Northwestern University

2001 Sheridan Road 580 Leverone Hall Evanston, IL 60208-2014 USA

Discussion Paper #1070

November, 1993

"Multilateral Tariff Cooperation During the Formation of Customs Unions"

Kyle Bagwell
Northwestern University

Robert W. Staiger
University of Wisconsin

www.kellogg.nwu.edu/research/math

CMS-EMS
The Center for
Mathematical Studies
in Economics &
Management Sciences

Discussion Paper No. 1070

Mulitlateral Tariff Cooperation During the Formation of Customs Unions*

by

Kyle Bagwell[1]

and

Robert W. Staiger[2]

November 1993

---

* We have benefited from helpful discussions with Don Davis, John Kennan, Karl Scholz, and seminar participants at Harvard, the NBER 1993 Summer Institute, CEPR's 1993 Workshop on the Political Economy of Trade Negotiations, and IGIER. Staiger gratefully acknowledges financial support from Stanford's Center for Economic Policy Research, and as an Alfred P. Sloan Research Fellow.

1 Northwestern University

2 The University of Wisconsin and NBER

Abstract

We study the implications of customs union formation for multilateral tariff cooperation. We model cooperation in multilateral trade policy as self-enforcing, in that it involves balancing the current gains from deviating unilaterally from an agreed-upon trade policy against the future losses from forfeiting the benefits of multilateral cooperation that such a unilateral defection would imply. The early stages of the process of customs union formation are shown to alter this dynamic incentive constraint in a way that leads to a temporary "honeymoon" for liberal multilateral trade policies. We find, however, that the harmony between customs unions and multilateral liberalization is temporary: Eventually, as the full impact of the emerging customs union becomes felt, a less favorable balance between current and future conditions reemerges, and the liberal multilateral policies of the honeymoon phase cannot be sustained. We argue that this is compatible with the evolving implications of the formation of the European Community customs union for the ability to sustain liberal multilateral trade policies under the General Agreement on Tariffs and Trade.

I. Introduction

The economics of customs unions has formed a central arm of the study of international commercial policy since Viner's (1950) classic treatment of the subject. Inspired by early discussions on the formation of a European customs union, Viner described a taxonomy of the effects of customs union formation which featured his classic distinction between trade diversion and trade creation, but which also included the effects of the formation of customs unions on tariff bargaining with non-member countries, as well as other aspects such as the possibility of significant administrative economies associated with the removal of internal tariffs. The early literature that followed Viner's work focused on the issue of trade diversion versus trade creation, paying little attention to the other elements of Viner's taxonomy. Yet the formation of the European Community (EC) appears to have had important effects that go well beyond the trade creation/trade diversion distinction.

Of particular importance is the impact that the formation of the EC seems to have had on multilateral tariff bargaining. Since its inception, the General Agreement on Tariffs and Trade (GATT) has maintained a somewhat uneasy coexistence with regional exceptions to its Most Favored Nation (MFN) principle through Article XXIV, which permits the formation of both free trade areas and customs unions (which also have common external tariffs) subject to certain stipulations. Nevertheless, the future prospect of an integrated EC market devoid of internal barriers but with common external tariffs appears to have been a major stimulus to the Kennedy Round of multilateral negotiations under GATT initiated in 1963. For example, appearing before the Joint Economic Committee to speak on the Trade Expansion Act of 1962 (which provided U.S. negotiating authority for the Kennedy Round), former Secretary of State

Christian Herter summarized the challenge posed by the newly forming EC customs union and the U.S. response to that challenge as follows:

> ...If we are to go in one direction and Europe go in the other, inevitably, you will find trade barriers growing as between two large free trade areas. With these trade barriers growing, you would find ...the slowing down of trade, both imports and exports...So what is the alternative in the picture? The alternative, to my mind, is to reconcile our policies with those of Europe, with a view to increasing trade on both sides...[p. 12].

Impetus for the Tokyo Round of multilateral GATT negotiations initiated in 1974 can be similarly linked in large part to EC enlargement to include the United Kingdom and others. Indeed, Viner (1950, p. 57) anticipated the potential for multilateral trade-liberalizing effects of customs union formation, arguing that a major goal in the creation of a European customs union was to create an economic entity large enough to deal effectively with the "American menace," including such aspects as "the high and ever-rising American tariff," and "the refusal of the United States to participate in the network of commercial treaties by which European tariffs were being lowered or at least their rise checked."[1]

Recently, there has been renewed interest in the welfare effects of customs unions, and of regional trade agreements more generally, as countries turn with increasing frequency to regional options regarding trade policy.[2] EC 1992 and the signing of the North American Free Trade Agreement (NAFTA) are but the most prominent examples of regional approaches to trade liberalization that have come about over the last decade. Ironically, much of the interest in the effects of such agreements seems to reflect a growing concern that their recent proliferation, with the United States in particular now actively engaged in the pursuit of regional trade agreements, could serve to undermine multilateral cooperation under GATT, even though the early experience with the formation of the EC seems to suggest just the opposite. Indeed, Bhagwati (1991) has observed that the perception in the 1960s was of a general compatibility between regional agreements and GATT, and that only recently has the view that regionalism might be antithetical to multilateral cooperation gained prominence.

We offer here an explanation for this apparent change in perception. We present a model of multilateral tariff cooperation in the presence of customs union formation which predicts a nonstationary relationship between customs union formation and multilateral tariff cooperation. The nonstationarity of this relationship is consistent with the changing perceptions of the compatibility between regional agreements and multilateral tariff cooperation described above. In particular, we argue that early experience with the formation of customs unions and their effects on multilateral tariff cooperation may be a poor guide to the impact that their formation has on sustainable multilateral tariff cooperation in the long run. On the contrary, the model we develop below suggests that the early stages of customs union formation will be associated with a temporary "honeymoon" for liberal multilateral trade policies that cannot be sustained.

We adopt the view, as in Bagwell and Staiger (1990), that (i) enforcement issues are central to an understanding of the dynamic behavior of trade intervention in a world where countries attempt to maintain cooperative trade policies, and that, (ii) in practice, the enforcement of agreed-upon behavior under GATT is limited by the severity of retaliation that can be credibly threatened against an offender by its trading partners. Specifically, we view cooperation in multilateral trade policy as involving a constant balance between, on the one hand, gains from deviating unilaterally from an agreed-upon trade policy, and on the other, the discounted expected future benefits of maintaining multilateral cooperation, with the understanding that the latter would be forfeited in the trade war which followed a unilateral defection in pursuit of the former. In such a setting, changes in current conditions or in expected future conditions can upset this balance, requiring changes in existing trade policy that will bring incentives back into line. In Bagwell and Staiger (1990), we adopted this view of tariff-setting to explore the implications of temporary import surges for the ability to maintain low cooperative tariffs, and interpreted managed trade as a cooperative response to deal with temporary incentive problems in the presence of volatile trade swings. We explore in this and a companion paper (Bagwell and Staiger, 1993) the sense in which the formation of regional trade agreements upsets the balance between current and future conditions, and trace through the dynamic ramifications of these effects for multilateral cooperation.

A crucial focus of our analysis is the period of transition, during which the regional agreement is being negotiated, and then implemented. Both because regional trade agreements typically involve a lengthy staging of internal tariff reductions and, in the case of customs unions, external tariff harmonization, and because trade patterns take time to reflect changes in trade barriers in any event, there will inevitably be a lag between the conclusion of negotiations and ratification of the agreement on the one hand, and final changes in internal and external tariffs and trading patterns reflecting the fully implemented regional agreement on the other.[3] Together with the period of negotiation, this lag creates a period of transition within which, at least initially, the formation of the regional trade area has its biggest impact on the expected future behavior of tariffs and trade patterns rather than on current conditions. It is this basic observation that is central to our results.

To understand our main findings, it is helpful first to categorize two of the principal effects of a regional agreement. A first consequence of such an agreement is the *trade diversion effect*, whereby the removal of internal tariffs between member countries acts to enlarge intra-member trade volume and reduce the volume of trade between member and nonmember countries.[4] A second effect is the *market power effect*. While the trade diversion effect arises for both free trade agreements and customs unions, the market power effect is particular to the formation of customs unions: Under a customs union, the member countries adopt a common external tariff on imports, and this in turn enables them credibly to impose a higher import tariff on their multilateral trading partners than if their external tariff were not harmonized, should such a punitive tariff be desired.[5]

In our companion paper, we argue that the trade diverting effect of free trade agreements leads to higher multilateral tariffs and acute multilateral trade tensions during the transition period over which such agreements are negotiated and implemented. Intuitively, during the period of transition, trade volume between member and non-member countries is still large, as internal tariffs between member countries have not yet been eliminated. On the other hand, member and nonmember countries recognize that they will trade less with one another in the future, once the agreements are implemented. Thus, during the transition phase, the incentive to deviate unilaterally is large as compared to the now smaller discounted future value of cooperation. To maintain some measure of cooperation between member and nonmember countries it is then necessary to raise the transition-period tariffs between the two sets of countries, reducing the volume of their trade and the associated incentive to defect.

In the present paper, we consider the formation of customs unions. While generally both trade diversion and market power effects will be associated with customs union formation, we provide a model in which the market power effect is isolated. In this setting, we show that the emergence of customs unions will be associated with temporarily reduced multilateral trade tensions between member and non-member countries, and consequently, to a temporary "honeymoon" for liberal multilateral trade policies. This easing of tensions arises during the period of transition, when the current degree of market power possessed by each member country (and hence the current incentive to deviate unilaterally from an agreed-upon multilateral tariff) is more or less unchanged at the same time that the expected future degree of market power possessed by each member country (and hence the value to non-member countries of maintaining future multilateral cooperation) has increased. Intuitively, under such conditions, non-member countries are less apt to take a confrontational stance in trade disputes with member countries of the emerging customs union, as the risks of a possible trade war with such countries now pose a greater deterrent to confrontation than they once did. Our results suggest, however, that the harmony between customs unions and multilateral liberalization is temporary: Eventually, as the impact of the emerging customs union on the degree of market power becomes felt, a less favorable balance between current and expected future conditions reemerges, and liberal multilateral trade policies cannot be sustained.

Together, the two papers offer two distinct explanations for the changing perception of the consequences of greater regionalism for multilateral tariff cooperation. Our companion paper suggests that the major episodes of regionalism in recent years, embodied in EC 1992 and NAFTA, are fundamentally different from the original formation and subsequent enlargement of the EC customs union with regard to their implications for multilateral tariff cooperation. In particular, the recent wave of regionalism would seem to embody primarily a trade diversion effect but not a market power effect since its primary focus is on dropping internal barriers rather than harmonizing external trade policy; thus, the process of implementing these regional agreements should be associated with high multilateral trade tensions. The present paper, by contrast, suggests that the current disillusionment with respect to the compatibility of multilateral tariff cooperation and regional agreements in general reflects the passing of a honeymoon period associated with the original EC customs union formation and its subsequent enlargement, over which the market power effect associated with customs union formation made possible lower multilateral tariffs.

The remainder of the paper is devoted to establishing and elaborating on these points. The next section sets out the basic model within which we will study the formation of customs unions, and establishes several properties in a stationary setting that will be useful in the dynamic non-stationary analysis to follow. Section III then characterizes the dynamic behavior of equilibrium multilateral tariffs in the non-stationary environment of emerging customs unions. Section IV derives various comparative statics results , while Section V concludes with a discussion of the original EC customs union formation and the design of Article XXIV in light of our results.

II. Multilateral Tariff Determination in Stationary Environments

In this section we develop and explore the properties of a stationary model of multilateral tariff formation in the presence of customs unions. In the next section we will then describe the nonstationarities that arise when the process of customs union formation is explicitly considered.

A. A Static Customs Union Model

To direct attention to the main effects, we analyze a simple, partial-equilibrium exchange economy. There are two types of countries, "foreign" countries denoted by a "*," of which there are a total of K, and "domestic" countries denoted by the absence of a "*," of which there are also K. The K foreign countries are grouped symmetrically into R customs unions or "regions," while the K domestic countries are similarly grouped symmetrically into R regions.[6] Thus, $k \equiv K/R$ gives the number of member countries per region.

There are only two goods, referred to as the domestic and the foreign export goods, respectively, and there exist two units of each good in total in the world.[7] Each domestic country is endowed with $2/K$ units of the domestic export good and none of the foreign export good, while each foreign country is endowed with $2/K$ units of the foreign export good and none of the domestic export good. On the demand side, we assume that demand functions are symmetric across countries and that the demand for any product i in any country j is independent of the price of the other good in country j. Thus, we suppose that each country j has demand for good i of

$$(1) \qquad C(P_j^i) = \frac{1}{K}[\alpha - \beta P_j^i]; \quad C(P_j^{i*}) = \frac{1}{K}[\alpha - \beta P_j^{i*}]$$

where P_j^i is the price of good i in domestic country j, and P_j^{i*} is the price of good i in foreign country j.

We note that our assumptions on endowments and demands imply that domestic countries do not trade with each other, and likewise that foreign countries do not trade with each other. Thus, customs union formation among domestic countries and among foreign countries entails no trade diversion, but occurs rather among competing suppliers of a common export good, and competing demanders of a common import product. This property of the model allows us to abstract from trade diversion/trade creation issues that are common to both free trade areas and customs unions (and which we treat in our companion paper, Bagwell and Staiger, 1993), and allows us instead to highlight the aspect of customs union formation that distinguishes it from the formation of free trade areas, namely, the harmonization of external tariffs and the market power effects that this implies.

With domestic and foreign regions symmetric, we now proceed to characterize the static equilibrium from a domestic region's perspective. Recalling that a common import tariff is selected by all members of a customs union or region, we may let τ_r represent the (specific) import tariff levied by region r, where $r = 1,...,R$.[8] Given the endowment structure described above, we also may simplify the notation and define prices as follows: P_{xr} is the price in region r of the good that r exports, and P_{mr} is the price in region r of the good that r imports. Finally, since import taxes are assumed not to discriminate across sources and since regions are otherwise symmetric, a given export good will face the same constellation of import tariffs regardless of the good's region-of-origin. Thus, the export good will have a single price in all exporting regions, and so we may remove the r subscript and let P_x denote the export good's price in any export region. Of course, import prices may differ across regions, as different regions may select different import tariffs.

Now, for any given good and associated importing region r, we have $P_{mr} = P_x + \tau_r$, provided τ_r is non-prohibitive. This, together with the market equilibrium condition

$$2 = \alpha - \beta P_x + \sum_{r=1}^{R} (1/R)(\alpha - \beta P_{mr})$$

gives the equilibrium prices, $\hat{P}_x(R,\bar{\tau})$ and $\hat{P}_{mr}(R,\bar{\tau})$, and per-region import quantities, $\hat{M}_r(R,\bar{\tau}) \equiv kC(\hat{P}_{mr}) = (1/R)(\alpha-\beta\hat{P}_{mr})$, for the given good, when it faces the vector of import tariffs $\bar{\tau} \equiv (\tau_1,\ldots,\tau_R)$. These solutions are:

(2) $$\hat{P}_x(R,\bar{\tau}) = \frac{\alpha-1}{\beta} - \frac{1}{2R}\sum_{\ell=1}^{R}\tau_\ell \quad ; \qquad \hat{P}_{mr}(R,\bar{\tau}) = \frac{\alpha-1}{\beta} - \frac{1}{2R}\sum_{\ell=1}^{R}\tau_\ell + \tau_r$$

(3) $$\hat{M}_r(R,\bar{\tau}) = \frac{1}{R} - \frac{\beta}{R}[\tau_r - \frac{1}{2R}\sum_{\ell=1}^{R}\tau_\ell]$$

With (2) and (3) in place, we can define regional welfare per member country. Specifically, a domestic country's welfare when domestic and foreign regions respectively select import tariffs $\bar{\tau} \equiv (\tau_1,\ldots,\tau_R)$ and $\bar{\tau}^* \equiv (\tau_1^*,\ldots,\tau_R^*)$ is given by

(4) $$W(R,\bar{\tau};\bar{\tau}^*) = \int_{\hat{P}_{mr}(R,\bar{\tau})}^{\alpha/\beta} C(P)dP + \int_{\hat{P}_x(R,\bar{\tau}^*)}^{\alpha/\beta} C(P)dP + \int_{0}^{\hat{P}_x(R,\bar{\tau}^*)} (2/K)dP + \tau_r[R/K]\hat{M}_r(R,\bar{\tau})$$

which corresponds to the consumer surplus received on the foreign export good, the consumer surplus received on the domestic export good, the producer surplus received on the domestic export good, and the tariff revenue received on the foreign export good, respectively, when r is the domestic region to which the country belongs. Welfare is defined symmetrically for any foreign country.

The main features of W are summarized as follows. First, W is maximized at the regional tariff choice

(5) $$\tau_r^D(R,\sum_{\substack{\ell=1\\ \ell\neq r}}^{R}\tau_\ell) = \frac{2R}{\beta(4R^2-1)} + \frac{1}{(4R^2-1)}\sum_{\substack{\ell=1\\ \ell\neq r}}^{R}\tau_\ell,$$

which is the best "defect" tariff for the country - and equivalently for the region r to which the country belongs - when other domestic regions select import tariffs τ_ℓ, where $\ell\neq r$. Thus, the optimal tariff is positive, and it is also independent of the tariff levels selected by trading partners (namely <u>foreign</u> regions). The latter property is a consequence of our assumptions that demands are independent across goods and that export taxes are not possible.

Second, notice that the optimal tariff for the domestic region r is increasing in the tariffs selected by other domestic regions. Intuitively, this is because the respective domestic regions "compete" for imports: When other domestic regions raise their import tariffs, additional import volume is released for region r, and this increases the incentive for region r to raise its own tariff and receive even greater tariff revenue.

Third, an interesting pattern of externalities is apparent. Examination of (4) reveals that an increase in a foreign-region tariff reduces the welfare of any domestic country, as the domestic country then receives lower producer surplus. On the other hand, there is a positive externality between similarly-endowed countries: As a domestic region raises its tariff, more import volume is directed to other domestic regions, and the countries in these regions experience a welfare gain. When domestic and foreign regions all select the same tariff, however, it is direct to show that the former effect dominates, in that each country's welfare increases as the symmetric tariff is reduced.

Finally, let us now define the static tariff game to be the game in which each region simultaneously selects an (external) import tariff in order to maximize its welfare per member country. Calculations reveal that the symmetric Nash equilibrium for the static tariff game occurs when all regions select the positive import tariff:

$$\hat{\tau}^N(R) = \frac{2}{\beta(4R-1)} \tag{6}$$

This expression exposes the "market power effect" of customs union formation: When customs unions expand in the sense of being fewer in number and larger in size (i.e., as R decreases), the newly-joined similarly-endowed countries internalize their joint incentive for higher import tariffs, and consequently the Nash tariff rises.[9]

B. A Stationary Dynamic Customs Union Model

We now consider a stationary dynamic tariff game, which is defined by the infinite repetition of the static tariff game described above. In each period the regions observe all previous import tariff selections and simultaneously choose import tariffs. For the reasons given above, we continue to assume that each region applies the same tariff to imported goods from all sources in any given period. The game is stationary in the sense that none of the model's parameters changes through time. Let $\delta \varepsilon (0,1)$ denote the discount factor between periods.

In order to express our ideas in a simple manner, we focus on a particular class of subgame perfect equilibria for the stationary dynamic tariff game. Specifically, we consider equilibria in which (i) symmetric stationary non-negative import tariffs are selected along the equilibrium path, meaning that in equilibrium all regions select the same import tariff in each period, and (ii) if a deviation from this common tariff occurs, then in the next period and forever thereafter the regions revert to the Nash equilibrium tariffs of the static tariff game. We then refer to the most-cooperative equilibrium of the stationary dynamic tariff game as the subgame perfect equilibrium which yields the lowest possible equilibrium tariff while satisfying restrictions (i) and (ii). The corresponding import tariff is then termed the most-cooperative tariff for the stationary dynamic tariff game.[10]

In a dynamic model, regions have the possibility of supporting a cooperative tariff, τ^c with $\tau^c < \hat{\tau}^N(R)$, since any attempt to raise the current-period tariff will be greeted with retaliatory (Nash) tariffs from other regions in future periods. Intuitively, a cooperative tariff

τ^c can then be supported in an equilibrium for the stationary dynamic tariff game if the one-time incentive to cheat is sufficiently small relative to the future value of maintaining a cooperative relationship among trading regions.

To formalize this intuition, let us first examine the incentive a region has to cheat. For a fixed cooperative tariff $\tau^c < \hat{\tau}^N(R)$, and given the class of subgame perfect equilibria upon which we focus, if a region is to deviate and select a tariff other than τ^c, then it will deviate to its best-response tariff, as defined in (5). We now simplify the notation slightly and use $\tau^D(R,\tau^c)$ to represent the best-response tariff for a given region when all other same-type regions are selecting the cooperative tariff, τ^c. The per member country gain when the associated region cheats is then given by:

(7) $$\Omega(R,\tau^c) \equiv W(R,(\tau^D,\bar{\tau}^c);\bar{\tau}^c) - W(R,\bar{\tau}^c,\bar{\tau}^c),$$

where τ^c is a scalar and $\bar{\tau}^c$ is a vector in which the scalar τ^c is present in each component.[11] Intuitively, Ω is the difference between (i) the per-member country welfare when the region to which the country belongs selects the best-defect tariff while all other regions - domestic and foreign - continue to select the cooperative tariff τ^c, and (ii) the per-member country welfare when the country's region cooperates in choosing the same cooperative tariff as do all other regions.

When a region cheats, however, it also causes future welfare to drop, and we now examine this cost of cheating. Define the one-period value to cooperation per member country to be:

(8) $$\omega(R,\tau^c) \equiv W(R,\bar{\tau}^c;\bar{\tau}^c) - W(R,\bar{\hat{\tau}}^N(R);\bar{\hat{\tau}}^N(R)).$$

where $\bar{\hat{\tau}}^c(R)$ is an R dimensional vector in which the scalar $\hat{\tau}^N(R)$ as defined in (6) is present in each component. Then the cost to cheating is $\delta/(1-\delta)\cdot\omega(R,\tau^c)$, since once a region defects and selects a high import tariff, cooperative tariffs are thereafter replaced by the higher Nash tariffs.

Using (7) and (8), the fundamental "no-defect" condition is that the benefit of cheating be less than the discounted future value of cooperation, or:

$$(9) \qquad \Omega(R,\tau^c) \leq \frac{\delta}{1-\delta}\omega(R,\tau^c).$$

Any cooperative tariff τ^c that satisfies (9) can be supported in a subgame perfect equilibrium of the stationary dynamic tariff game.

Our interest lies in the most-cooperative tariff $\hat{\tau}^c$, which is the smallest nonnegative tariff that satisfies (9). To characterize this tariff, we first investigate the properties of $\Omega(R,\tau^c)$ and $\frac{\delta}{1-\delta}\omega(R,\tau^c)$. Calculations reveal that:

$$(10) \qquad \Omega(R,\tau^c) = \frac{4-4\beta(4R-1)\tau^c+\beta^2(4R-1)^2(\tau^c)^2}{8\beta K(4R^2-1)}$$

Using (10), it follows that:

$$(11) \qquad \frac{\partial\Omega(R,\tau^c)}{\partial R} < 0 \text{ if } \tau^c < \hat{\tau}^N(R)$$

$$(12) \qquad \frac{\partial\Omega(R,\tau^c)}{\partial\tau^c} < 0 \text{ and } \frac{\partial^2\Omega(R,\tau^c)}{\partial^2\tau^c} > 0 \text{ if } \tau^c < \hat{\tau}^N(R).$$

Thus, a decrease in R, which corresponds to more concentrated customs union formation, acts to raise the benefit from defection, since a given customs union has greater power to affect world prices with its tariff increase. Notice also that lower cooperative tariffs heighten the incentive to cheat, because a deviation to the Nash tariff then represents a more significant tariff increase.

Calculations also reveal that:

$$(13)\qquad \frac{\delta}{1-\delta}\,\omega(\mathrm{R},\tau^c)=\frac{\delta}{1-\delta}\,\frac{4-\beta^2(4\mathrm{R}-1)^2(\tau^c)^2}{4\beta\mathrm{K}(4\mathrm{R}-1)^2}>0 \text{ if } \tau^c<\hat{\tau}^N(\mathrm{R}).$$

Using (13), it follows that:

$$(14)\qquad \frac{\partial}{\partial \mathrm{R}}\left(\frac{\delta}{1-\delta}\,\omega(\mathrm{R},\tau^c)\right)<0 \text{ if } \tau^c<\hat{\tau}^N(\mathrm{R})$$

$$(15)\qquad \frac{\partial}{\partial \tau^c}\left(\frac{\delta}{1-\delta}\,\omega(\mathrm{R},\tau^c)\right)<0 \text{ and } \frac{\partial^2}{\partial {\tau^c}^2}\left(\frac{\delta}{1-\delta}\,\omega(\mathrm{R},\tau^c)\right)<0 \text{ if } \tau^c>0$$

Thus, as R falls and customs unions become fewer in number and larger in size, the non-cooperative Nash tariff rises and so the cost of a trade war grows. However, higher cooperative tariffs lower the discounted value of future cooperation.

The determination of the most-cooperative tariff is now easily illustrated by Figure 1. Observe in Figure 1 that the no-defect condition (9) is satisfied for all $\tau^c \varepsilon [\hat{\tau}^c_s(\mathrm{R}), \hat{\tau}^N(\mathrm{R})]$. These are the set of tariffs that are supportable as subgame perfect equilibrium tariffs for our stationary dynamic tariff game, given the class of equilibria upon which we focus. Solving (9) for the tariff that gives equality yields the most-cooperative tariff, which is given by:

$$(16)\qquad \hat{\tau}^c_s(\mathrm{R})=\hat{\tau}^N(\mathrm{R})\left\{\frac{(4\mathrm{R}-1)^2(1-\delta)-2\delta(4\mathrm{R}^2-1)}{(4\mathrm{R}-1)^2(1-\delta)+2\delta(4\mathrm{R}^2-1)}\right\}.$$

Two observations can be made about the equilibrium most-cooperative tariff in the stationary dynamic game. First, note that $\hat{\tau}^c_s(\mathrm{R})$ is decreasing in δ, with $\hat{\tau}^c_s = 0$ at $\delta = \dfrac{(4\mathrm{R}-1)^2}{2(4\mathrm{R}^2-1)+(4\mathrm{R}-1)^2} \equiv \delta^*(\mathrm{R})$. This decreasing relationship is intuitive: As δ increases, the discounted value of future cooperation is enhanced, and so a lower tariff can be supported (despite the consequent greater incentive to cheat). To avoid cases in which the most-cooperative tariff corresponds to either of the extreme polar outcomes of free trade or the non-cooperative tariff $\hat{\tau}^N(\mathrm{R})$, we assume in what follows that $\delta \varepsilon (0, \delta^*(\mathrm{R}))$ for all R that we consider.

Second, whether or not the existence of smaller numbers of larger customs unions is good or bad for multilateral tariff cooperation between regions in the stationary dynamic game depends on the discount factor δ. This makes sense, since customs union formation (a falling R) increases both the onetime benefit from cheating $(\Omega(R,\tau^c))$ and the cost of a tariff war $(\frac{\delta}{1-\delta}\omega(R,\tau^c))$. If countries don't weigh the future too heavily, then the effect of customs union formation on the one-time benefit from cheating dominates its effect on the cost of a tariff war, and the most-cooperative stationary tariff must be raised to keep the incentive constraint in check. In the limit, when $\delta = 0$, the most-cooperative tariff is simply $\hat{\tau}^N(R)$, which by (6) is declining in R. On the other hand, for δ sufficiently high, customs union formation may be good for multilateral tariff cooperation in the stationary dynamic game, as the effects of customs union formation on the added benefits from cheating become overwhelmed by the added cost of a tariff war. In the limit, as δ approaches $\delta^*(R)$, customs union formation must be good for multilateral cooperation, since $\delta^*(R)$ in increasing in R; i.e., customs union formation lowers the discount factor at which free trade is sustainable.

The dependence of the most-cooperative stationary tariff on the discount factor is illustrated in Figure 2, where customs union expansion corresponds to a reduction in the number of customs unions from R_0 to R_1, where $R_0 > R_1$. As the figure indicates, when countries do not weigh the future heavily, customs union expansion results in a higher most-cooperative stationary tariff; but, when the discount factor is large, customs union expansion is reflected in a lower most-cooperative stationary tariff. It follows that a critical discount factor, $\tilde{\delta}(R_0,R_1)$, must exist at which the most-cooperative stationary tariff is neutral with respect to customs union expansion, and below which customs union expansion results in a higher most-cooperative stationary tariff.[12]

In what follows we assume $\delta < \tilde{\delta}(R_0,R_1)$ so that customs union formation is bad for multilateral cooperation in the stationary dynamic game. We choose to do this for several reasons. First, for simple cases of customs union formation, this captures most of the relevant range of discount factors.[13] For example, in the case where customs union formation takes R from $R_0 = 2$ to $R_1 = 1$, we require that $\delta < \delta^*(R_1) = 0.60$ to ensure that free trade cannot be sustained when $R=1$. If $\delta < \tilde{\delta}(R_0,R_1) = 0.58$ also, the customs union formation represented by the move from $R_0=2$ to $R_1=1$ will be bad for multilateral cooperation in the stationary dynamic game. Second, as we will show below, even adopting this "pessimistic" view of the stationary effect of customs unions on multilateral cooperation, there will nevertheless be a "honeymoon" phase during which multilateral tariffs first fall before they later rise.

C. Summary

We summarize the results of this section with the following proposition:

Proposition 1:

(a) The Nash equilibrium of the static tariff game occurs when each region sets an external import tariff of $\hat{\tau}^N(R) = \dfrac{2}{\beta(4R-1)}$.

(b) The most-cooperative equilibrium of the stationary dynamic game occurs when each region sets an external tariff of $\hat{\tau}^c_s(R) = \hat{\tau}^N(R)\left\{\dfrac{(4R-1)^2(1-\delta)-2\delta(4R^2-1)}{(4R-1)^2(1-\delta)+2\delta(4R^2-1)}\right\}$, provided that $\delta\varepsilon(0,\delta^*(R))$.

III. The Formation of Customs Unions

We turn now to a dynamic model in which, at some point in time, customs union expansion occurs, in that the number of domestic regions and also the number of foreign regions decreases. While understanding the timing of customs union formation is important in its own right, it is a problem that has many dimensions, and a proper treatment is well beyond the scope of this paper. Instead, we assume that the process of customs union expansion occurs randomly and for exogenous, political reasons. The possibility that the number of regions may change through time introduces a nonstationarity into the dynamic interaction between countries, and our focus here is on how customs union expansion affects the ability of domestic and foreign regions to cooperate multilaterally in the setting of low tariffs.

A. The Customs-Union Model

We envision a trading relationship that passes through three phases. In phase 1, there are R_0 domestic regions and also R_0 foreign regions. The domestic and foreign regions trade with one another just as above. The countries are aware, however, that a time may come at which it becomes politically feasible for customs union expansion to occur, both among domestic countries and among foreign countries. Phase 2 corresponds to a transition phase, in which there are still R_0 regions of each type, but in which customs-union-expansion discussions have already commenced. Finally, in phase 3, the customs-union-expansion talks are completed, the new regions are fully implemented, and there are now R_1 regions of each type, where $R_1 < R_0$. This final set of trading patterns then persists into the infinite future.

Equilibria of repeated games with nonstationarities are often difficult to characterize. For tractability, therefore, we impose two assumptions. First, we assume that the transition process obeys a constant-hazard-rate (i.e. stationary-Markov) property. Namely, if the countries are in phase 1 at any date t, then $\rho \varepsilon (0,1)$ is the constant probability that they will be in phase 2 at date $t+1$. Similarly, $\lambda \varepsilon (0,1)$ is the fixed conditional probability of transition from phase 2 to phase 3. Note that ρ and λ are assumed independent of the tariff history between countries. As will become clear, while the constant-hazard-rate assumption is not completely general, it does make possible some very precise predictions. Second, we assume that the domestic and foreign countries pass through their respective phases at the same dates. This enables us to exploit symmetry between the two country types, and thereby simplifies the analysis. We return to this assumption in Section V, and indicate the manner in which our results extend under asymmetric customs union expansion.

The customs-union game is now defined as the infinite-period game, in which countries pass through the described three phases, and regions select tariffs in each period with the goal of maximizing the welfare of their respective current-member countries, where at any given date all regions are perfectly informed as to past tariffs and the current phase of the game. For this game, we examine a class of subgame perfect equilibria, for which (i) along the equilibrium path, in any given phase of the game, the domestic and foreign regions select a common import tariff for all dates within that phase; and (ii), if at any point in the game a deviation from the equilibrium tariff for the corresponding phase occurs, then in the next period and forever thereafter all regions select the Nash equilibrium tariffs of the relevant static tariff game.[14]

For such equilibria, there will be three cooperative tariff levels, with each corresponding to a different phase. Let τ_1^c, τ_2^c and τ_3^c refer to the cooperative tariff levels in phases 1, 2, and 3 respectively. Once again, we look for a most-cooperative equilibrium, and we solve for the associated most-cooperative tariffs, $\hat{\tau}_1^c$, $\hat{\tau}_2^c$ and $\hat{\tau}_3^c$. The most-cooperative tariffs may be found using a recursive solution approach. Specifically, we first identify the no-defect condition for phase 3 and find the lowest tariff that can be supported in this phase in an equilibrium of the desired class. With $\hat{\tau}_3^c$ thus determined, we next turn to phase 2, represent the relevant no-defect condition for this phase, and then solve for the phase-2 most-cooperative tariff, $\hat{\tau}_2^c$. Finally, having solved for the most-cooperative tariffs in phases 2 and 3, we characterize next the no-defect condition for phase 1 and solve for the lowest tariff in this phase that doesn't invite cheating. The resulting tariff is the most-cooperative phase-1 tariff, $\hat{\tau}_1^c$. This recursive method does indeed identify the most-cooperative tariffs for the overall game, since the discounted value of cooperation as viewed from any given phase rises as future cooperative tariffs drop. Thus, by selecting the lowest possible cooperative phase-3 tariff, we raise the cost to countries of defecting and igniting a trade war in phases 2 and 1, and we thereby make possible lower tariffs in these phases as well. Similarly, a lower phase-2 tariff makes it possible to support lower cooperative tariffs in phase 1.

We are now ready to formally represent the no-defect conditions for each of the three phases. Let us begin with phase 3. At any date within this phase, there are R_1 regions of each kind, the future is known to be stationary, and the no-defect condition is:

(17) $$\Omega(R_1, \tau_3^c) \leq \frac{\delta}{1-\delta}\omega(R_1, \tau_3^c).$$

This has the same form as (9), the no-defect condition in our stationary model, except that the number of regions is now R_1. Thus, it follows that $\hat{\tau}_3^c = \hat{\tau}_s^c(R_1)$; in other words, the phase-3 most-cooperative tariff is the most-cooperative stationary tariff for a world in which there are R_1 domestic and foreign customs unions, respectively.

Consider now phase 2. The no-defect condition for this phase is:

$$(18a) \qquad \Omega(R_0,\tau_2^c) \leq \delta \sum_{n=1}^{\infty} \lambda(1-\lambda)^{n-1} \left[\sum_{q=1}^{n-1} \delta^{q-1} \omega(R_0,\tau_2^c) + \sum_{k=n}^{\infty} \delta^{k-1} \omega(R_1,\hat{\tau}_3^c) \right]$$

where n indexes the period at which phase 3 begins, with $n=1$ meaning that phase 3 begins in the next period, and where q and k correspond to periods within phases 2 and 3, respectively.[15] Observe in phase 2 that there are R_0 domestic and foreign regions, respectively, and this is reflected in the left hand side of (18a). With some further simplification, the phase-2 no-defect condition may be written as:

$$(18b) \qquad \begin{aligned} \Omega(R_0,\tau_2^c) &\leq \frac{(1-\lambda)\delta}{[1-(1-\lambda)\delta]} \omega(R_0,\tau_2^c) + \frac{\lambda\delta/(1-\delta)}{[1-(1-\lambda)\delta]} \omega(R_1,\hat{\tau}_3^c) \\ &\equiv V_2(\tau_2^c; \lambda, \delta, R_0, R_1) \end{aligned}$$

where V_2 is defined to be the expected discounted value to future cooperation, as viewed in phase 2. Intuitively, V_2 is a weighted average of $\omega(R_0,\tau_2^c)$ and $\omega(R_1,\hat{\tau}_3^c)$, since a defection in phase 2 induces a trade war, thus sacrificing the cooperative welfare that could have been received in the remainder of phase 2 (at the tariff level τ_2^c) as well as the cooperative welfare that would have been forthcoming once phase 3 was entered (at the tariff level $\hat{\tau}_3^c$). The lowest tariff satisfying (18b) defines $\hat{\tau}_2^c$.

Finally, we come to the phase-1 no-defect condition:

$$(19a) \qquad \Omega(R_0,\tau_1^c) \leq \delta \sum_{s=1}^{\infty} \rho(1-\rho)^{s-1} \left[\sum_{t=1}^{s-1} \delta^{t-1} \omega(R_0,\tau_1^c) + \delta^{s-1}(\omega(R_0,\hat{\tau}_2^c) + V_2(\hat{\tau}_2^c; \lambda, \delta, R_0, R_1)) \right]$$

where s indexes the period at which phase 2 begins, with $s=1$ meaning that phase 2 begins in the next period, and where t represents periods within phase 1. Using (18b), we may rewrite (19a) as:

$$\Omega(R_0,\tau_1^c) \le \frac{(1-\rho)\delta}{[1-(1-\rho)\delta]}\,\omega(R_0,\tau_1^c) + \frac{\rho\delta}{[1-(1-\rho)\delta]}\,\frac{\{\omega(R_0,\hat{\tau}_2^c) + [\lambda\delta/(1-\delta)]\,\omega(R_1,\hat{\tau}_3^c)\}}{[1-(1-\lambda)\delta]} \tag{19b}$$

$$\equiv V_1(\tau_1^c;\rho,\lambda,\delta,R_0,R_1)$$

where V_1 gives the expected discounted value to future cooperation as viewed from phase 1. Note now that V_1 is a weighted average of $\omega(R_0, \tau_1^c)$, $\omega(R_0,\hat{\tau}_2^c)$ and $\omega(R_1, \hat{\tau}_3^c)$, reflecting the fact that a deviation in phase 1 sacrifices the ability to cooperate in the remainder of phase 1 as well as throughout phases 2 and 3. The smallest tariff satisfying (19b) is then defined to be $\hat{\tau}_1^c$.

B. Characterization of the Most-Cooperative Tariffs

We are prepared now to characterize the three most-cooperative tariff levels, so that their relative magnitudes may be determined. In this way, we will be able to assess the consequences of customs union expansion for multilateral tariff cooperation.

The tariffs are characterized in a recursive fashion, beginning with the phase-3 most-cooperative tariff. As discussed above, this tariff is simply the most-cooperative stationary tariff for a world in which there are R_1 customs unions of each country type:

Lemma 1: $0 < \hat{\tau}_3^c = \hat{\tau}_s^c(R_1) < \hat{\tau}^N(R_1)$.

Thus, over the range of discount factors which we consider, the phase-3 most-cooperative tariff lies between free trade and the Nash tariff (for R_1 regions).

Consider now the phase-2 most-cooperative tariff. To characterize this tariff, we first record the following:

Lemma 2: $\omega(R_1,\hat{\tau}_s^c(R_1)) > \omega(R_0,\hat{\tau}_s^c(R_0))$.

This lemma states that the per-period value of cooperation at the most-cooperative stationary tariff is highest when there are fewer regions. Intuitively, two forces are at work in the proof of this lemma. First, for any fixed cooperative tariff, a smaller number of regions results in a greater value of cooperation, as (14) demonstrates, since under the market power effect a trade war is more damaging when regions are larger in size and fewer in number. Second, under the range for δ which we consider, the most-cooperative stationary tariff is higher when there are fewer regions, and, as (15) indicates, this higher cooperative tariff acts to diminish the gain from cooperation. The lemma establishes that the direct effect of a smaller number of regions outweighs the indirect effect of a higher cooperative tariff, and so the per-period value of cooperation at the most-cooperative stationary tariff is higher when there are fewer regions. A proof of this lemma is found in the Appendix.

With this lemma in place, we can record some properties of the V_2 function.

(20) $$\frac{dV_2}{d\tau_2^c}(\tau_2^c;\lambda,\delta,R_0,R_1)<0$$

(21) $$V_2(\hat{\tau}_s^c(R_0);\lambda,\delta,R_0,R_1)>\frac{\delta}{1-\delta}\,\omega(R_0,\hat{\tau}_s^c(R_0)).$$

Intuitively, the discounted value of future cooperation as viewed from phase 2 is lower when the cooperative tariff in phase 2 is higher, since in this case cooperation in phase 2 is already modest, and so a trade war instigated in this phase would result in less welfare loss during any remaining periods of phase 2. More formally, (20) follows directly from (18b) and (15). As for (21), observe from Lemma 1 and (18b) that $V_2(\hat{\tau}_s^c(R_0);\cdot)$ is a weighted average of

$\omega(R_0,\hat{\tau}^c_s(R_0))$ and $\omega(R_1,\hat{\tau}^c_s(R_1))$, the per-period values of cooperation at the most-cooperative stationary tariffs when there are R_0 and R_1 regions, respectively. Now, Lemma 2 tells us that the per-period value of cooperation at the most-cooperative stationary tariff is greatest when the number of regions is small, and it thus must be that the discounted value of future cooperation as viewed from phase 2 exceeds that obtained in a stationary setting with a large number of regions.[16]

It is convenient to ensure that free trade is never supportable, and that the no-defect condition therefore always binds with equality. To guarantee this, we further require that:

(22) $\Omega(R_0,0) > \delta/(1-\delta)\omega(R_1,0)$

which in turn implies that $\Omega(R_0,0) > V_2(0;\cdot)$, indicating that free trade is not supportable in phase 2.[17] Observe that (22) will clearly hold under our existing assumptions if $R_0 - R_1$ is small, since in that event (22) basically requires again that free trade not be supportable in stationary environments. More generally, (22) is satisfied if δ is restricted to lie below some critical level, $\delta^{\dagger}(R_0,R_1)$.[18] Thus, (22) may be understood as a further strengthening of our small-δ orientation.

We are now prepared to characterize $\hat{\tau}^c_2$, which is the lowest tariff for which $\Omega(R_0,\tau^c_2) \leq V_2(\tau^c_2;\lambda,\delta,R_0,R_1)$. As Figure 3 illustrates, given (20), (21) and (22), $\hat{\tau}^c_2$ must lie strictly between 0 and $\hat{\tau}^c_3$:

Lemma 3: $0 < \hat{\tau}^c_2 < \hat{\tau}^c_3$.

Thus, a "honeymoon" phase occurs while the customs union expansion is being negotiated and phased in, as multilateral tariffs are low throughout this process. Once the customs-union expansion is fully implemented, however, multilateral tariff cooperation deteriorates and higher tariffs remain.

Lemma 3 may be understood in the following intuitive terms. If the number of regions were stationary through time, with R_0 regions of each type of country, then the regions could support a tariff level of $\hat{\tau}_s^c(R_0)$, as this is the tariff that just balances a region's immediate incentive to cheat against the long-term consequences of a trade war. This situation now may be contrasted with that which arises in the transition phase of the customs-union game. At this point, the various countries are still aligned with R_0 regions of each country type, and so the incentive to cheat is the same as in the associated dynamic stationary game, but the countries are also aware that customs-union expansion - and the greater punishment that this makes possible - will soon occur. Thus, as compared to the stationary environment that supports $\hat{\tau}_s^c(R_0)$, in the transition phase of the customs union game, the countries perceive the expected discounted value of future cooperation now to be higher (as (21) states). It follows, therefore, that the most-cooperative stationary tariff, $\hat{\tau}_s^c(R_0)$, is easily supported in the transition phase of the customs-union game. In fact, the balance between the incentive to cheat and the expected discounted value of future cooperation is not restored until a lower phase-2 cooperative tariff is selected and the incentive to cheat is correspondingly raised to a level commensurate with the expected discounted value of future cooperation. Hence, it must be that $\hat{\tau}_2^c < \hat{\tau}_s^c(R_0)$. The final step now is to recall that $\hat{\tau}_s^c(R_0) < \hat{\tau}_s^c(R_1) = \hat{\tau}_3^c$, and so phase-3 most-cooperative tariff must exceed the phase-2 most-cooperative tariff.

More generally, the honeymoon prediction may be understood as a reflection of the evolution of market power throughout the customs-union game. While customs unions are being negotiated and phased in, countries recognize that once these larger regions are in place, the world will have better enforcers, because under the market power effect larger regions can credibly impose higher Nash tariffs, should such punitive tariffs be called for. The prospect that a trade war initiated today might reach such proportions in the future then makes countries reluctant to pursue unilateral objectives in the present, and so the recognition of eventual customs union expansion gives rise to a honeymoon period in which low multilateral tariffs can be supported. This honeymoon eventually gives way, however, since once the customs-unions are actually expanded, each country realizes that its region's enhanced market power also makes possible a large welfare gain from defection to a higher import tariff. Thus, after the customs union expansion is finalized, multilateral tariffs must rise to diminish the incentive to cheat and restore balance.

We turn next to the initial-phase tariff, $\hat{\tau}_1^c$, which is the lowest tariff such that $\Omega(R_0,\tau_1^c) \leq V_1(\tau_1^c;\rho,\lambda,\delta,R_0,R_1)$. To characterize this tariff, we first show that:

Lemma 4: $\omega(R_1,\hat{\tau}_s^c(R_1)) > \omega(R_0,\hat{\tau}_2^c)$.

Recalling that $\hat{\tau}_s^c(R_1) = \hat{\tau}_3^c$, we see that this lemma states that the per-period equilibrium value of cooperation rises from phase 2 to phase 3. Once again, there are two effects, as phase 3 involves a smaller number of regions, which acts to raise the per-period value of cooperation, and yet phase 3 also entails a higher most-cooperative tariff (by Lemma 3), which works to reduce the per-period value of cooperation in phase 3. As above, however, the direct effect of a smaller number of regions dominates, and thus the per-period equilibrium value of cooperation is higher in phase 3. This lemma is proved in the Appendix.

Three key properties of the V_1 function may now be reported:

$$\frac{dV_1}{d\tau_1^c}(\tau_1^c;\rho,\lambda,\delta,R_0,R_1)<0 \tag{23}$$

(24) $V_1(\hat{\tau}_2^c;\rho,\lambda,\delta,R_0,R_1) < V_2(\hat{\tau}_2^c;\lambda,\delta,R_0,R_1)$

(25) $V_1(\hat{\tau}_s^c(R_0);\rho,\lambda,\delta,R_0,R_1) > \delta/(1-\delta)\ \omega(R_0,\hat{\tau}_s^c(R_0))$.

As before, a higher cooperative tariff reduces the fear of a trade war, at least during the associated phase, and therefore reduces the overall expected discounted value of future cooperation. More formally, (23) is direct from (19b) and (15). To understand (24), observe that, under Lemma 4, the per-period equilibrium value of cooperation is higher in phase 3 than in phase 2; consequently, the expected discounted value of future cooperation is higher in phase 2 than in phase 1 (at the relevant tariff), because the transition to phase 3 is more imminent when countries begin in phase 2.[19] Finally, to gain some insight into (25), recall that $V_1(\hat{\tau}_s^c(R_0);\cdot)$ is a weighted average of the associated per-period values of cooperation across the three phases, namely, $\omega(R_0,\hat{\tau}_s^c(R_0))$, $\omega(R_0,\hat{\tau}_2^c)$ and $\omega(R_1,\hat{\tau}_s^c(R_1))$; however, the per-period value of cooperation in the third phase exceeds that in the first by Lemma 2, and the per-period value of cooperation in the second phase also exceeds that in the first, since the most-cooperative second-phase tariff is lower than the most-cooperative stationary tariff that is specified in (25) for the first phase. Thus, (25) clearly must hold.[20]

With this groundwork done, we may return to Figure 3 and, using (23), (24) and (25), conclude that:

Lemma 5: $\hat{\tau}_2^c < \hat{\tau}_1^c < \hat{\tau}_3^c$.

Thus the initial-phase tariff is lower than the final-phase tariff, but it is not as low as the tariff that occurs during the transition phase.

The intuition underlying Lemma 5 is easily related. Consider first why tariffs are lower in the initial phase than in the final phase. When the countries are in phase 1, there are

R_0 regions, and each of these small regions has some incentive to cheat. Balancing against this defection incentive is the cost of a future trade war, and if the world were stationary with R_0 regions of each country type forever, then the incentive to cheat would be just balanced against the cost of a trade war at the most-cooperative stationary tariff, $\hat{\tau}_s^c(R_0)$. In phase 1 of the customs-union game, however, the countries recognize that (i) the very cooperative honeymoon phase will arrive soon, and (ii) eventually customs-union expansion will occur and there will be only R_1 regions of each country type. For both of these reasons, the expected discounted value of future cooperation (i.e., the expected cost of a trade war) is quite high as viewed from phase 1 (as (25) indicates), and so the most-cooperative stationary tariff with R_0 regions is easily supported in phase 1 of the customs-union game. In fact, an even lower cooperative tariff (with the concomitant larger defection incentive) can be supported at this phase, and it therefore follows that $\hat{\tau}_1^c < \hat{\tau}_s^c(R_0) < \hat{\tau}_s^c(R_1) = \hat{\tau}_3^c$.

At a more general level, before countries begin negotiations on customs union expansion, each region has little market power, since each region is comprised of only a few countries. Given this, there is little benefit to the countries in a region from cheating and selecting a high tariff. On the other hand, the countries correctly perceive that larger regions with great market power are on the horizon; thus, a trade war begun today would sacrifice the very cooperative tariffs that would otherwise be enjoyed while the customs union expansion was being negotiated and phased in and it would also culminate in a very costly tariff war once the large regions were actually in place. This imbalance between a low current gain from cheating and a large future cost to a trade war enables the countries to support a very low cooperative tariff in the period of time that precedes customs-union negotiations. Once the customs unions are fully implemented, however, the natural balance between the incentive to cheat and the expected discounted value of future cooperation is restored, since each region then has substantial market power in the present, with the corresponding greater incentive to cheat. Thus, once customs unions are fully implemented (in phase 3), the cooperative tariff must rise above that found prior to customs-union negotiations (in phase 1).

Finally, consider why the phase-2 most-cooperative tariff is lower than the phase-1 most-cooperative tariff. In both phases, the incentive to cheat is small, since a region has little market power, being comprised of only R_0 countries. The essential difference between the initial and transition phases is that actual customs union expansion can be expected to occur sooner when countries are already in the negotiation or transition phase. As (24) indicates, this in turn means the expected discounted value to future cooperation is higher in phase 2, since a trade war initiated in that phase will be exacerbated more quickly by the higher Nash tariffs that large regions are prone to select. Accordingly, a lower cooperative tariff can be supported in phase 2.

Our main results may now be summarized in the following proposition:

Proposition 2:

For the customs-union game, in the most-cooperative equilibrium, the domestic and foreign regions set the most-cooperative import tariffs, $\hat{\tau}_1^c$, $\hat{\tau}_2^c$ and $\hat{\tau}_3^c$, in phases 1, 2, and 3, respectively, and these tariffs may be ranked as follows: $0 < \hat{\tau}_2^c < \hat{\tau}_1^c < \hat{\tau}_3^c$.

These rankings are captured in Figure 4, which depicts the most-cooperative tariffs in the three phases of the customs-union game. As the figure illustrates, the prospect of a future customs union expansion serves to lower current cooperative tariffs, and particularly so as the expansion becomes more imminent. Once the customs union expansion is fully implemented, however, multilateral tariffs must rise to maintain cooperation.

IV. Comparative Statics

The customs-union model developed above has a variety of parameters, and it is important to assess the sensitivity of the most-cooperative tariffs to these parameters. In addition, some of these parameters may be loosely associated with aspects of GATT policy toward customs unions as embodied in Article XXIV. Thus, in this section, we present comparative-statics results, and in the concluding section we discuss their implications with regard to the design of Article XXIV. Examining the respective no-defect conditions presented above, it is apparent that the most-cooperative tariffs have the following functional dependencies: $\hat{\tau}_3^c = \hat{\tau}_3^c(\delta, R_1)$, $\hat{\tau}_2^c = \hat{\tau}_2^c(\lambda, \delta, R_0, R_1)$ and $\hat{\tau}_1^c = \hat{\tau}_1^c(\rho, \lambda, \delta, R_0, R_1)$.

Before proceeding, it is important to record the following corollary:

Corollary 1: $\omega(R_1, \hat{\tau}_3^c) > \omega(R_0, \hat{\tau}_2^c) > \omega(R_0, \hat{\tau}_1^c)$.

This corollary indicates that the per-period equilibrium value of cooperation increases through time. The proof is direct, as the first inequality is simply a restatement of Lemma 4 (recall that $\hat{\tau}_3^c = \hat{\tau}_s^c(R_1)$), while the second inequality follows from Proposition 2 and (15). This corollary will be important for the proofs of the comparative-statics results derived below.

Our central set of results is contained in the following proposition:

Proposition 3:

For the customs-union game, the most-cooperative tariffs satisfy the following relationships:

(i) $\hat{\tau}_1^c(\rho,\lambda,\delta,R_0,R_1)$ is decreasing in ρ, λ and δ, and it is increasing in R_1.

(ii) $\hat{\tau}_2^c(\lambda,\delta,R_0,R_1)$ is decreasing in λ and δ, and it is increasing in R_1.

(iii) $\hat{\tau}_3^c(\delta,R_1)$ is decreasing in δ, and decreasing in R_1 if δ is sufficiently small.

The proof of this proposition is in the Appendix.

To gain some intuition, let us start with the effect of an increase in ρ on the phase-1 most-cooperative tariff. As ρ rises, countries that are currently in phase 1 recognize that the transition to phase 2 and, ultimately, the transition to phase 3 will occur sooner. This in turn raises the expected discounted value to cooperation (i.e., V_1) as viewed from phase 1, since Corollary 1 establishes that the per-period equilibrium value of cooperation is larger in later phases. With the perceived cost to a trade war thereby increased, lower phase-1 tariffs (with the associated higher incentive to cheat) can be supported.

The consequences of a larger value for λ are similar, although the argument is slightly more involved. Consider first the phase-2 most-cooperative tariff. As λ rises, the transition to the final phase is expedited, and, using Corollary 1 once more, it follows that the expected discounted value of future cooperation (i.e., V_2) as viewed from phase 2 rises. Thus, a higher value for λ acts to lower the phase-2 most-cooperative tariff. In phase 1, a higher λ also results in a lower most-cooperative tariff, though now for two reasons. First, as before, when λ is increased, the final phase is reached more quickly, and using Corollary 1 this enables a lower phase-1 most-cooperative tariff. Second, an increase in λ lowers the phase-2 most-cooperative tariff, as argued just above, and the anticipation of this more-cooperative phase-2 behavior in turn acts to raise the expected discounted value of future cooperation as viewed from phase 1, thereby making possible a lower phase-1 most-cooperative tariff. Thus, for both direct and indirect reasons, a higher value for λ results in a lower phase-1 most-cooperative tariff.

An increase in δ has the anticipated effect on the most-cooperative tariffs, as all such tariffs then decline. Intuitively, the direct effect of an increase in δ is that the future is valued more, and so countries are more reluctant to sacrifice cooperation and enter into a trade war. Thus, a higher value for δ serves to raise the expected discounted value to future cooperation, and thereby makes possible the support of lower most-cooperative tariffs in all phases. Second, for phases 1 and 2, an increase in δ also has beneficial indirect effects. For example, in phase 2, when δ increases, countries recognize that the phase-3 most-cooperative tariff will be lower as a result, and so the cost of a future trade war is raised for this reason as well. Similar indirect effects arise in phase 1.

Finally, we may view R_0 as an initial condition and investigate the consequences of greater customs union expansion by allowing R_1 to be smaller. For sufficiently small δ, a decrease in R_1 raises the phase-3 most-cooperative tariff. Intuitively, as Lemma 1 indicates, the phase-3 most-cooperative tariff is simply the most-cooperative stationary tariff for a world with R_1 regions of each country type. Further, as we argued in Section II, in stationary environments a reduction in the number of regions enhances each region's market power, and this results in a higher most-cooperative stationary tariff if countries discount the future sufficiently. Thus, given our small-δ orientation, it must be that the phase-3 most-cooperative tariff rises as customs union expansion becomes more significant.[21]

The effect of greater customs union expansion generates a rather different consequence for the phase-1 and phase-2 most-cooperative tariffs. As we show in the Appendix while proving Lemma 2, a key feature of our model is that $\omega(R, \hat{\tau}_s^c(R)) \equiv f(R)$ is declining in R, so that the per-period value of cooperation at the most-cooperative stationary tariff is always higher when there are fewer regions. In other words, the direct positive effect that a reduction in the number of regions has for the per-period value of cooperation always outweighs any possible negative indirect effect from a consequent increase in the most-cooperative stationary tariff.[22] Using Lemma 1, it follows immediately that a reduction in R_1 acts to raise the per-period equilibrium value of cooperation in phase 3, and this in turn implies that a lower phase-2 most-cooperative tariff can be supported when greater customs union expansion is anticipated. Given that greater customs union expansion has the direct effect of raising the per-period equilibrium value of cooperation in phase 3 and the indirect effect of lowering the phase-2 most cooperative tariff, it follows that the expected discounted value of future cooperation (i.e., V_1) as viewed from phase 1 rises for two reasons, and a lower phase-1 most-cooperative tariff can be supported as a result.

Thus, in general, customs union expansion has a beneficial effect on multilateral tariff cooperation before the expansion is fully implemented. Further, any parameter change that speeds up the transitional process that leads to customs union expansion, or which increases the extent of the eventual expansion, will result in even greater multilateral tariff cooperation prior to the full implementation of the regional agreements. On the down side, however, if countries are sufficiently impatient, greater regional expansion can have negative consequences for multilateral tariff cooperation once the expansion is fully implemented.

V. Conclusion

We have presented a model of customs unions which predicts that the early stages of the process of customs union formation will lead to a temporary "honeymoon" for liberal multilateral trade policies which ultimately must be reversed as the customs union becomes fully implemented. We have highlighted the special effects of customs union formation as distinct from the formation of free trade areas by constructing a model that isolates the market power effect which comes with customs union formation, and abstracts from the trade diversion effect which is common to both customs unions and free trade agreements. Since a comparison of our results here to those of our companion paper (Bagwell and Staiger, 1993), which considers free trade agreements and thus the trade diversion associated with their formation, establishes that trade diversion effects of regional agreements run opposite to market power effects in terms of their implications for multilateral tariff cooperation, we can only claim to have captured the implications of customs union formation for multilateral tariff cooperation when the market power effect dominates the trade diversion effect.[23] Nevertheless, these two papers together highlight the two distinct forces that determine the impacts of regional integration on multilateral tariff cooperation, with free trade agreements reflecting primarily the trade diversion effects and customs unions in general reflecting some combination of both trade diversion and market power effects.

Moreover, our results can provide an interpretation of the apparent consequences of the original EC customs union formation and subsequent enlargement for multilateral liberalization: Provided that market power effects of EC customs union formation were sufficiently important relative to trade diversion effects, our model predicts that multilateral tariff liberalization would be initially stimulated by the formation and subsequent enlargement of the EC customs union, but that this complementarity would ultimately give way to a more dissonant relationship and multilateral tariff cooperation would suffer as a result.

In fact, in light of the prominence of the EC experience, it may be useful to depart from the symmetric customs union formation that we have considered throughout the formal analysis and to illustrate the workings of our model for a specific example in which a single group of countries forms together into a customs union. Figure 5 depicts such a case with specific parameter values (repeated in the legend) chosen to yield the stylized tariff movements that emerge generally under the symmetric customs union formation on which our results in Proposition 2 are based.[24] As the figure depicts, with one domestic region and two foreign regions initially, the domestic region initially enjoys an asymmetrically high cooperative tariff and faces correspondingly low cooperative tariffs from its two foreign trading partners, reflecting the larger bargaining power that its size initially imparts. Upon the initiation of the transition period during which the two foreign regions will ultimately form a single foreign customs union, multilateral liberalization occurs, with all countries reducing their cooperative tariffs. For the domestic region, this newfound ability to liberalize multilaterally reflects the increased likelihood that a breakdown in multilateral cooperation would imply a trade war with a united trading bloc abroad. For the foreign regions, it reflects the increased stakes of maintaining multilateral cooperation in the presence of a now more cooperative domestic region. Finally, once the foreign customs union comes into force and the two foreign regions begin behaving as a unified bloc, symmetric in size to the domestic region, the liberal multilateral tariffs of the honeymoon period cannot be sustained and both blocs raise their multilateral tariffs. For the foreign bloc, this deterioration in the ability to maintain low multilateral tariffs stems from its rising temptation to defect and utilize its increased market power. For the domestic bloc, it reflects the decreased stakes of maintaining multilateral cooperation in the presence of the now less cooperative foreign bloc.

While this example is clearly stylized in its depiction of the evolution of multilateral cooperation, it does conform with the basic historical trends in liberalization under GATT as described, for example, by Bhagwati (1988). Bhagwati documents the rise in protectionism which began in the mid 1970's reversing its previous decline, and attributes this reversal to various elements of structural change including the rise of the newly industrialized countries, the relative decline of the United States, and the emergence of macroeconomic imblances. Our theory suggests adding another element of structural change to the story: the formation and the process of implementation of the EC customs union. Of course, our theory does not explicitly address the change in emphasis from tariff to non-tariff barriers which also occured at this time.[25] Nor does it capture the political economy dimensions on which Bhagwati focuses and which are undoubtedly important.[26] But it does highlight several of the effects of customs union formation on multilateral cooperation which are likely to play an important role in any more general theory.

Finally, we comment briefly on the institutional implications of our analysis with regard to the design of Article XXIV. In our companion paper (Bagwell and Staiger, 1993), we argued that the proscriptions placed on regional agreements by Article XXIV could be interpreted as an attempt to restrict the behavior of GATT member countries with regard to regional agreements in an effort to reduce (a) the frequency with which such agreements occur (related to a reduction in the parameter ρ), (b) the degree of trade diversion which accompanies such agreements, and (c) the length of the transition period to the fully implemented agreement (related to a rise in the parameter λ) from what these parameters would look like in an unrestricted world.[27] The comparative statics results here and in our companion paper suggest that Article XXIV could have important consequences for multilateral tariff cooperation. In particular, efforts to lower the frequency with which customs union agreements are negotiated will tend to diminish multilateral cooperation as long as those efforts are successful and customs union formation is deterred, and will delay the start of the harmonious transition phase associated with the early stages of customs union implementation, but will also postpone the post-customs-union high-tariff world, i.e., $\hat{\tau}_1^c$ will be higher than it would have been, and $\hat{\tau}_2^c$ (which is below $\hat{\tau}_1^c$) will be delayed, but $\hat{\tau}_3^c$ (which is the highest of the three) is also postponed. At the same time, efforts to shorten the transition period during which a customs union agreement is implemented will tend to boost multilateral cooperation up until the final implementation is achieved, but such efforts will also hasten the arrival of the post-customs-union high-tariff world, i.e., $\hat{\tau}_1^c$ and $\hat{\tau}_2^c$ will fall, but $\hat{\tau}_3^c$ (which is the highest of the three) will arrive sooner. Finally, efforts to reduce trade diversion will increase the relative importance of market power effects associated with customs unions, and accentuate the distinction between customs unions and free trade agreements that we have drawn out in this and our companion paper. The distinct implications of Article XXIV proscriptions for multilateral tariff cooperation that emerge from a comparison across our two papers suggests a possible rationale for very different treatment of these two forms of regional agreements within Article XXIV itself.

References

Arndt, Sven W., "Customs Union and the Theory of Tariffs," American Economic Review 59, 1969, pp. 108-118.

Bagwell, Kyle, and Robert Staiger, "A Theory of Managed Trade," American Economic Review 80, September 1990, pp. 779-795.

Bagwell, Kyle, and Robert Staiger, "Multilateral Cooperation During the Formation of Free Trade Areas," NBER Working Paper No. 4364, 1993.

Bhagwati, Jagdish, Protectionism, Cambridge, MA: MIT Press, 1988.

Bhagwati, Jagdish, The World Trading System at Risk, Princeton University Press, Princeton, New Jersey, 1991.

Bond, Eric W., and Costas Syropoulos, "Optimality and Stability of Regional Trading Blocs," unpublished manuscript, November 1992.

Bond, Eric and Costas Syropoulos, "Trading Blocs and the Sustainability of Inter-Regional Cooperation," unpublished manuscript, March 1993.

Deardorff, Alan, and Robert Stern, "Multilateral Trade Negotiations and Preferential Trading Arrangements," in Alan Deardorff and Robert Stern (eds.) Analytical and Negotiating Issues in the Global Trading System, Ann Arbor, University of Michigan Press, 1987, 1991 and forthcoming.

Grossman, Gene, and Elhanan Helpman, "Trade Wars and Trade Talks," NBER Working Paper No. 4280 (1993).

Herter, Christian, Statement Before The Joint Economic committee of the Congress of the United States, December 4-14, 1961, pp. 8-13.

Kowalczyk, Carsten, and Tomas Sjostrom, "Bringing GATT into the Core," unpublished manuscript, Dartmouth College, November 1992.

Kowalczyk, Carsten, and Ronald Wonnacott, "Substitute and Complement Trading Clubs," Dartmouth College, Dept. of Economics, WP No. 91-16, 1991.

Krause, Lawrence B., European Economic Integration and the United States, Brookings Institute, 1968.

Kreuger, Anne O., "Free Trade Agreements as Protectionist Devices: Rules of Origin," NBER Working Paper No. 4352, April 1993.

Krugman, Paul R., "Is Bilateralism Bad?" in Elhanan Helpman and Assaf Razin (eds.) International Trade and Trade Policy, The MIT Press (Cambridge, MA and London, England), 1991a, pp. 9-23.

Levy, Philip I., "A Political-Economic Analysis of Free Trade Agreements," unpublished manuscript, Stanford University, January 1993.

Ludema, Rodney, "On the Value of Preferential Trade Agreements in Multilateral Negotiations," unpublished manuscript, May 1992.

Viner, Jacob, The Customs Union Issue, New York: Carnegie Endowment for International Peace, 1950.

Figure 1

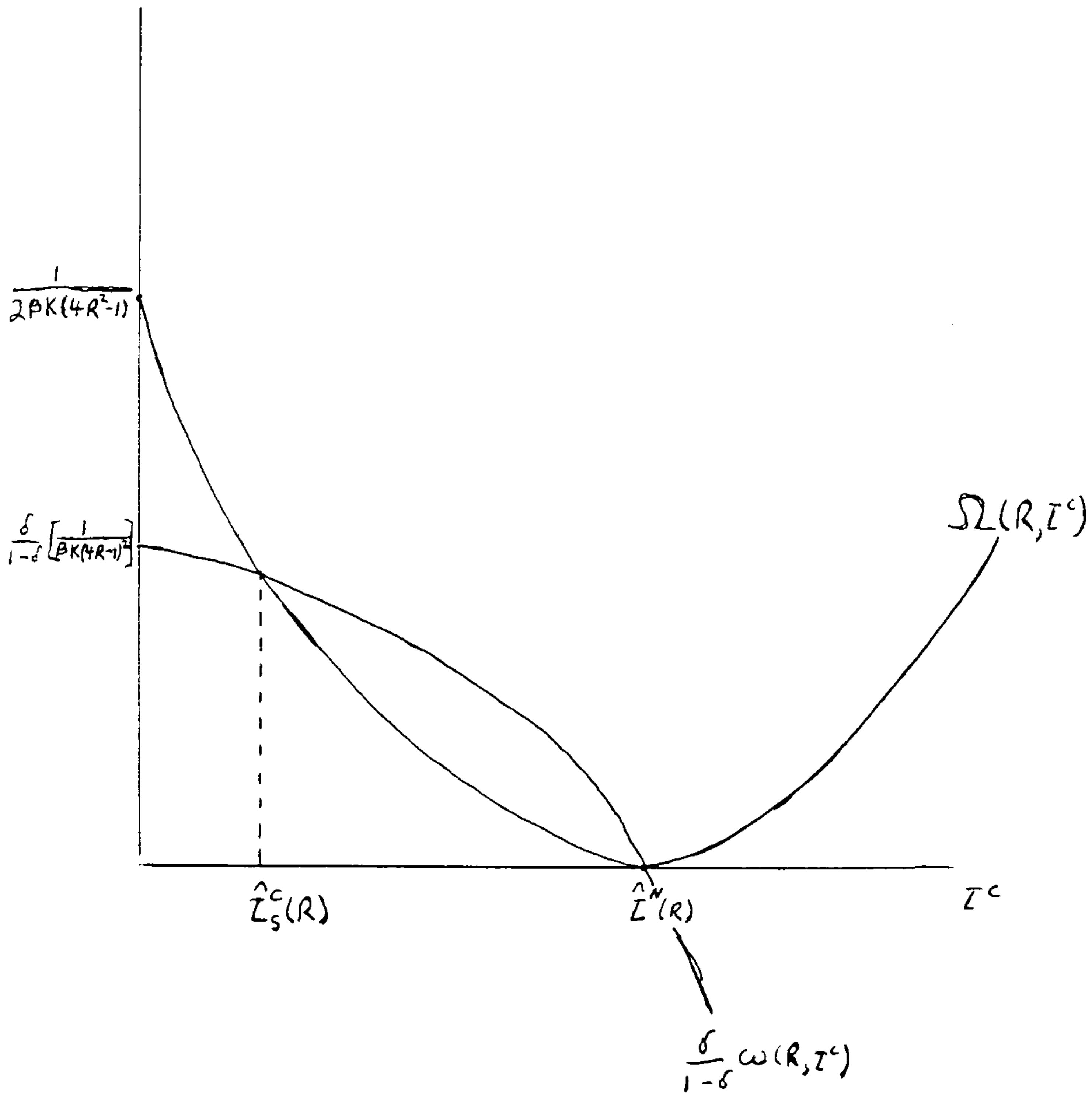
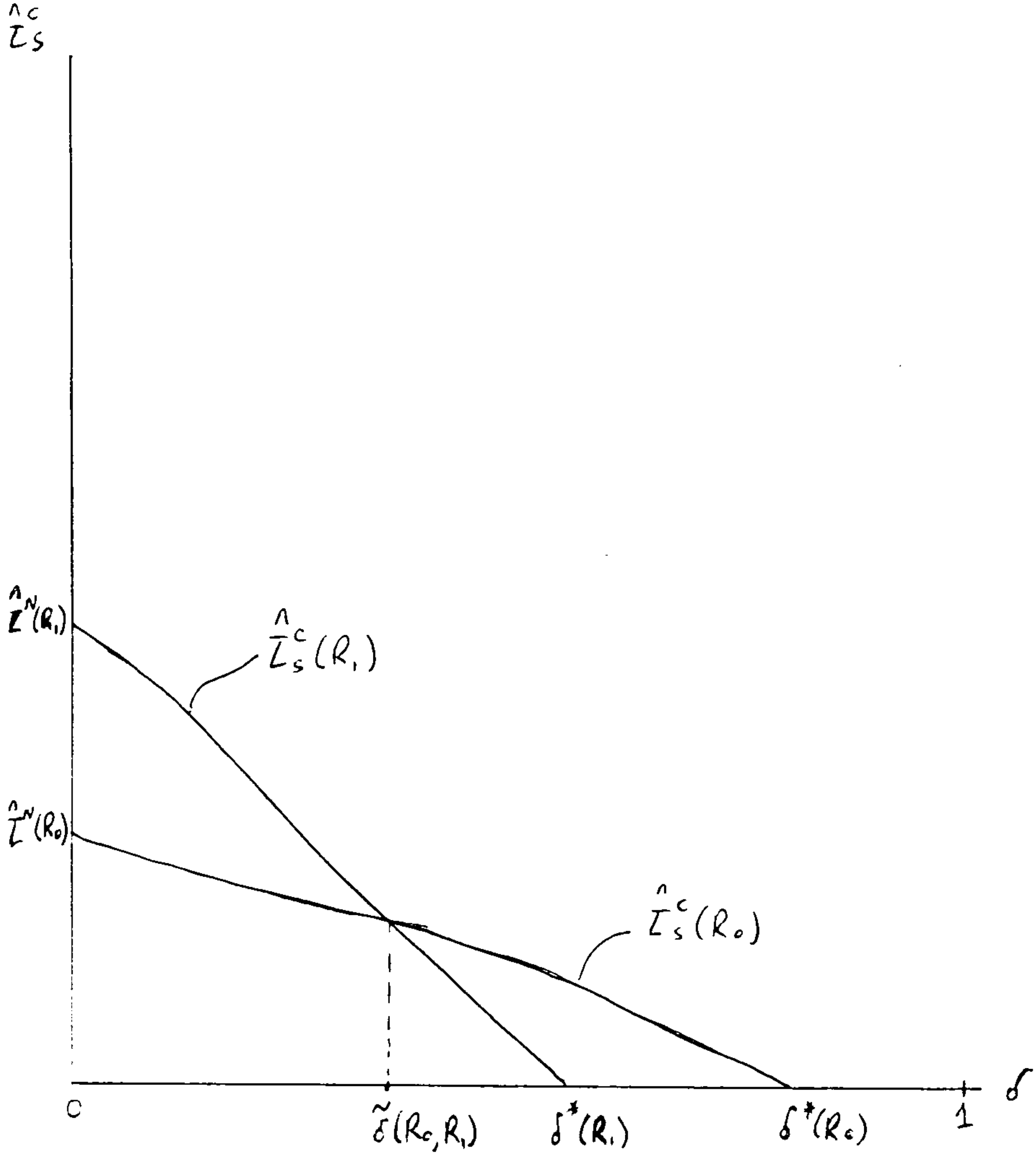
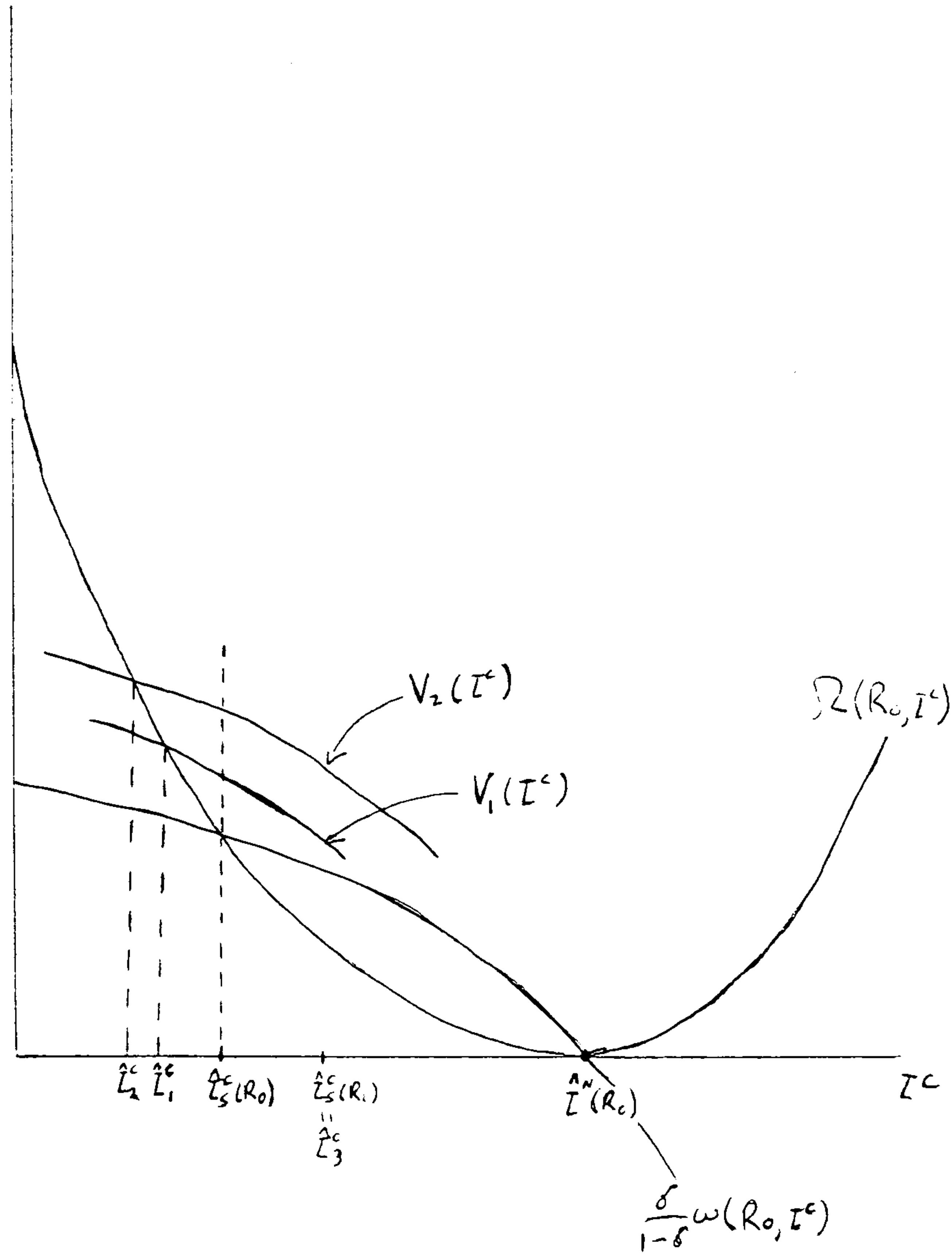

$\frac{1}{2\beta K(4R^2-1)}$
$\frac{\delta}{1-\delta}\left[\frac{1}{\beta K(4R-1)^2}\right]$
$\Omega(R,I^c)$
$\hat{I}^c_s(R)$
$\hat{I}^N(R)$
I^c
$\frac{\delta}{1-\delta}\omega(R,I^c)$

Figure 2



$\hat{L}_s^c$
$\hat{L}^N(R_1)$
$\hat{L}^N(R_0)$
$\hat{L}_s^c(R_1)$
$\hat{L}_s^c(R_0)$
0
$\tilde{\delta}(R_0,R_1)$
$\delta^*(R_1)$
$\delta^*(R_0)$
1
δ

Figure 3



$V_2(I^c)$
$V_1(I^c)$
$\Omega(R_0, I^c)$
$\hat{I}_2^c$
$\hat{I}_1^c$
$\hat{I}_S^c(R_0)$
$\hat{I}_S^c(R_1)$
$=$
$\hat{I}_3^c$
$\hat{I}^N(R_0)$
I^c
$\frac{\delta}{1-\delta}\omega(R_0, I^c)$

Figure 4

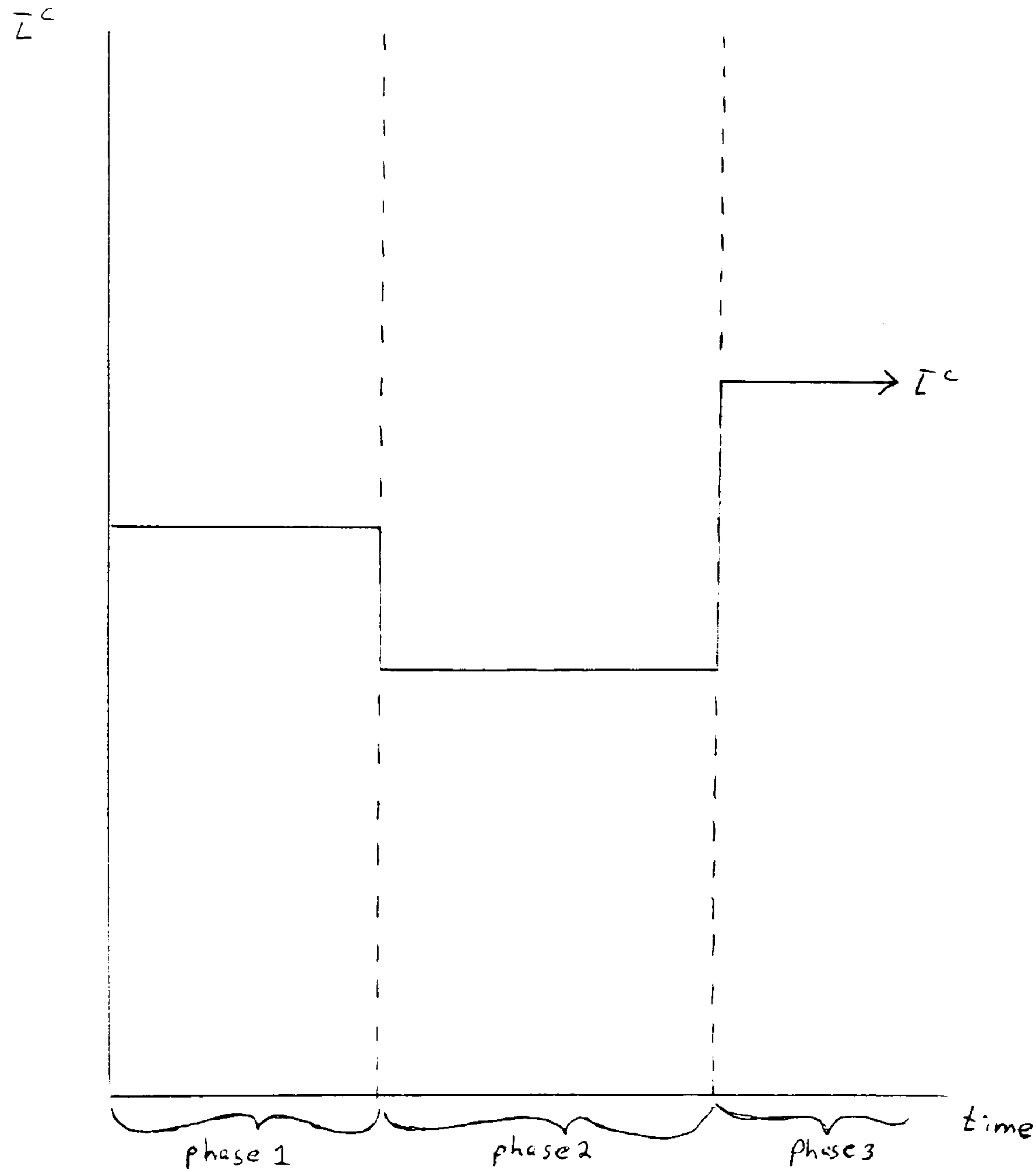
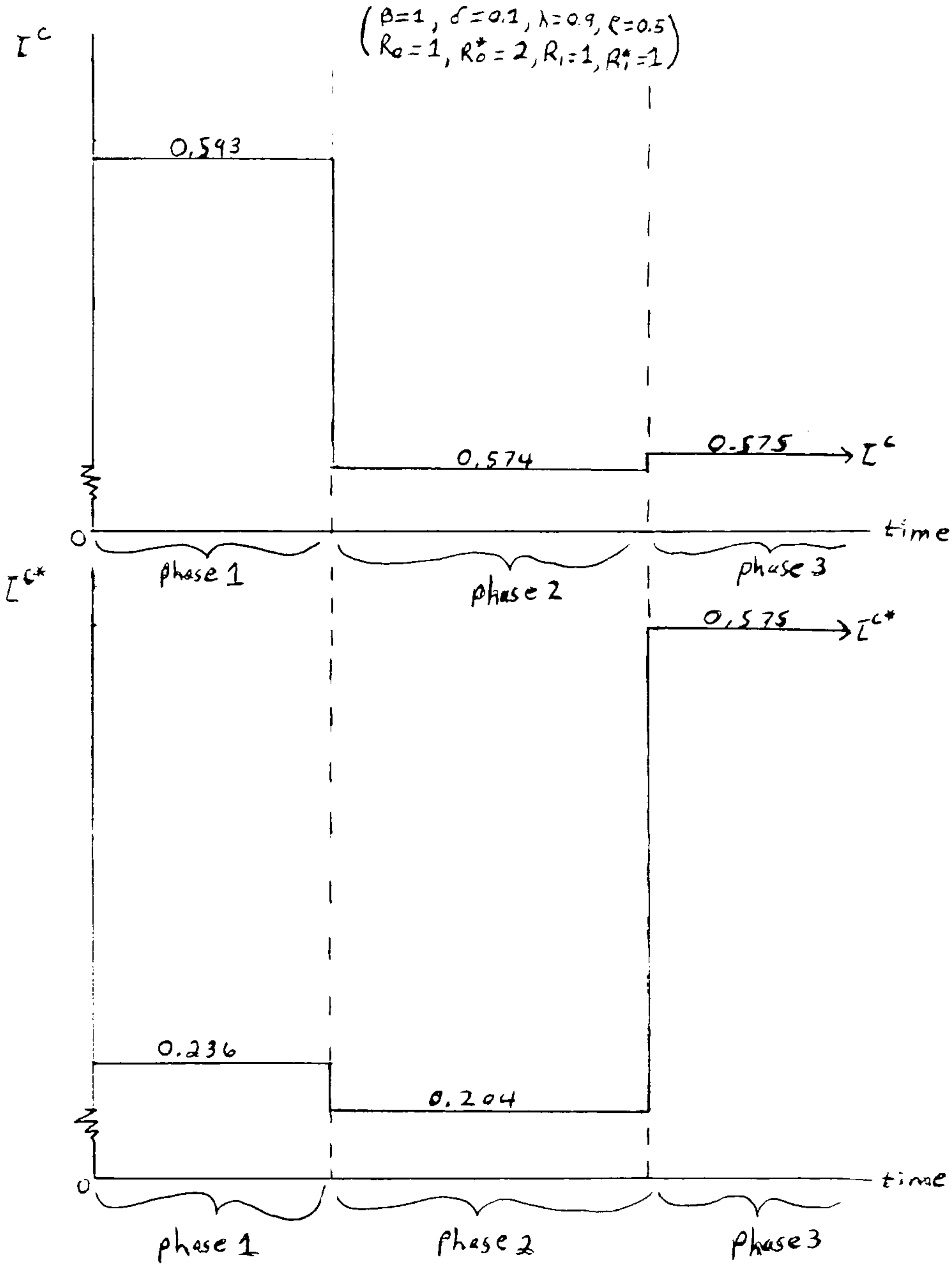


Figure 5



(β=1, δ=0.1, λ=0.9, ρ=0.5
$R_0=1$, $R_0^*=2$, $R_1=1$, $R_1^*=1$)
I^c
0.593
0.574
0.575
I^c
0
time
phase 1
phase 2
phase 3
I^{c*}
0.575
I^{c*}
0.236
0.204
0
time
phase 1
phase 2
phase 3

Appendix

Proof of Lemma 2: $\omega(R_1, \hat{\tau}_s^c(R_1)) > \omega(R_0, \hat{\tau}_s^c(R_0))$

Observe that

$$\text{(A1).} \quad \frac{d}{dR}\,\omega(R, \hat{\tau}_s^c(R)) = \frac{\partial\omega\,(R, \hat{\tau}_s^c(R))}{\partial R} + \frac{\partial\omega}{\partial\tau^c}\,(R, \hat{\tau}_s^c(R))\,\frac{\partial\hat{\tau}_s^c(R)}{\partial R}$$

To sign this expression, we first calculate that

$$\text{(A2).} \quad \omega(R, \hat{\tau}_s^c(R)) = \frac{8\delta(1-\delta)(4R^2-1)}{K\beta[(4R-1)^2(1-\delta)+2\delta(4R^2-1)]^2}$$

Further calculations then yield that

$$\text{(A3).} \quad \frac{d}{dR}\,\omega(R, \hat{\tau}_s^c(R)) = \frac{64\delta(1-\delta)\{(8R^3-7R+2)\delta-(16R^3-9R+2)\}}{K\beta\,[(4R-1)^2(1-\delta)+2\delta(4R^2-1)]^3} < 0$$

where the inequality follows since the bracketed expression in the numerator is negative for all $\delta\ \varepsilon(0,1)$ and $R \geq 1$.

Proof of Lemma 4: $\omega(R_1, \hat{\tau}_s^c(R_1)) > \omega(R_0, \hat{\tau}_2^c)$

Define $D(\tau_2^c) \equiv \omega(R_1, \hat{\tau}_s^c(R_1)) - \omega(R_0, \tau_2^c)$. Observe that $D(\tau_2^c)$ is increasing for $\tau_2^c > 0$; thus, since $\hat{\tau}_2^c > 0$, Lemma 4 is sure to hold if $D(0) = 0$. Assume then that $D(0) < 0$. In this event, using Lemma 2, there exists a unique $\tau^*\ \varepsilon(0, \hat{\tau}_s^c(R_0))$ at which $D(\tau^*) = 0$. The lemma is thus proved if $\hat{\tau}_2^c > \tau^*$.

Let $\hat{\tau}_2^c(\lambda=1)$ denote the most-cooperative phase-2 tariff when $\lambda=1$. Since under (22) the no-defect condition (18b) must hold with equality, we may use $D(\tau^*) \equiv \omega(R_1,\hat{\tau}_s^c(R_1)) - \omega(R_0,\tau^*) = 0$ and Lemma 1 to conclude that

$$\text{(A4).} \quad \Omega(R_0, \hat{\tau}_2^c(\lambda=1)) = \frac{\delta}{1-\delta}\,\omega(R_0, \tau^*)$$

But $\tau^*\ \varepsilon(0,\hat{\tau}_2^c(R_0))$ then implies that $\hat{\tau}_2^c(\lambda=1)\ \varepsilon\ (\tau^*,\hat{\tau}_s^c(R_0))$ (See Figure 1.). Thus, the lemma holds for $\lambda=1$ and $D(\hat{\tau}_2^c(\lambda=1)) > 0$.

Examining (18b) and (19), it is apparent that $\hat{\tau}_2^c(\lambda=0) = \hat{\tau}_s^c(R_0) > \tau^*$. Thus,

$D(\hat{\tau}_2^c(\lambda=0)) > 0$ is also true, and the lemma holds when $\lambda=0$.

We next compute $\frac{\partial \hat{\tau}_2^c}{\partial \lambda}$. Since (18b) holds with equality, this is given by

$$\text{(A5).} \quad \frac{d\hat{\tau}_2^c}{\partial \lambda} = \frac{\dfrac{\partial V_2(\hat{\tau}_2^c; \lambda,\delta,R_0,R_1)}{\partial \lambda}}{\dfrac{\partial \Omega(R_0,\hat{\tau}_2^c)}{\partial \tau_2^c} - \dfrac{\partial V_2(\hat{\tau}_2^c; \lambda,\delta,R_0,R_1)}{\partial \tau_2^c}}$$

The denominator of (A5) is negative, since V_2 cuts ω from below at $\hat{\tau}_2^c$, as Figure 3 illustrates. It follows that

$$\text{(A6). sign } \frac{\partial \hat{\tau}_2^c}{\partial \lambda} = -\text{sign } \frac{\partial V_2(\hat{\tau}_2^c; \lambda,\delta,R_0,R_1)}{\partial \lambda}$$

Using (18b) and Lemma 1, we find that

$$\text{(A7).} \quad \frac{\partial V_2(\hat{\tau}_2^c; \lambda,\delta,R_0,R_1)}{\partial \lambda} = \frac{\delta}{(1-(1-\lambda)\delta)^2}[\omega(R_1,\hat{\tau}_s^c(R_1)) - \omega(R_0,\hat{\tau}_2^c)],$$

from which it follows that

$$\text{(A8).} \quad \text{sign } \frac{\partial \hat{\tau}_2^c}{\partial \lambda} = -\text{sign } D(\hat{\tau}_2^c)$$

Now suppose that $\lambda^{\dagger}\varepsilon(0,1)$ exists for which $D(\hat{\tau}_2^c(\lambda^{\dagger})) = 0$. From (A8), it follows that $\hat{\tau}_2^c(\lambda) = \tau^*$ for all $\lambda \geq \lambda^{\dagger}$. But this contradicts $\hat{\tau}_2^c(\lambda=1) > \tau^*$. It thus must be that $D(\hat{\tau}_2^c(\lambda)) > 0$ for all $\lambda\, \varepsilon[0,1]$, which proves Lemma 4. Note also from (A8) that $\hat{\tau}_2^c$ decreases in λ.

<u>Proof of Proposition 3</u>: Comparative Statics.

Begin with part (iii). Using Lemma 1 and (16), straightforward calculations reveal that $\hat{\tau}_3^c$ decreases in δ for $\delta\, \varepsilon(0,\delta^*(R_1))$. Further $\frac{\partial \hat{\tau}_3^c}{\partial R_1}$ has sign which is (i). quadratic in δ, (ii). zero at $\tilde{\delta}(R_1,R_1\text{-}\varepsilon)$ for some small $\varepsilon > 0$, (iii). negative at $\delta = 0$, and (iv).

positive at $\delta = 1$. It follows that $\hat{\tau}_3^c$ decreases in R_1 for $\delta\ \varepsilon(0,\tilde{\delta}(R_1,R_1\text{-}\varepsilon))$.

Consider next part (ii). The proof of Lemma 4 establishes that $\hat{\tau}_2^c$ decreases in λ. Arguing as in that proof, it is apparent that $\hat{\tau}_2^c$ decreases in δ, if V_2 increases in δ when τ_2^c is fixed at $\hat{\tau}_2^c$. Calculations reveal that

$$\frac{\partial V_2(\hat{\tau}_2^c;\lambda,\delta,R_0,R_1)}{\partial \delta} = \frac{\lambda\delta/(1-\delta)}{1-(1-\lambda)\delta}\ \frac{\partial\omega(R_1,\hat{\tau}_3^c)}{\partial\tau_3^c}\frac{\partial\hat{\tau}_3^c}{\partial\delta} + \frac{\lambda[1-\delta^2(1-\lambda)]}{(1-\delta)^2(1-(1-\lambda)\delta)^2}\,\omega(R_1,\hat{\tau}_3^c)$$

$$+\frac{(1-\lambda)}{(1-(1-\lambda)\delta)^2}\,\omega(R_0,\hat{\tau}_2^c) > 0,$$

Thus, $\hat{\tau}_2^c$ declines in δ. Similarly, $\hat{\tau}_2^c$ is increasing in R_1, since using Lemma 1 and (A3),

$$\frac{\partial V_2(\hat{\tau}_2^c;\lambda,\delta,R_0,R_1)}{\partial R_1} = \frac{\lambda\delta/(1-\delta)}{1-(1-\lambda)\delta}\ \frac{d}{dR_1}\ \omega(R_1,\hat{\tau}_s^c(R_1)) < 0$$

Consider finally part (i). Using Corollary 1,

$$\frac{\partial V_1(\hat{\tau}_1^c;\rho,\lambda,\delta,R_0,R_1)}{\partial\rho} = \frac{\delta}{(1-(1-\rho)\delta)^2}\left\{\frac{1-\delta}{1-(1-\lambda)\delta}\left[\omega(R_0,\hat{\tau}_2^c) + \frac{\lambda\delta}{1-\delta}\,\omega(R_1,\hat{\tau}_3^c)\right] - \omega(R_0,\hat{\tau}_1^c)\right\}$$

$$> \frac{\delta}{(1-(1-\rho)\delta)^2}\left\{\frac{1-\delta}{1-(1-\lambda)\delta}\left[\omega(R_0,\hat{\tau}_1^c) + \frac{\lambda\delta}{1-\delta}\,\omega(R_0,\hat{\tau}_1^c)\right] - \omega(R_0,\hat{\tau}_1^c)\right\}$$

$$= 0,$$

and so $\hat{\tau}_1^c$ is decreasing in ρ. Next, using Corollary 1 again,

$$\frac{\partial V_1(\hat{\tau}_1^c;\rho,\lambda,\delta,R_0,R_1)}{\partial\lambda} = \frac{\rho\delta}{(1-(1-\rho)\delta)(1-(1-\lambda)\delta)^2}\left\{(1-(1-\lambda)\delta)\frac{\partial\omega(R_0,\hat{\tau}_2^c)}{\partial\tau_2^c}\frac{\partial\hat{\tau}_2^c}{\partial\lambda}\right.$$

$$\left. + \delta\left(\omega(R_1,\hat{\tau}_3^c) - \omega(R_0,\hat{\tau}_2^c)\right)\right\} > 0,$$

and so $\hat{\tau}_1^c$ decreases in λ.

To evaluate the dependence of $\hat{\tau}_1^c$ on δ, re-write V_1 as

$$V_1(\tau_1^c;\rho,\lambda,\delta R_0,R_1) =$$

$$\left[\frac{\rho\delta}{(1-(1-\rho)\delta)(1-(1-\lambda)\delta)}\right]\left[\omega(R_0,\hat{\tau}_2^c)+\frac{\lambda\delta}{1-\delta}\,\omega(R_1,\hat{\tau}_3^c)\right]+\left[\frac{(1-\rho)\delta}{1-(1-\rho)\delta}\right]\omega(R_0,\tau_1^c)$$

Calculations reveal that each bracketed term increases with δ, so that V_1 increases with δ.

Thus, $\hat{\tau}_1^c$ declines as δ increases. Finally, using Lemma 1 and (A3) we have that

$$\frac{\partial V_1(\hat{\tau}_1^c;\rho,\lambda,\delta,R_0,R_1)}{\partial R_1}=\frac{\rho\delta}{1-(1-\rho)\delta}\left[\frac{\dfrac{\partial\omega(R_0,\hat{\tau}_2^c)}{\partial\tau_2^c}\dfrac{\partial\hat{\tau}_2^c}{\partial R_1}+\dfrac{\lambda\delta}{1-\delta}\dfrac{d}{dR_1}\omega(R_1,\hat{\tau}_s^c(R_1))}{1-(1-\lambda)\delta}\right]<0,$$

It thus must be that $\hat{\tau}_1^c$ increases in R_1.

Endnotes

1. Arndt (1969) was apparently the first to attempt to formalize the effect of customs-union formation on the tariffs of non-member countries, concluding that: "An interesting by-product of the foregoing analysis is the suggestion that optimum tariff strategy may dictate that some countries remaining outside the union reduce their prevailing tariffs. Much depends upon the extent of tariff warfare prior to formation of the union, and upon the dynamics of tariff competition about which very little is known. It is nevertheless intriguing to speculate about the extent to which the existence of the European Economic Community facilitated the Kennedy Round" (p. 117).

2. See, for example, the recent analyses of Krugman (1991), Deardorff and Stern (1991), Bond and Syropoulos (1992, 1993), Ludema (1992), Kowalczyk (1990), Kowalczyk and Wonnacott (1991), Kowalczyk and Sjostrom (1992), and Levy (1993).

3. For example, the staging of internal tariff reductions and harmonization of external tariffs in the original formation of the EC customs union lasted 10 years.

4. In principle, increasing income associated with regional integration could stimulate trade to such an extent that inter-region trade could actually grow relative to what it would have been. In practice, though, a substantial drop in inter-bloc trade is likely with regional integration. For example, Krause (1968) estimated that U.S. exports of manufactures to the EC market would eventually drop by 15% as a result of the formation of the EC. Note also that our use of the term "trade diversion" here refers simply to a reduction in interbloc trade volume. This differs from the standard usage of the term as defined by Viner (1950), which in addition to changing trade patterns, emphasizes the move from a more to a less efficient allocation of resources.

5. We follow the standard convention of defining a customs union as a free trade area with a common external tariff policy. As such, the market power effect which we have described arises by definition only in the context of customs unions, and is absent from free trade areas. In practice, this distinction between customs unions and free trade areas is more a matter of degree than our simple dichotomy would suggest: the formation of a free trade area may implicitly serve to harmonize external tariffs of member countries to some degree through its choice of rules of origin. For example, absent internal transport costs within a free trade area, very weak rules of origin would tend to funnel imports across the border with the lowest external tariffs. Alternatively, as Kreuger (1993) has noted, the adoption of sufficiently strict rules of origin might allow a member country with high external tariffs on an imported intermediate product to effectively extend protection on that intermediate good to the entire free trade area, if final goods producers in other member countries were induced to switch from low-cost non-member country suppliers of the intermediate product to suppliers in the high-tariff member country in order to meet the rule of origin for the free trade area. Despite these qualifications, we find a sharp distinction between free trade areas and customs unions to be conceptually useful for sorting out the impacts of various kinds of regional agreements, with the understanding that free trade agreements in practice may contain small amounts of the market power effects that we associate with customs unions.

6. We discuss the generalization of our results to asymmetric cases in the concluding section.

7. Each country is also endowed with a traded numeraire good that provides constant marginal utility and is always consumed in positive amounts by inhabitants of every country. Trade in the numeraire good ensures overall trade balance.

8. We abstract in the formal model from the possibility of inter-regional tariff discrimination based on import source. Non-discriminatory tariffs are natural in the setting where non-member countries with which a country trades are all symmetric. However, we are ruling out discriminatory tariffs as possible punishments, an issue to which we return in the next section.

9. The finding that static Nash tariffs increase with customs union formation is a robust implication of the market power effect of customs union formation that we have isolated in this model, but it can be overturned in certain circumstances with the introduction of sufficient trade diversion. (See Krugman, 1991, for a model of customs union formation with both market power and trade diverting effects, in which customs union formation leads to higher static Nash tariffs, and Bond and Syropoulos, 1992, for an example of how trade diverting effects can overturn this result.) We return in the concluding section to discuss more broadly the sensitivity of our results to the relative strengths of the market power and trade diversion effects associated with the formation of customs unions.

10. We could consider other forms of symmetric punishment, some of which would allow for greater levels of cooperation than the infinite Nash reversion considered here. However, the qualitative nature of our dynamic results concerning the behavior of cooperative tariffs is unlikely to be affected. Moreover, infinite reversion is not an entirely implausible representation of actual tariff wars: The high U.S. tariffs on imports of light-duty trucks imposed as a result of the "chicken war" with the EC in 1963, for example, are still in place 30 years later. Our restriction to symmetric punishments, however, could be more substantive. In particular, it will be important for our results that customs union formation enhances the severity of punishments. But if punishments involved high tariffs only between a defector and its trading partners, with low tariffs continuing between all other regional trading pairs, then customs union formation could fail to facilitate punishments, and could even make them more difficult to inflict. On the one hand, in the face of such punishments there would continue to be a market power effect of customs union formation, which would tend to make punishment more severe as customs union formation proceeded. But on the other hand, as customs union formation proceeded, fewer competing importing and exporting regions would remain outside the defecting region, and this would tend to undercut the severity of such punishments. Since the characterization of such asymmetric punishments can become quite complex, and since their inclusion would in any event simply introduce a complicating factor that would have to be weighed against the features we have captured here, we abstract from them in what follows.

11. The notation in (7) is slightly awkward: $\overline{\tau}^c$ is an R dimensional vector in each appearance, except for when $(\tau^D,\overline{\tau}^c)$ is written, in which case $\overline{\tau}^c$ is an R-1 dimensional vector. This latter case is meant to symbolize that the domestic region of interest defects to τ^D while all other

domestic regions continue to each select the tariff τ^c.

12. Figure 2 has been drawn to illustrate the case in which there is a unique δ under which the most-cooperative stationary tariff is neutral with respect to customs union expansion. If more than one such δ exists over the relevant range--and there could be at most two--then we define $\tilde{\delta}(R_0,R_1)$ as the smallest such δ.

13. In a related context to our stationary dynamic environment, Bond and Syropoulos (1993) argue that this is the relevant case as well.

14. Note that the constant hazard rate assumption enables us to look for a single tariff for all dates within a phase. Observe also that the level of the static Nash tariff will depend upon the phase, since as shown above the Nash tariff is sensitive to the number of existing regions.

15.

$$\sum_{q=1}^{0} \delta^{q-1}\omega \equiv 0 \text{ is understood here.}$$

16. Formally, Lemmas 1 and 2 imply

$$V_2(\hat{\tau}_s^c(R_0);\cdot) \equiv \{(1-\lambda)\delta/[1-(1-\lambda)\delta]\}\,\omega(R_0,\hat{\tau}_s^c(R_0)) + \{[\lambda\delta/(1-\delta)]/[1-(1-\lambda)\delta]\}\,\omega(R_1,\hat{\tau}_3^c))$$
$$> \{(1-\lambda)\delta/[1-(1-\lambda)\delta]\}\,\omega(R_0,\hat{\tau}_s^c(R_0)) + \{[\lambda\delta/(1-\delta)]/[1-(1-\lambda)\delta]\}\,\omega(R_0,\hat{\tau}_s^c(R_0))$$
$$= [\delta/(1-\delta)]\,\omega(R_0,\hat{\tau}_s^c(R_0))$$

17. To see this, observe that $V_2(0;\cdot)$

$$\equiv \{[(1-\lambda)\delta]\,/\,[1-(1-\lambda)\delta]\}\,\omega(R_0,0) + \{[\lambda\delta/(1-\delta)]\,/\,[1-(1-\lambda)\delta]\}\omega(R_1,\hat{\tau}_3^c)$$
$$< \{[(1-\lambda)\delta]\,/\,[1-(1-\lambda)\delta]\}\,\omega(R_0,0) + \{[\lambda\delta(1-\delta)]\,/\,[1-(1-\lambda)\delta]\}\,\omega(R_1,0)$$
$$< [\delta/(1-\delta)]\,\omega(R_1,0),$$

which uses (14) and (15).

18. Specifically, (22) holds if $\delta < \delta^\dagger(R_0,R_1) \equiv [R_1(4R_1-1)^2]/[R_1(4R_1-1)^2 + 2R_0(4R_0^2-1)]$.

19. Formally, using Lemma 4, calculations reveal that:

$$V_2(\hat{\tau}_2^c;\lambda,\delta,R_0,R_1) - V_1(\hat{\tau}_2^c;\rho,\lambda,\delta,R_0,R_1) =$$
$$\{\lambda\delta/\{[1-(1-\lambda)\delta][1-(1-\rho)\delta]\}\}[\omega(R_1,\hat{\tau}_s^c(R_1)) - \omega(R_0,\hat{\tau}_2^c)] > 0.$$

20. Formally, Lemmas 1 and 2 and also $\hat{\tau}_2^c < \hat{\tau}_s^c(R_0)$ and (15) imply that $V_1(\hat{\tau}_s^c(R_0);\cdot)$

$$
\begin{aligned}
&\equiv \{(1-\rho)\delta \,/\, [1-(1-\rho)\delta)]\}\ \omega(R_0,\hat{\tau}^c_s(R_0)) \\
&+ \{\rho\delta/[1-(1-\rho)\delta]\}\{\omega(R_0,\hat{\tau}^c_2) + [\lambda\delta/(1-\delta)]\ \omega(R_1,\hat{\tau}^c_3)\} \,/\, [1-(1-\lambda)\delta] \\
&> \{(1-\rho)\delta \,/\, [1-(1-\rho)\delta)]\}\ \omega(R_0,\hat{\tau}^c_s(R_0)) \\
&+ \{\rho\delta \,/\, [1-(1-\rho)\delta]\}\{\omega(R_0,\hat{\tau}^c_s(R_0)) + [\lambda\delta \,/\, (1-\delta)]\ \omega(R_0,\hat{\tau}^c_s(R_0))\} \,/\, [1-(1-\lambda)\delta] \\
&= [\delta/(1-\delta)]\ \omega(R_0,\hat{\tau}^c_s(R_0)).
\end{aligned}
$$

21. Heretofore, we have assumed that δ is sufficiently small that a customs union expansion in which the number of regions of each country type changed from R_0 to R_1 would result in a higher most-cooperative stationary tariff. This defines a range of δ, which we stated as $\delta \leq \tilde{\delta}(R_0,R_1)$, and we argued in Section II that this range seemed to rule out few δ, other than those which we had already ruled out in assuming that free-trade is not viable. The comparative static investigated here is slightly different, since R_0 is held fixed and we examine what happens to the most-cooperative stationary tariff as R_1 declines. Thus, if $\delta \leq \tilde{\delta}(R_1,R_1\text{-}\varepsilon)$, then the most-cooperative stationary tariff will rise as R_1 declines. This additional small-δ restriction appears to rule out few if any additional values for δ; for example, when $R_1 = 2$, $\tilde{\delta}(2,2\text{-}\varepsilon) = .59$. Recall that $\tilde{\delta}(2,1) = .58$. Nevertheless, because $\delta \leq \tilde{\delta}(R_1,R_1\text{-}\varepsilon)$ is not a maintained assumption on δ, as it is employed only in deriving the comparative static for the phase-3 most-cooperative tariff as a function of R_1, we have included in the statement of Proposition 3 that this comparative static need hold only for sufficiently small δ.

22. This result does not require any additional assumption on the range of permissible δ (for example, $\delta \leq \tilde{\delta}(R_1,R_1\text{-}\varepsilon)$), and so parts (i) and (ii) of Proposition 3 are stated without additional restrictions on δ.

23. As discussed in note 8, sufficient trade diversion could also overturn our results through the possible effect on the relationship between static Nash tariffs and customs union formation.

24. We solve for the cooperative tariffs that emerge when both incentive constraints bind in every phase.

25. The move in the mid 1970's to non-tariff barriers, and to other "grey-area" measures outside of GATT rules, corresponded with the rise of protectionism and the reversal of previous episodes of liberalization under GATT. While the current paper considers only tariffs as instruments of protection, this change in instruments could be loosely interpreted within our general framework as reflecting an effort on the part of member countries to maintain the credibility of the rest of the GATT system in the face of mounting protectionist pressures which overwhelmed the existing safety valves explicitly provided for in GATT (see Bagwell and Staiger, 1990).

26. Domestic political economy considerations clearly are important determinants of trade policy, but the introduction of political economy motives into the analysis does not mean that the terms-of-trade effects upon which we have focussed would therefore cease to be important. In the political economy model of Grossman and Helpman (1993), for example, even a "purely political" government that cared only for political contributions would still take into account its ability to affect the terms of trade when determining unilaterally-optimal trade policies.

27. An important omission from this discussion is the enforcement of Article XXIV itself. While we treat the proposed changes to Article XXIV as enforceable, whether or not they are in fact self-enforcing would depend on what the private gains would be to negotiating a free trade agreement which violated Article XXIV, as compared to the future multilateral punishment suffered as a result. We abstract from such considerations here, and simply assume that the relevant incentive constraints associated with changes in Article XXIV are met.