the first-best. But if not, they will be moving from a second-best to a third-best regime.

The other view is that the formation of trading blocs is a sensible way of breaking down the existing barriers and moving forward to a regime with significantly fewer barriers to trade (Lawrence,

¹⁰ The so-called "Memorial Drive" school as identified by Bhagwati (1993, p. 29).
1991). If countries find it convenient to group themselves into blocs in order to reduce any barriers between them, and if they can do that without feeling the need to raise the barriers between blocs, then that will be a net gain. And if, having done that, they can go on to accept any new members who are prepared to abide by the same free trade rules as the existing members — where the new members may be previously uncommitted countries or from other blocs — then they will necessarily reach a far greater degree of free trade than they started with. If GATT is slow moving, then it is at least possible that creating free trading blocs and spreading sideways from there is preferable to forcing things through successive GATT-negotiating rounds. Such negotiations are very complex and the benefits highly contingent because the sanctions on free-riders and on those who make concessions to special interests are not usually applied and may not even be enforceable. On this view, regional trading arrangements may be complementary to GATT (Lawrence, 1994).

II.3 How Does the New Trade Theory Fit into This?

At this point, it becomes important to examine why regional trading blocs form in the first place. Only if we do that, do we have a proper yardstick to judge whether groups of economies would rationally want to form a bloc; what they could expect to achieve by doing so, and the likelihood of them actually doing so. We can address these questions in two ways: using the new trade theory results "as is" — effectively a welfare approach, given the implicit preferences people use. Or we can look at the incentive structure of the different trade policy strategies (a game theory or political-economy approach).11

A Welfare Approach

The theme which distinguishes the new from the traditional trade theory is that markets are now recognized to be imperfectly competitive or to suffer certain domestic distortions, or to show significant scale economies in production (Richardson, 1992). If the problem is simply a domestic price distortion, then the correct remedy would be to correct that distortion and changes in the trade regime are unlikely to be an effective way to address this issue. So that doesn't help us explain regionalism. But if the issue is scale economies or imperfectly competitive markets then regionalism can make sense because it increases market size and reduces average costs in the first case and reduces market power of individual

11 Those two approaches correspond roughly to Paul Krugman's "narrow" and "broad" arguments for free trade. See Krugman (1993).
firms (and hence prices) in the second—provided that a bloc-centric industrial policy doesn’t go along with it, and provided that the processes of forming a coalition doesn’t “close” the bloc at its external boundary. In other words, "scale economies and oligopoly increase the potential gains from trade" (Krugman 1989, p. 361) and trading blocs provide an effective policy regime to explore these gains (e.g., through intra-industry specialization and the increase of competition).

Nevertheless, there are also some traditional arguments which might move us towards larger but more closed blocs along the lines of "optimal tariff" reasoning. In imperfectly competitive markets, market power allows "large" firms to keep prices above costs. This plays exactly the same role as an optimal tariff which, by limiting the volume of trade, allows "large" nations to raise national income. Indeed the optimal tariff is a function of the price elasticity of foreign import demand (Johnson, 1953). Hence monopoly pricing and an optimal tariff are equivalent if we leave aside their redistribution effects. On the other hand, gains made by one player will come at the expense of others in the market, which immediately invites retaliation. That in turn will leave the players worse off on average, even if one of them is individually better off.

From this perspective, it’s more attractive to seek a larger market in which you retain a degree of market power, and then maintain an external tariff or its monopoly pricing equivalent. That may invite retaliation from outside the bloc, but if the bloc is more closed than the original economy, this will matter less in terms of lost earnings elsewhere. That would imply a net gain, providing an indication of another type of incentive for forming blocs. Likewise, a bloc may be in a better position to pursue an aggressive industrial policy targeting "high-return" industries and influencing the world distribution of such industries in its favor. Needless to say, the capacity of any trading bloc to effectively pursue

---

12 Empirical work shows that growth is correlated negatively with market power, and positively with productivity growth. Both results suggest freer trade is better where there are imperfectly competitive markets. See, for example, Richardson (1992).

13 Note that this does not change the fact that unilateral liberalization would probably be first-best from a welfare perspective. The point made here is simply that regionalism may be welfare-improving as compared to the pre-integration status quo of trade barriers, and that scale economies and oligopolies tend to strengthen this result.

14 Reinforced by the fact that scope for intervention (by tariff or by exercising market power) falls as spillovers become more diffuse or more equal (since retaliation cancels what you do). So a trade policy of not cooperating becomes less destructive with larger, more diverse blocs.
welfare-enhancing "strategic" trade policies or optimal-tariff strategies has not been established in practice.

**A Game-Theoretic Approach**

The arguments above offer a rationale for forming RIAs, but they depend very much on the advantages which accrue to particular agents or sectors in the prospective blocs. That suggests that rent seeking and special interest groups will have a special role. But it does not provide a satisfactory explanation of across-the-board moves to regional trading arrangements as illustrated by the recent proliferation of regional initiatives.

Secondly, it is noticeable how often these explanations end up basing themselves on the benefits of cooperative behavior. Indeed, we argue that free trade (or rather the GATT as its outward and visible commitment, or punishment, mechanism) is an explicitly cooperative regime. In the case of large producers, cooperation is going to produce the best approximation to the Pareto optimality of free trade when producers are large in size but small in number. That would be a welfare-oriented approach too. But to get regionalism into that story we have to look at the political economy aspects in more detail, to find out why the intended cooperation might break down, and why players might get "diverted" from any tendency to regress back to a regime of ordinary unilateral noncooperative policies to form RIAs — i.e., coalitions — instead. That we do in section III.

Meanwhile we can start by observing that, when the number of producers/decision makers (n) is large — so that individually they have no incentive to respond to one another — and when the decision makers all have exactly the same objectives such as maximizing earnings, profits or utility, then noncooperative decision making will nevertheless produce socially (i.e., Pareto) optimal outcomes. In other words, under the usual assumptions of perfectly competitive product and factor markets, "small" agents, and full information, the individual and noncooperative-decision making of a free trade regime will be first-best optimal.¹

¹ The proof is based on an optimal control version of a dynamic game where n tends to infinity in the solution. See Hughes Hallett and Rees (1983).
But the crucial point is that, by using exactly the same framework, you find that as n gets small and the players become large (so they respond to each other's actions in different ways), or when their objectives differ\textsuperscript{16}, then only cooperation will maintain those Pareto optimal outcomes and the problem has a core of solutions which strictly dominates all others with noncooperation in them. Thus cooperative decision making actually reproduces all the desirable welfare features that free trade would have generated (strict perfect competition excepted).

There are therefore two options here. We must either create a trade policy regime which generates cooperation explicitly (i.e., a "perfect" GATT system). Or we must create one which simulates the mechanisms that produce the gains which we would have obtained from an explicitly cooperative framework. But what might those mechanisms be? There are several possibilities: a reallocation of decisions to exploit a wider "portfolio" of comparative advantage characteristics as far as possible; sequencing those decisions best (i.e., specializing to exploit comparative advantage over time); reducing/internalizing externalities by "trading off" policy interventions between players; and improving domestic policy effectiveness by reducing spillovers (i.e., creating scale economies in interventions).\textsuperscript{17} One way to implement this — at least in theory — would be to have RIAs constructed as a sequence of expanding mini-GATTs leading to a worldwide "perfect" GATT.

These are exactly the desirable characteristics of the policy regime which Paul Krugman associates with the new trade theory, with all its attention to market structure and oligopolistic behavior (Krugman 1987, 1993). He argues that strict free trade (as conventionally defined) is no longer first-best optimal. However, something very like it will be — specifically a regime with these cooperative characteristics which rule out the predictable excesses of noncooperative behavior. Hence holding to a conventional free trade regime won't involve much loss.

The crucial distinction here is that cooperation is needed to secure free market access for the existing producers, whether they are perfectly competitive or not. Free trade as conventionally

\textsuperscript{16} Their objectives may differ because they face different structures in their input or product markets for example; or because both producers and consumers are involved in the game.

\textsuperscript{17} The proposition that cooperation improves policy effectiveness, as well as its other potential benefits, are laid out in Cooper (1969) and Hughes Hallett (1986). Policy effectiveness can be equated with scale economies to the extent that it implies that the incremental gains in the achievement of policy targets per unit change in the instruments are larger under cooperation.
understood will require something more since the existing market structures (which may entail market power for some producers) will not always support perfect competition. It is not clear that, by imposing the rules of strict free trade when "n" is in fact small, we will actually impose only small losses on the players individually (as opposed to on average). Strictly speaking, this is the wrong approach anyway since you can always compute the Pareto optimal cooperative solution explicitly in any given case. However that objection may not matter much in practice since if markets are not fully competitive you cannot impose those free trade rules without either extensive regulation (which you can't do because there’s no international law to enforce that regulation), or without extensive and carefully coordinated interventions to rig prices/supplies to generate free trade outcomes out of imperfectly competitive markets (which you can't do because it's not incentive compatible since, by construction, free trade is less than Pareto optimal in this case). So what you actually have to do is to reinvent a GATT-type free trade structure by which governments pledge to use their own national regulation systems to produce the first outcome. Or you have to invent a series of explicit negotiating frameworks capable of forcing the cooperative trade policies you need (to give the second outcome). That is usually most easily done with your nearest trading partners, leading to the formation of regional trading blocs.

This reformulation of Krugman's broad, or political economy arguments for "free trade," doesn't produce free trade as such, but a cooperative multilateral trading regime. That is certainly consistent with, and could possibly be achieved through, a system of mini-GATT arrangements for groups of economies. It is also a result which does not depend on an appeal to the informally defined second-best approximations used in Krugman's arguments. Nevertheless, we can see that one crucial element is missing. There is nothing in this view of the world which ensures that the distribution of the benefits will be either incentive compatible for all, or acceptable in the sense of reflecting their (perceived) market power. So the outcomes could well turn out to be contestable, either to the participating countries or to the private sector interests within them. Why, for instance, would the players want to commit to this open (and expanding) regime of mini-GATTs if neither free trade rules nor explicit cooperation favor them as much as being able to use their market power, either individually or as a member of a coalition? The point here is to draw a distinction between a multilateral cooperative regime which guarantees free market access to a small or large number of producers, and a free-trade/perfect-competition regime where there are only large numbers of producers so that prices settle at marginal costs.
III. GATT AND RIAs

At this stage, it is worth asking why countries have found it so difficult to go right through to global cooperative solutions? Is it because they find it hard to commit themselves to regimes where there are few effective sanctions against those who try to free ride? It may well be that regional blocs work better as pre-commitment devices because they are seen to lock in favorable trading practices between close partners to a degree not possible in a global scheme; because the losses for a player disciplined or ostracized from his bloc are more immediate and tangible; and because a smaller bloc can more easily create a network of interlocking economic benefits, beyond just trading arrangements, such that reneging on the latter can trigger sanctions which deny the “sinner” the other benefits as well.

We start our analysis by focussing on the problems of the GATT system as a commitment device. After that we look at inter-bloc trade relations (and related institutions) and to what extent regional blocs can be used as “building” blocs towards a new global cooperative regime.

III.1 The GATT System

The GATT system (which in the near future will become part of the World Trade Organization) amounts to a commitment to a cooperative regime in which there is limited advantage in unilateral liberalization — but the advantages of multilateral liberalization accrue to all only so long as all “play the game.” The threat of discrimination against those who broke ranks, in particular more difficult conditions of access to the large US market, sustained this process of liberalization for many years. But, like any cooperative regime, there is in fact often little sanction against individuals who revert to their best noncooperative policies — and quite possibly none at all against those who form a coalition with market power.

First, participants may judge retaliation to be uncertain, unreliable, and costly for the injured party(ies). Second, countries may be reluctant to incur the costs of retaliation against dissidents whose unilateral actions do not affect them much. Similarly, those who would be left at a disadvantage by unilateral action elsewhere will be reluctant to cooperate with those who, in their own interest, fail (or have failed) to play the game. That in itself will start to sort countries out into coalitions of like-minded partners, and once there are two or more larger “players” in operation the pressure for freer trade will fade, since the competitive offering of access to the coalitions’ markets will secure much of the
cooperative benefits for most of the participants, while reciprocal discriminatory trade policies will ensure that free trade in a wider sense is always denied to some group(s). Indeed it is quite possible that some coalitions will form which can secure greater benefits for its members than would have been possible under full cooperation, but at the cost of worse outcomes for those outside. Those outsiders may then form a coalition in self-defence.

These last remarks reveal an important reason why the GATT system will continue to face challenges in the post-Uruguay Round world. Once countries have made the calculation that their interests are better served by making free trade arrangements with their immediate trade partners and by allowing more noncooperative policies guide their relations with non-members, they would only not move towards a system of regional blocs if the sanctions that could be levied against them within the GATT system itself, or in retaliation by other (excluded) governments, were sufficiently powerful.

History shows that GATT-sanctioned retaliations are the exception rather than the rule. Actually, the only example of GATT authorized retaliation was issued in 1952, as the result of a complaint by the Netherlands against GATT-illegal barriers imposed by the United States upon dairy products. More recent requests for retaliatory authority (by the EC and Canada against the United States in 1988-89) have been blocked by the defendant country. This should not come as a surprise given the fact that GATT's dispute-settlement procedures (as everything else in GATT) operate under a consensus rule (Hudec, 1990, p. 184).

As argued by Finger (1988, p. 13), the "GATT-founding fathers" most likely had as their objective to minimize the probability of GATT-authorized trade sanctions. In this context, GATT's dispute-settlement mechanism was framed to foster consultations among the involved parties and to work mainly through moral suasion (peer pressure). The threat of trade retaliation was, of course, there, but as history has shown its implementation is, to say the least, unlikely. If one adds to this the failure of Article XXIV to effectively discipline the formation of blocs, the potential for non-cooperative behavior among trade blocs becomes evident.

---

18 It is worth noting that neither did the Netherlands act on this authorization to retaliate, nor did the United States change its trade practice. See Jackson (1989, p. 96).
Moreover, GATT's effectiveness is further complicated by the fact that many trade restrictions which are now commonly used do not properly fall under the disciplines of the GATT treaty. This allows governments to maintain a reasonably GATT-consistent face to the world while operating policy instruments which have a protectionist impact in practice. Administered protection (anti-dumping and countervailing duties) or "voluntary" export restraints, for example, can be targeted rather precisely to thwart foreign competitors. But their redistributive effects (via rent transfers) are non-transparent both inside and outside the country imposing the trade restriction.

Against this background, it seems clear that countries with their own interests in bloc-wise noncooperative behavior could easily find ways of doing so without apparently violating GATT disciplines (e.g., through anti-dumping actions that discriminate against firms of non-member countries). Once countries realize that is the case, the most powerful reason for wanting to form free trade blocks within the GATT system becomes apparent. It is a defensive move, since there is no other way that countries can credibly commit to behave in a properly cooperative manner within GATT.\(^9\) It is therefore better to withdraw into a smaller coalition of your immediate partners where an explicit agreement (and the implicit threat of expulsion and loss of market access in general, rather of access to particular markets) can "lock in" liberal trading policies and market access where it matters. Unfortunately, the very need to make such an arrangement sustainable, and the commitments or sanctions credible, suggests that the bloc might well turn into a "closed" trading arrangement.

The successful conclusion of the Uruguay Round in December 1993 indicates that the multilateral system is trying to respond to these challenges. The Uruguay Round extended and deepened the network of variables covered by multilateral rules. This can be interpreted as an attempt to generate sufficient incentives and potential sanctions to sustain the multilateral trading regime. In any case, one might argue as we do that even with strong regional blocks, the GATT (or the future World Trade Organization) will still be needed as a means of preventing backsliding, of handling trade disputes (particularly, among blocs), and of extending the scope of the multilateral rules to ensure that incentives for multilateral cooperation do not disappear with the growing importance of RIAs. An important area of research would be to see how that might be done. But all of this depends on GATT maintaining effective and credible dispute/sanctions procedures.

\(^9\) A similar point is made in Hindley and Messerlin (1993).
The Dispute Settlement Understanding of the Uruguay Round will improve the ability of the GATT to adopt panel reports against countries that infringe multilateral disciplines. In the past, the consensus rule often impeded the implementation of the findings of a dispute settlement panel (since the "guilty" contracting party could block the adoption of the panel's results). Under the new rules, unless all contracting parties agree not to adopt the results of the panel, they will be adopted within 60 days of the release of the report. This is in principle an important improvement in terms of GATT's institutional capacity to discipline "GATT-illegal" behavior. Yet, given the problems of punishing non-cooperative behavior at a multilateral level in a timely fashion, this institutional improvement will not erase the appeal of RIAs as pre-commitment devices, as argued below.

III.2 On Credible Commitments to a Regime and Institutional Reform

Is there anything intrinsic to RIAs that make them work better as pre-commitment devices than a global regime such as the GATT system? Is there any reason to suppose that the participants would want to commit in some absolute sense even to a regional cooperative arrangement? There are two arguments here: first with respect to the "absolute" incentive to commit to any cooperative regime (global or regional), and second to show the relative ease of making credible commitment within a system of (smaller) regional blocs. Those are important arguments because they highlight a fundamental issue: whether a regime of regional blocs would be sustainable in the long term.

The usual justification for expecting commitment to any regime is that if you cheat and get away with it the gains are only small (convex "losses" towards the private optimum). But it is open to anyone who fears that such a lack of commitment to the new regime would damage them significantly to cheat back, and that would impose larger losses on you since it involves convex losses away from your private optimum. Since in a continuing game those threatened losses would be larger than the potential gains, it is clear that you would be better off by committing to the rules of the regime and not cheating. That's one version of why cooperative trade policies would be sustainable: punishment losses will outweigh the temptation gains. Indeed, triggering a trade war would push both players off their contract curve and

---

20 A property of Pareto optimality when the other players gain. What is more, cheating back can be done at anytime (preemptively, simultaneously, or retrospectively) and by anyone (individuals or coalitions). Bond and Syropoulos (1992, 1993) have different results, but they are dealing with cheating between blocs, in the cases where such blocs already exist. Here we are talking of commitment within blocs, where market power is not an issue.
(if they are of similar size) leave them both worse off and therefore persuade them to commit to the cooperative/free trade solution. But this result will be fragile if a bloc has significant market power (see section III.3); you may have to make blocs of a similar size, and maintain something of a closed structure to maintain credible punishment options should someone cheat, in order to sustain it. Altogether that makes free trade hard to maintain — particularly if one of the blocs is large enough and the transgression principally affects the more remote trading partners, or where retaliation has uncertain impacts. The difficulty which the authors of the GATT system had in getting agreement on the principle that participants should be able and willing to take collective and effective action against signatories who violate the agreement’s provisions, illustrates the significance of this qualification.

The sustainability of cooperative trading arrangements may therefore depend on whether countries find it easier to commit to regional blocs than to a global regime. A system of regional blocs, in which each player maintains several different targets simultaneously, not only focusses the commitment problem on the costs for (and retaliation from) one’s immediate partners. It also allows us to replace the individual targets of more remote partners in one arena with a wider range of targets for the immediate partners in various different arenas. It is the latter point which matters here. The need for cooperation is more apparent because to get the additional gains in efficiency we have to be able to negotiate a trade off between an improvement in the target of one player in one arena against that in the target of another player in another arena. This we do by playing out the implied simultaneous game. That is the attraction of deepening instead of widening. Moreover the costs of cheating and retaliation will also be clearer since the former will show up in all of a player’s targets, and retaliation can be undertaken by any of a player’s instrument variables. In addition, since we are now dealing with a self-supporting coalition of players, punishments will be exacted by a larger numbers of players (who want to pursue all the interests of the bloc) rather than just the parties injured by one target failure.

There are many examples where the formation of regional blocs has naturally led to a network of linked targets between members, in place of one set of similar targets across nonmembers. Eichengreen and Frieden (1993) refer to the European Monetary Union (EMU) deal as bargaining between members which involves implicit or explicit links between EMU and other issues. Garrett (1993) highlights the case of Germany because the Germans find EMU relatively unattractive, but go along with it in order to secure other goals such as the acceptance of German reunification within the EC, or achieving increased exports to the rest of the EC. What of course is happening is that coercion, where a country’s interests might be marginal or negative, is being substituted by gains in that country’s other
targets to maintain the necessary incentives.

Forming a regional bloc implies new or revised institutions and it is the ability to design new institutions which take a whole range of issues under the competence of one regulatory body which make countries more prepared to commit themselves to regional blocs. That in fact is a special case of a more general proposition (Martin, 1993; Cohen, 1993). Institutions, in this case the newly designed regional ones, will play a powerful role by changing the incentives that countries face. Consequently making a denser network which binds policy linkages together, and by shrinking cooperation onto a larger number of variables affecting a smaller number of countries, makes the bloc both more attractive because where there are potential losses there are now more issues at stake with which to make side-payments (more "quids" per "quo"), and more robust (because there exists a greater need to exact punishment to redress the losses per transgression). That is why countries might be prepared to commit to regional (but not global) cooperation. Sustainability therefore is attained whenever blocs can provide the institutional backing necessary to deter independent governments from breaking bargains that turn out to be inconvenient in some respect.

III.3 Is the Proliferation of RIAs Consistent with a Move to Global Free Trade?

Recent analytical work has focused on modelling the interactions between blocs in more detail; specifically on how the prospects for free trade between blocs fare, and how the welfare of the blocs themselves varies, as bloc sizes increase or become more asymmetric. These results are important for our arguments because they illustrate some of the incentives for forming, then expanding or deepening the cooperation within a bloc; also for what is likely to happen in terms of free trade, cooperation or continued barriers between blocs. That is important because by tracking the welfare gains of the different players under the different regimes or bloc patterns we can say something about where the present trend to regionalism is likely to end up; and to evaluate the argument that progressing through a sequence of mini-GATTs of increasing size is an easier or more effective way of reaching a full free trade regime than attempting to broaden out the present GATT system with its lack of credible enforcement or commitment mechanisms.

Lawrence (1994) makes the same point. Of course this advantage is likely to come at the expense of outsiders — a denser network of targets created to underwrite the interests of members typically marginalize those of nonmembers.
A convenient starting point for an analysis of this problem is the contribution offered by Kemp and Wan (1976). Kemp and Wan demonstrated that, in a completely unrestricted world (i.e., competitive trade between any number of countries in any number of goods and where tariffs, export taxes, etc are permitted), a subset of those countries can always form a customs union with a common external tariff structure and a system of internal side-payments such that everyone (whether a member of that union or not) is no worse off — and in general better off — than before the formation of that union. This result is important because it illustrates that a move to regionalism (in the guise of customs unions) can be seen as a move towards global free trade (Lawrence, 1994). In other words, progressing through a sequence of mini-GATT blocs (new RIAs being formed or, more likely, old ones being enlarged at each stage) is indeed an alternative and perhaps easier way of creating a system-wide cooperative free trade regime, as intended by the authors of GATT.

But notice two crucial qualifications. First a system of side-payments within the union is typically needed to get the result that everyone is made better off. That can only be fixed by explicit cooperation within each union. Hence the progression to a full cooperative regime through a sequence of mini-GATTs can only be achieved in general with explicit intra-bloc cooperation. If for any reason the incentives for that cooperation were to fail, including when creating the final worldwide bloc, then this strategy of creating free trade outcomes via a sequence of free trade blocs would become infeasible.

Second, this desirable equilibrium is not going to be unique. In particular there is no assumption that each player will have maximized their own interests in making each step through to the full free trade regime. It is also perfectly possible, indeed highly likely, that some of them could combine into a union which, by exerting market power at the expense of weaker or unorganized nonmembers, could produce greater gains for its members than were available either before that union or under worldwide free trade. And if that is the case, then, there will come a stage in the progression through mini-GATTs to full free trade, where no defensive coalition can prevent the larger customs union to appropriate those gains for itself and the movement towards a free trade regime will simply stop. Unless, of course, there are sufficient non-economic losses, or sanctions imposed by some global institution or hegemon interested in world as opposed to private welfare, which offset the incentives to form such a blocking customs union.

---

22 This situation is highly likely because otherwise it would not be necessary to offer side-payments each time new entrants (at least at the final stage) joined the existing union.
That outcome seems likely because the Kemp and Wan proof depends on nonmembers' trade being fixed at its pre-union level at each step, so that nonmembers suffer no trade diversion as a result of the union. If trade barriers within the union then fall, there will be trade creation in the union but no diversion outside. A Pareto improvement versus the status quo ante. In fact, greater trade within the union means high incomes and hence some spillover of extra demand for nonmembers at the preexisting interbloc trade levels. All players actually gain therefore, provided side-payments within the union ensure a net gain for all members as the internal barriers come down.

But this is only one possible construction of the gains from forming or extending a union, and as such it is not necessarily the self-maximizing construction. Greater gains (for the union as a whole) will typically emerge when the constraint that nonmembers trade must remain fixed at its previous level is removed, and the union is free to maximize its collective revenue. Exploiting market power like that tends to raise union prices and tariffs (offsetting any tendency for prices to fall internally) at the expense of nonmembers who now lose out because of interbloc trade diversion without any compensating side payments. But side-payments to new members will also typically be necessary to ensure that they share in some of the gains relative to what they could have achieved outside the union. At the last stage of enlarging blocs to a single free-trade bloc, a union (with market power at the penultimate stage) would lose those remaining interbloc tariffs which, sustained by that market power, were generating the union’s excess gains over cooperation/free trade. Self-interest would therefore prevent such a union going through to the last stage.

Krugman’s (1991a) "Is bilateralism bad?" paper provides an alternative insight on the prospects for global free trade when blocs form and/or become bigger in size and maximize their own interests. Krugman assumes full cooperation (free trade) within each bloc, but no cooperation (optimal tariffs) between blocs, and allows the size of each bloc to dictate market power in raising those (external) tariffs in a single-shot game. In this game all countries are identical but each produces a different good, so that there are no comparative advantages. Krugman then shows that individual welfare increases (slowly) with bloc size. But small inter-bloc tariffs generate significant trade diversion so that world welfare decreases more sharply with bloc size until there are just three blocs, and then increases rapidly to its first best level at free trade (a single bloc). That in itself is a warning of what may happen, but the underlying model is a highly specialized one and the results may not be so robust.
Bond and Syropoulos (1992, 1993) have, in a series of papers, given a different picture of what happens as the blocs grow in size. They first allow each country to have comparative advantage in producing one good and the same advantage in producing all the others. They also allow the interbloc tariffs to be set from a repeated game; if the implicitly cooperative behavior thus generated is violated, the punishment strategies triggered are a single period noncooperative tariff game before cooperation can be reestablished.

Bloc size can then be varied in two ways. First equal sized blocs are formed and their number reduced. That means within bloc trade rises but the relative bloc sizes remain constant. It also means the share of "home" consumption in total output rises, so that each bloc's (absolute) market power rises as well as the optimal tariff if noncooperative interbloc behavior is triggered. In this case, increasing bloc size has two effects: (i) it increases the welfare gain which one bloc can make by defecting from free trade to maximize its private interests (since larger blocs have greater market power vs. the rest of the world); but (ii) larger blocs means that the welfare level reached in the noncooperative punishment phase which follows is lower (because noncooperation with larger blocs means their own policies become less effective - which is Cooper's (1969) standard result). Thus the incentives to cheat go up, but so does the punishment for cheating, which leaves it ambiguous whether the incentive to go noncooperative rises or falls with bloc size. Bond and Syropoulos, however, show by simulation that the cheating gains always dominate in their model, so free trade becomes harder to sustain with larger blocs.

Second one can fix the bloc size for a given numbers of blocs, and then increase the size of the first one to see the effects of increasing relative (rather than absolute) bloc size. Asymmetry presents the larger bloc with increased incentives (greater market power) to cheat relative even to the increasing costs from the subsequent retaliation and punishment. In other words, inter-bloc cooperation becomes even more difficult with asymmetric blocs.

These two results are, like Krugman's, derived from a particular model of production, consumption and trading patterns between blocs. So the question of the robustness of the results to alternative specifications naturally arises. However, Bond and Syropoulos are able to demonstrate that they hold firm with increasing degrees of comparative advantage and increasing degrees of risk aversion (i.e., curvature in the country/bloc welfare functions). In fact, as one might expect, increasing risk aversion makes it harder to support free trade as blocs fear cheating more and hence raise their evaluation of the gains to be made from cheating on free trade by more than they would raise their
evaluation of the punishment costs. Hence the former dominates the latter by even more and it becomes harder to support free trade since everyone perceives larger downside risks/costs from defecting. Raising the degree of comparative advantage, however, has the opposite effect; it makes free trade easier to support, but not by enough to overturn any of the existing results. Comparative advantage evidently increases the private gains which a bloc can capture by increasing its size and behaving noncooperatively. It also has the effect of decreasing the welfare attained under the noncooperative (punishment) phase, so the costs of defecting from free trade are greater because less effective policies — or more offset between them — mean that tariffs have to be used harder to try to capture those large private gains. Of these two, the effect on punishment costs is larger, so that free trade becomes relatively easier to support with stronger comparative advantages — but not by enough to remove the incentive for blocs to follow competitive trade policies.

These results suggest that once trading blocs become of significant size it is increasingly unlikely that the world could maintain a cooperative (free trade) regime between them. That implies losses compared to a full free trade regime, although it is an open question as to whether that kind of loss is large or not. It can be argued that, compared to the gains from free trade within each bloc, those losses would not be large — but we know of no serious empirical (or theoretical) evidence to back that claim up in practice. Yet, as mentioned in Krugman (1991b), even if global welfare losses are not serious, individual countries left behind could be significantly affected.

The process of forming and enlarging blocs would not lead to a "single bloc" free trade regime unless the presence of those losses, and the fact that it is a repeated game, persuades the remaining blocs of the advantages of, and each players eventual commitment to, interbloc cooperation. But for that even to be possible (ignoring for the moment that, given the earlier results, it would probably not be possible to persuade participants that these large blocs would ever commit to interbloc cooperation), the gains from interbloc cooperation have to remain positive as blocs increase in size but each bloc continues to maximize its own "private" welfare.

Bond and Syropoulos (1993) in fact show that as one bloc gets larger, not only do its gains from defecting from free trade yield higher welfare, but its welfare, given noncooperative behavior with respect to the rest of the world (having defected), is also higher. At that point one of the blocs has become a dominating coalition which cannot be blocked by the rest of the world; with market power it can do better as a coalition than in a free trade regime, but at the cost of the rest of the world. So only do large
blocs have an incentive to defect from free trade in the sense that they gain more from cheating than they lose from noncooperation subsequently; they may even gain from noncooperation itself if they are large enough.

In subsequent work, Bond and others (1993) showed that this result continues to hold even if several smaller blocs attempt to form a blocking coalition against a sufficiently large dominating coalition. Hence free trade achieved through a sequence of mini-GATTs is not a feasible strategy, even if a credible commitment mechanism were created (and guaranteed) through some multilateral agency able to impose sufficient penalties on those who try to violate the agreed free trade or market access rules.\(^2\)

The only ways out of that difficulty are if: (a) the dominant bloc can be persuaded through altruism, or because other (noneconomic) interests would otherwise be forfeit, to play cooperatively through to the end despite the incentive to do otherwise (for example, it can be argued that the United States did this in the post-war period because it had ideological and security concerns which led it to preserve free trade policies in the West through GATT, rather than push its own economic interests as it could have done); or (b) if the blocs themselves link up a network of other issues to free trade, such that if the latter were to fail the other issues (on some of which the economically dominant coalition is not dominant) would be dealt with noncooperatively to the cost of the dominant bloc; or (c) if the smaller blocs, realizing the potential cost of being dominated in this sense, make it their policy to maintain a blocking coalition of sufficient size before the larger blocs get to the stage of excluding them (but it is not clear why they would not try to join the larger bloc’s bandwagon first.); or (d) the smaller blocs lobby for a multilateral agency with blocking powers; or (e) the smaller countries form a coalition in a specific sector where they have market power (e.g. in certain commodity markets). In any of these cases, the larger bloc would cooperate; otherwise not.

IV. PROTECTIONISM AND THE POLITICAL ECONOMY OF REGIONALISM

To understand the process of regional integration we have contrasted multilateralism (cooperation) with protectionism (noncooperation) and regionalism (coalitions). In order to uncover the mechanisms

\(^2\) All the results in this section are derived using tariffs as the instrument of trade policy. But to the extent that other instruments (quotas, export restraints, nontariffs barriers) can be written in terms of their tariff equivalents by calculating their impacts in raising prices at the border, these results can be generalized.
which make regional coalitions relatively attractive, we argued that there is a natural tendency to "close" the coalition at its boundaries (or to deepen it) in order to maintain incentive compatibility and maximize the local gains, and to maintain the credibility of the commitment to free trade. That has to do with the coalition's market power. But as we noted at the start, a second possibility is that political pressure from special interest groups will produce a similar outcome by distorting trade policy to suit their particular industries. The importance of special interest in this connection was already apparent in the incentives to "close" the bloc at its external borders, in the trade-off between the gains in some industries and losses in others, and in the need to provide insurance against the bloc breaking down. This section therefore considers how special interests would influence the formation and the policies of regional arrangements.

We have to recognize that, while governments try to improve the welfare of their electorates, they are not immune to political pressure and offers of support from special interest groups. Moreover those interest groups will interact strategically with their governments — and between themselves because their interests will necessarily clash internationally and domestically since greater assistance to any particular special interest would (ceteris paribus) cut into general welfare and cause governments to limit the total amount of assistance on offer. Other special interest groups would then suffer.

On the other hand, a regional arrangement carries the potential for diluting the interest groups' impact and hence for improving welfare; the larger the association, the less sensitive is policy to private interests which do not coincide with general welfare. But whether that actually happens or not also depends on how the interest groups interact (i.e., whether they cooperate too); on how the governments interact (competition may dilute their policies, and make them more dependent on political support); on what the external policies are; on the differences in preferences and policy effectiveness between governments and interest groups; and on the weight which governments give to satisfying special interests.

To determine the final outcome, we have to set the problem within a framework which incorporates strategic interactions between special interests and governments, as well as between the governments themselves; that is a series of simultaneous games. Unfortunately, the literature has little to say about that. However, recent work by De Melo and others (1993) provides a simple model which focuses on the interactions between a single lobby for private sector interests and a single government.
Two scenarios are considered: the case where the single government is a national government acting in isolation, and the case where that single government is a regional institution constructed from an "equal shares" bargain between two national governments. There are therefore no strategic interactions, or indeed even spillovers, between the two countries or between different interest groups; all the action is between the private sector lobby and the government in question, with a view to showing the difference in policy settings when the latter is a regional institution from when it is a national government.

The burden of the results is to show that the dilution effect is always present, but will be offset by asymmetries in preferences between the partners in the regional institution. So it is not clear, even in this restricted world, if regional integration would make the participating economies better off, or that the governments themselves would feel better off and therefore have an incentive to integrate. On its own the dilution effect would reduce the distorting influence of the lobbies, implying an improvement in performance (and efficiency) in terms of the underlying targets of policy. On the other hand, if the aims of policy (which must be interpreted here as depending on the relative priorities put on the underlying targets, and on the relative ease of achieving those targets) differ sufficiently between countries, then those improvements in performance will be offset to the point, eventually, where one or both countries will be worse off than without integration.

That is the basic insight, but there are other points. First any "institutional engineering" which would control the parameters of these lobbying policy-making interactions (i.e., the relative priorities on targets and political support; and the fact that the governments play follower without attempting to influence their lobbies further, and appear to ignore any policy conflicts between themselves or between lobbies) is more easily done when the regional institutions are being set up, rather than later. The question of institutional design (or reform) is therefore crucial. Second, it is easy to show that the preference asymmetry effect becomes stronger the more members the regional association contains, making it progressively more difficult for each participant to obtain a net improvement in performance on the strength of the dilution effect. That explains why the new regional groupings typically have a smaller number of like-minded members rather than the more diverse membership of the global postwar institutions. It tells us, in particular, that a series of mini-GATTs may appear to be progressing while GATT itself appears to move only slowly. Finally, the negative and possibly dominant preference asymmetry effect naturally implies that getting convergence will become an objective in its own right for
the regional institution, but not for the member governments. The debates over subsidiarity (or compliance) reflect that tension.

Nevertheless, one must recognize that the decision model used to derive these results is extremely restricted. By definition none of the decision makers are "small" within their regional blocs. Hence governments (or pressure groups) will recognize that their actions have significant spillovers on their counterparts in other countries and start to act strategically between themselves as well as between pressure groups and between governments. Similarly, when governments realize that, by intervening, they alter the limits of what is feasible and desirable, they will want to influence what the pressure groups seek to obtain instead of just following the lobbyists' lead. Then there is the question of what happens if interventions may be different in different parts of the integrated bloc, or if different governments attach different weights to the importance of the lobbyists political support, or if the lobbyists do not benefit equally from a given policy change in their own economies? On these matters we have no guidance, but we can see that solving the simultaneous games set out in the Appendix provides a way of analyzing them in a framework which allows any kind of asymmetries in preferences or structures, and any kind of strategic behavior between governments or between governments and pressure groups.

Using an alternative approach, Grossman and Helpman (1993a, b) focus on the effects of strategic behavior between governments, rather than on that between pressure groups and governments. The pressure groups seek to influence government policy in their own favor with political contributions. Government objectives meanwhile consist of the economic welfare of their electorate, and the support (contributions) offered by pressure groups. One can then show how optimal tariffs vary with the size of the contributions and with the importance (to the government) of those contributions when there is retaliation (a tariff war); and then the same again when there is negotiation in the form of trade talks, both cases being compared to the case where governments are immune to political pressure.

A number of interesting results emerge, but again the limits on the interactions between players suggest that this kind of analysis needs to be placed in the context of simultaneous/overlapping games for one to grasp the real scope of the results. For example, the Grossman-Helpman model has the optimal tariffs depending on the optimal political contributions from each lobby, where the latter are determined taking that dependence into account (i.e., a Stackelberg equilibrium with the government as follower).
Having solved that part of the problem, domestic tariffs become a function of foreign tariffs and vice versa. So to solve the tariff game requires a Nash equilibrium between governments. However, the contribution decisions made by each lobby in the first stage are quite independent of the spillover effects they may have on any other lobby, domestic or foreign. That means there are no interactions between lobbies, except indirectly, since the contributions of any lobby will be a function of the expected tariffs which are themselves determined in response to other political contributions. Hence there is a Nash element to the final solution of the contribution levels, but only through the agency of a third player: government. That suggests an overlapping game formulation. That would certainly be appropriate for the cooperative "trade talks" scenario. For the noncooperative "trade war" case, it is a free-for-all; there is no linkage through a single third party because governments no longer cooperate. A simultaneous game formulation would be the appropriate framework for analyzing that kind of problem, therefore.

To give an idea of the results, Grossman and Helpman show that, in the trade war case, import tariffs and export subsidies will be higher (export taxes lower) when interest groups exert political pressure on, or make contributions to governments. Moreover, that outcome is exaggerated if the political contributions become more important and the immediate concern for general welfare is reduced. In that case the import tariff for industry A (a net importer domestically) will rise, while export subsidies to industry A abroad also rise; protectionism starts to rise between blocs.

Meanwhile, other sectoral interests in each economy will also interact. What happens to industry B? If it too is an importer in the home country, it will expect the same solution. But it will also see that the additional pressure on the government for assistance would cut further into general welfare and cause the government to hold back on granting additional tariffs or subsidies, given what industry A is already obtaining, in order to preserve the optimal mix of welfare and aggregate protection. Thus, industry A’s contributions have a perceptible spillover onto industry B, and (in the absence of any adjustments by industry A) one would expect industry B to take account of that either by reducing the contributions it would have made in isolation so that the average tariff/subsidy provision (and hence industry B’s share) is maintained at a higher level (cooperation between lobbies); or by increasing its contributions in order that a larger share of the then declining average tariff/subsidy provision comes its way (no cooperation between lobbies). Meanwhile industry A will be making exactly parallel calculations, so that they will in fact reach an optimal noncooperative equilibrium. Industry B’s possible actions are therefore standard
noncooperative responses, and the strategy actually selected will depend on whether reducing average tariffs would have a bigger (positive) impact on general welfare, or whether maintaining them would have a larger (positive) impact on governments’ evaluations of their political support.

If however industry B is an exporter at home and an importer abroad, the story (in isolation from industry A) would be reversed. That might change the results when industries recognize the spillovers between them and the potential for retaliation. For instance, industry B’s will realize that industry A’s import tariff will trigger a foreign tariff in retaliation. That will tend to increase industry B’s pressure for a yet larger export subsidy. Again that could be done by reducing contributions to maintain a higher average subsidy at home (since subsidies will be subject to a budget constraint in aggregate), or by increasing them to attract a higher share of the available subsidies, depending on the contribution multipliers and their priority in government objective functions. But industry A will be aware of that and will wish to moderate its own contributions for the same reasons. The net outcome of these changes is not easy to predict. As in the previous case, we would have to solve the simultaneous game in contributions and in tariffs/subsidies to find out.

It is particularly important to consider sectoral interests, and the potential conflicts between them, since none of these stories explicitly consider the interests of consumers and those hurt by protectionism. In general some interests will exert political pressure to get protection while others try to get it lifted. That is the point of industry B playing the opposite role to industry A. Nevertheless in a general noncooperative set up, players will neglect the response of their opponents, and any of these solutions will result in an oversupply of contributions and hence of tariffs or subsidies. Trade negotiations (cooperation between the government players, and hence the simultaneous game version) would lead to a lower level of protection. And that is what happens with potential welfare benefits being higher, the stronger the political imbalance in the trade war case. If political power is roughly equal between groups and countries, the negotiated solution reverts towards free trade.

Hence, on the face of it, an internally cooperative regional bloc is likely to adopt a less protectionist stance than the members might have done in the trade war case, unless the special interest groups recognize their own interdependence or learn to cooperate, or unless the trade negotiations involve problems of incentive compatibility which can be fixed by side-payments generated through external
tariffs, or unless the coalition indulges in noncooperative behavior with respect to the rest of the world in order to increase its "market" share (as discussed in section III.3). In all those cases, a regional bloc would tend to be more protectionist than its members would have been alone.24

In addition to those results, Grossman and Helpman (1994) also show that free trade agreements are most likely to emerge where there is relative balance in internal trade; and where the protectionism of a high price partner can be captured by the exporting industries of other countries, rather than when the low prices of a liberalized member would be transferred to all competing industries within the bloc. Whether the latter would actually happen depends on the market power of the industries in the low price/low tariff economies, or equivalently on whether they can supply all or only part of the high price/high tariff country’s demand, and on the bloc’s policy towards external tariffs — all of which will be a matter of particular cases, with the general implication that trading blocs will tend to occur where they are least helpful in terms of maintaining/raising aggregate economic welfare. That of course reinforces our earlier conclusions that regional blocs may well prove to be more protectionist than is members would have been alone, and that the proliferation of regional blocs increases the need for GATT-type disciplines. Thus this "political economy" approach tends to strengthen the conclusions already reached in the context of the "strategic behavior" analysis of section III.

V. CONCLUDING REMARKS

The picture that emerges from our analysis is not an optimistic one with respect to the future role of RIA’s in advancing a liberal trade order. We agree with the proposition that regionalism is here to stay (actually, it has never left the stage in the post-World War II period). What is not as clear is the future role of these arrangements in paving the way for multilateral cooperation. It is true that RIA’s did not choke the process of international economic integration in the past.25 The future, however, may not be

24 Given our observations at the end of section III. 3, it may be interesting to extend this Grossman - Helpman type analysis to 3-country, 3-sector interactions where one of the players is a sector based (rather than country based) actor, such as the Cairns group (the coalition of countries in favor of agricultural trade liberalization in the Uruguay Round). One could then analyze how the coalition’s behavior would change as the trading rules are extended.

25 For details see Anderson and Norheim (1993).
a replay of the past given the emergence of new trading blocs with market power, the evident need for a credible commitment mechanism, and lingering doubts about the capacity of the GATT system in disciplining "cheaters" in an effective way. In short, it seems unlikely that RIAs could work as building blocs to a "perfect" GATT.

The propensity of larger RIAs to follow noncooperative behavior will depend on the relative importance of extra-bloc trade. If the benefits from extra-bloc trade remain significant, export-oriented interests will serve as a counterweight to interventionist forces. Actually, there is evidence that regionalism has not reversed the process of international economic integration. In other words, although intra-regional trade has been growing in the post-World War II period this has not choked integration between regions (Anderson and Norheim, 1993). As economic interdependence increases, episodes of noncooperation may generate enough costs to stimulate periods of liberalization. It can also be argued that if the bloc becomes big enough, intra-bloc competition may be sufficient to foster economic restructuring, diminishing the constituency for protection of “sunset” industries. The growing market power of the large trading blocs, however, does not bode well either for the fate of the multilateral system or for those left behind by the major RIAs.

The conclusion of the Uruguay Round diminished the likelihood that noncooperative behavior between large RIAs will escalate in the short run. Still, for developing countries, the potential impact of these large trading blocs on their trade and investment opportunities is a legitimate matter for concern. South-South RIAs, organized as defensive coalitions, are unlikely to be strong enough to insulate these countries from the effects of non-cooperative behavior by large blocs. Moreover, the alternative of climbing on the larger blocs bandwagon is only a real possibility (as illustrated by the EC-Eastern Europe Association Agreements and NAFTA) for a limited number of developing countries for the foreseeable future. Sector-oriented North-South coalitions, as illustrated by the Cairns Group with respect to agriculture, could be mentioned as an alternative strategy. The effective capacity of such coalitions to shield participating developing countries from noncooperative behavior by large blocs, however, also seems limited, except possibly where they can exert significant market power in a particular (economically strategic) commodity market. OPEC has attempted to play that role.

\[\text{\textsuperscript{26} For an analysis of the dynamic process behind protection, retaliation and trade wars, see Gould and Woodbridge (1993).}\]
Ironically, the best alternative for developing countries continues to be a working (even though imperfect) multilateral trade system, able to call attention to excesses by those with market power. Although our analysis qualifies the hope for a major strengthening of the multilateral regime, it can continue to play a fundamental role in making the negative effects of discrimination and protection more transparent. Moreover, the future WTO is the only hope for a multilateral agency able to constrain the actions of large trading blocs.
REFERENCES


Appendix: Simultaneous and Overlapping Games

Consider first a standard 3 player game, where each player tries to minimize some loss function $L_i$ defined over his/her policy targets, subject to a set of constraints which define what is (and is not) technically feasible for the available policy instruments to achieve:

\[
\begin{bmatrix}
  x \\
  y \\
  z
\end{bmatrix} =
\begin{bmatrix}
  R_{11} & R_{21} & R_{13} \\
  R_{21} & R_{22} & R_{23} \\
  R_{13} & R_{23} & R_{33}
\end{bmatrix}
\begin{bmatrix}
  u \\
  v \\
  w
\end{bmatrix}
+ 
\begin{bmatrix}
  s_1 \\
  s_2 \\
  s_3
\end{bmatrix}
\] (1)

where $x, y, z$ and $u, v, w$ are the targets and the instruments of players 1, 2 and 3 respectively (and the $S_i$ terms represent external shocks). Thus $R_{ii}$ describe the domestic target responses to policy changes and $R_{ij}$ ($i \neq j$) describe the spillover effects onto other players. Then:

(a) An overlapping game would take the same form except that $R_{13} = 0$ and $R_{31} = 0$ would be imposed with players 1 and 3 interacting solely through their interdependence with the link player, player 2. That interdependence is nevertheless clear to see; the inverse $R^{-1}$, for example, has no zero restrictions.

(b) A simultaneous game is rather different. In a 2-player case, (1) would become

\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  y_1 \\
  y_2
\end{bmatrix} = 
\begin{bmatrix}
  R_{11} & S_{11} & 0 \\
  ? & R_{22} & 0 & S_{22} \\
  T_{11} & 0 & P_{11} & ? \\
  0 & T_{22} & ? & P_{22}
\end{bmatrix}
\begin{bmatrix}
  u_1 \\
  u_2 \\
  v_1 \\
  v_2
\end{bmatrix}
+ 
\begin{bmatrix}
  s_1 \\
  s_2
\end{bmatrix}
\] (2)

where there are two separate theaters of rivalry (subscripts 1 and 2). There may be internal spillovers the two theaters for either player individually, (submatrices marked "?") but there are no spillovers from player 2 in theater 2 to player 1 in theater 1 (submatrices marked "0") and vice versa, or from player 1 in theater 2 to player 2 in theater 1, etc. Hence the $S$ and $T$ submatrices in (2) have been specified to be block diagonal, even if the $R$ and $P$ submatrices are not. Finally, the objectives will also be separable over theaters:

$L_1 = x'_1 Q_{11} x_1 + x'_2 Q_{12} x_2$ and $L_2 = y'_1 Q_{21} y_1 + y'_2 Q_{22} y_2$ where $y_1 \cap y_2 = 0$, $x_1 \cap x_2 = 0$, but $x_1 \cap y_1 \neq 0$.

Then the symmetrical version of (2) will have $R_i = P_i$, $S_i = T_i$;

while direct independence implies $S_i$ and $T_i$ are zero so that the games themselves become separable.
The solutions of all these games are available from standard results. The overlapping game was standard 3-player game with some zero restrictions, and our simultaneous game could be treated as a 4-player game with a (cooperative) coalition between \( x_1 \) and \( x_2 \) and another between \( y_1 \) and \( y_2 \). Hence the point here is to evaluate the pattern of interdependence rather than to construct new types of solutions. In an overlapping game ("widening" the bloc) the interactions are all indirect, via the adjustments that the actions which one player might induce in a link player with whom both players of interest interact directly. Those adjustments by the link player then call for a response by the third player. In a simultaneous game ("deepening" the bloc), the interactions are similarly indirect. Because both players know they interact in each of two theaters, there is scope for trading off concession by one of them in one theater for a gain by that player in the other theater — even if there is no direct spillovers between theaters.

Some examples: i) Widening — Greece and Norway (as a potential new entrant) would not affect each other much directly within the EC "bloc." But both affect Germany through the fiscal contributions which the latter has to make to the EC budget which supports Greece via its cohesion funds and Norway (on accession) through the special agricultural and fisheries program. Thus if Greece were to adjust its policies, so as to produce a better economic performance and higher growth, it would have less call on the cohesion funds. Those adjustments would therefore affect Germany, and Germany’s responses would affect Norway in turn: an overlapping game; ii) Deepening — The EC’s trade and industrial policy and its monetary union program are considered separate initiatives, which don't interact directly. (It is technically possible to have one without the other.) But Germany is prepared to surrender control over its currency for a (possibly inferior) European currency in order to guarantee free access to those wider European markets. A simultaneous game therefore.
<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
<th>Date</th>
<th>Contact for paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>WPS1331 The Myth of Monopoly: A New View of Industrial Structure in Russia</td>
<td>Annette N. Brown, Barry W. Ickos, Randi Rytorman</td>
<td>August 1994</td>
<td>M. Borg 36969</td>
</tr>
<tr>
<td>WPS1332 Poverty and Household Size</td>
<td>Potot Lanjouw, Martin Ravallion</td>
<td>August 1994</td>
<td>P. Cook 33902</td>
</tr>
<tr>
<td>WPS1333 A Test of the International Convergence Hypothesis Using Panel Data</td>
<td>Norman V. Loayza</td>
<td>August 1994</td>
<td>R. Martin 39026</td>
</tr>
<tr>
<td>WPS1334 Taxation, Public Services, and the Informal Sector in a Model of Endogenous Growth</td>
<td>Juan Braun, Norman V. Loayza</td>
<td>August 1994</td>
<td>R. Martin 39026</td>
</tr>
<tr>
<td>WPS1335 Labor Regulations and the Informal Economy</td>
<td>Norman V. Loayza</td>
<td>August 1994</td>
<td>R. Martin 39026</td>
</tr>
<tr>
<td>WPS1337 The Countrywide Effects of Aid</td>
<td>Howard White, Joke Luttik</td>
<td>August 1994</td>
<td>E. Khine 37471</td>
</tr>
<tr>
<td>WPS1338 Commodity Exports and the Adding-Up Problem in Developing Countries: Trade, Investment, and Lending Policy</td>
<td>Maurice Schiff</td>
<td>August 1994</td>
<td>A. Kim 33715</td>
</tr>
<tr>
<td>WPS1339 China's Emergence: Prospects, Opportunities, and Challenges</td>
<td>Andres Boltho, Uri Dadush, Dong He, Shigeru Otsubo</td>
<td>August 1994</td>
<td>J. Queen 33740</td>
</tr>
<tr>
<td>WPS1341 Explaining Pakistan's High Growth Performance Over the Past Two Decades: Can It Be Sustained?</td>
<td>Sadiq Ahmed</td>
<td>August 1994</td>
<td>A. Bhalla 84440</td>
</tr>
<tr>
<td>WPS1342 Winners and Losers in Transition: Returns to Education, Experience, and Gender in Slovenia</td>
<td>Peter F. Orazem, M'Jan Vodopivec</td>
<td>August 1994</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Author</td>
<td>Date</td>
<td>Contact for paper</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>WPS1346 Linking Competition and Trade Policies in Central and Eastern European Countries</td>
<td>Bernard M. Hoekman, Petros C. Mavroidis</td>
<td>August 1994</td>
<td>F. Hatab 35835</td>
</tr>
<tr>
<td>WPS1347 Antitrust-Based Remedies and Dumping in International Trade</td>
<td>Bernard M. Hoekman, Petros C. Mavroidis</td>
<td>August 1994</td>
<td>F. Hatab 35835</td>
</tr>
</tbody>
</table>