On the Welfare Effects of a Customs Union Versus Unilateral Trade Policies Under Internal Scale Economies and Product Homogeneity

Ramkishen S. Rajan*
The University of Michigan

Abstract

Notwithstanding some recent attempts at analyzing the impact of increasing internal returns to scale (IIRS) on CU theory, research on this issue is still at a highly rudimentary stage. The potentially important implications for CU theory of the introduction IIRS were in fact independently discussed in seminal, albeit relatively neglected papers by Max Corden [1972] and Donald Mead [1968]. This paper selectively surveys and applies the earlier papers to systematically establish conditions under which a CU could, from a single country's perspective, potentially pareto dominate unilateral trade policies. Relaxation of some restrictive assumptions maintained by Corden [1972] and Mead [1968] allow for the consideration of a wide-ranging and rich milieu of theoretical possibilities of non-negligible practical significance. (JEL Classification: F02, F12, F15)

1. Introduction

The development of the so-called 'new' trade theories in the last fifteen

* Correspondence Address: Department of Economics, The University of Michigan, Ann Arbor, MI 48109, U.S.A.; Detailed suggestions and criticisms on a series of earlier drafts by Professor Sven Arndt led to a marked improvement in the paper's quality, as did comments by an anonymous referee of this journal. The author is grateful to them for their assistance. Thanks are also due to Gail Greenfield for valuable editorial assistance. It goes without saying that any remaining errors are due solely to me.

©1996 – Institute for International Economics, Sejong Institution. All rights reserved.
years (see Krugman [1992a]) has been the impetus for the recent shift in Customs Unions (CU) theory from a perfectly competitive framework a la Jacob Viner [1950], to one emphasizing, among other things, increasing internal returns to scale (IIRS) (see for instance Smith and Venables, 1988a & b). These works have in turn largely provided the analytical bases for computations of the gains from European product market integration (EU) as espoused by the Chechini Report Commission of the European Community (see Emerson and associates [1992]). Indeed, it is estimated that exploitation of IIRS could contribute over one-third of the aggregate benefits of having a single European Common Market (Caballero and Lyons [1990], p. 805). The potentially important implications for CU theory of introducing IIRS were however not only recognized by earlier authors (see the succinct surveys of CU theory by Gunter [1989], Krauss [1972] and Pomfret [1986]), but were in fact the primary focus of seminal studies by Max Corden [1972] and Donald Mead [1968], both of which have been relatively neglected (especially the latter). On the other hand, Giorgio Basevi [1970], Richard Pomfret [1975] and William Tyler [1968] demonstrated early on the possibility of import protection in the presence of IIRS acting as an inducement for export promotion.

Rather than cover all the above earlier papers and related literature in a detailed fashion, this paper attempts to selectively survey and synthesize these studies, with the aim of systematically establishing conditions under which a CU could, from a single country’s perspective, potentially pareto dominate unilateral trade policies (including unilateral trade liberalization or UTL). Relaxation of some critical restrictions imposed by Corden and Mead (henceforth denoted by C-M), allow for the examination of a rich milieu of

1. A referee of this journal has correctly noted that empirical studies on the significance of IIRS are far from conclusive (see for instance, Tybout [1993]). However the purpose of this paper is not to survey these empirical concerns. Rather, insofar as there remains substantial empirical evidence to suggest the importance of IIRS on the one hand (see for instance, Backus, et. al. [1992]), and given its influence in business and public policy on the other (see Krugman [1984], p.187 and Pomfret [1986], pp. 455-6 in general; and Baldwin [1992], Flam [1992] and Holmes [1987] in the case of the EU in particular), it is imperative to highlight and extend the analytics behind CU theory in the presence of IIRS, thus providing the primary impetus for this paper.
theoretical possibilities of non-negligible practical significance. Further extension and application of the C-M theoretical framework reveal new insights regarding the welfare effects of forming a CU in the presence of IIRS and product homogeneity.\textsuperscript{2} A number of commonly-held generalizations are also negated.

The remainder of this paper is organized as follows. The next section explicitly sets out the analytical framework to be utilized in this paper. Consistent with the above-noted older literatures on trade theory, a geometric framework is utilized. Such a framework is especially useful in studies involving imperfectly competitive market structures (which are the necessary result of IIRS), as all assumptions used in the analysis are made explicit and unambiguous (though admittedly, an explicit model is necessary and made use of when a certain degree of precision is called for). Section III considers the consequences of made-to-measure (MTM) tariffs. This provides a useful starting point for the remainder of the analysis. Section IV discusses the welfare effects of UTL under IIRS. Section V concerns itself with the phenomenon of import protection-induced export promotion (IPEP). The penultimate section, which forms the bulk of the paper, examines in detail the welfare consequences of forming a CU. This section uses the basic C-M framework to develop a number of new and important insights for CU theory under IIRS. The final section provides some concluding observations.

II. Analytical Apparatus

Consistent with received CU theory, the analytical framework is partial equilibrium, geometric and static in nature, and, as in the case of the bulk of the CU literature, is limited to the pro-typical conditions of one commodity - the focus good and 'all other commodities', and three countries - $A$, $B$ and the rest of the world (henceforth denoted by ROW). Specifically, following C-M, we assume that in autarky countries $A$ and $B$ are each served by a single domestic monopolist (or a few producers acting in collusion), and the

\textsuperscript{2} For discussions on CU theory under external returns to scale in general, and in relation to the EU, see Choi and Yu [1984] and Caballero and Lyons [1990] respectively.
good in question is completely homogenous. We assume that production costs involve constant marginal costs (MC) and lumpy fixed or irreversible costs, such that the average cost curve is declining, thus not embodying any producers' surplus. This is one particular source of IIRS described by Silberston [1972]. We also implicitly assume that there is some degree of friction/inertia in adjustment (i.e. time-lags) to increases in market sizes. Accordingly, following Basevi [1970], the cost curves might be thought of as potential or long-run costs. Countries A and B are assumed to be small (i.e. price-taking) countries, thus allowing us to abstract from complications arising from terms of trade effects. We also assume international immobility of factors. Income growth effects on demand are ignored by assuming quasi-linear demand schedules (for a discussion of these potentially important income growth effects, see Baldwin [1989]). This assumption also ensures that consumer surpluses provide exact measures of welfare effects without having to appeal to Willig [1976]. Our primary focus will be on country A, though partner country (B) and CU-wide effects will also be highlighted where considered necessary.

Wonnacott and Wonnacott [1981, p.705] have noted that "in a world in which . . . obstacles to trade exist, it is meaningless to analyze the effects of freeing trade between CU members if we assume that there are no impediments to trade with outsiders (ROW)." Accordingly, following the Wonnacotts [1981], Krugman [1991, p. 19] and Frankel, et. al. [1993], we assume that countries A and B (which form the CU) are located 'sufficiently close' to each other, such that intra-country (union) transport costs (as discussed by Donnenfeld and Hirsch [1993]) are negligible (so-called 'natural' trading blocs), while their trade with the ROW incurs a fixed per unit transport cost. This is consistent with the observation by Grant, et. al. [1993] that "(g)eographic proximity . . . appears to be one of the most important factors

3. By assuming away product differentiation, we are admittedly ignoring the important, albeit uncertain impact of product diversity on economic welfare (see Flam [1992], pp. 21-2 for an informal discussion with reference to the European context). For an early though much-neglected attempt at formalization of the 'optimal' trade-off between exploitation of cost reductions due to IIRS on the one hand, and benefits of product diversity on the other, see James Meade [1974]. Corden [1972] does also touch on this issue in the latter part of his paper.
motivating bloc formation." Following the terminology introduced by Drysdale and Garnaut [1993] and Garnaut [1994], the above objective resistances to trade are usually compounded by subjective resistances, the latter arising due to lack of information, increased uncertainty and risk, and the like, all of which tend generally to increase with distance. Wonacott and Lutz [1989, pp. 69-70] have in fact argued that since "(g)roupings of distant nations may be economically inefficient because of the high transport costs, the farther apart the prospective members, the more critically an FTA should be studied." As in Scherer and Ross [1990, pp. 106-8], these costs are, for tractability, imputed into the price of goods to and from the ROW.

III. 'Made-to-Measure' (MTM) Tariffs

Following C-M, we assume initially that both countries impose MTM non-discriminatory per unit tariffs (T) on imports. By assuming initial production where price is set at the point where the demand curve cuts the average cost curve, Tyler implicitly maintains a similar supposition. Such an assumption may well describe reality, seemingly consistent with the prohibitive trade policies maintained by a number of European economies prior to the formation of the EU (see Flam [1992], p. 16). As will be elaborated upon below, insofar as both countries are assumed to have identical domestic demands and similar fixed costs but differing average costs, the assumption of MTM tariffs being in place, implies that both countries impose dissimilar tariffs. As such, the model is in broadly similar spirit to Wonacott and Wonacott [1981], and is in sharp contrast to orthodox CU theory. The latter either ignores partner country tariffs (see Wonacott and Wonacott [1981] and De Melo, et. al. [1993]), or assumes identical pre-union tariff rates in both countries (see C-M), both of which are highly unrealistic. Accordingly, the Wonacotts have rightly emphasized that the primary motivation for, and hence perceived importance of the formation of a CU, are the important export advantages it provides to member countries. This point has also

4. Admittedly there are always exceptions, the US-Israel FTA being a case in point. Bhagwati and Panagariya [1996a & b] show, within the traditional CU theory framework, that the nexus between trade diversion and geographical proximity is not a valid generalization.
been largely ignored by received CU theory, despite being of critical importance in contemporary international trading relations (see Bhagwati [1993]), and is an integral component of the analysis in this paper.

Referring to Figure 1, we see that the world price \( P_w \) – which as noted, incorporates the transport costs from the ROW to country A – lies below country A’s minimum average costs \( AC_a \). We assume that, at the limit, \( AC_a \) equals the country’s marginal production costs \( MC_a \). In other words, following Mead [1968], we assume that the CU market is large enough such that if a single producer supplied the entire market (pricing at average cost), \( AC \) would equal \( MC \). While we are following Mead [1968] and Pomfret [1975] in implicitly assuming that average cost curves are kinked or L-shaped (i.e. average costs become horizontal at the point of minimum efficient scale or MES), arguably a more plausible situation would be one where diseconomies of scale are experienced beyond the MES, as in Basevi [1970] and noted by Pomfret [1975]. On the other hand, while we are following the con-
Construction of Pomfret [1975], Corden [1972] and Mead [1968] by assuming constant marginal costs, Basevi [1970] and Tyler [1968] assume marginal cost curves are U-shaped. These limiting assumptions regarding the costs curves do not alter the analytics obtained. Rather, the constructions are purely for convenience, allowing us to obtain definite static equilibria despite the presence of infinitely elastic marginal costs, keeping in mind the need to ensure consistency with the stylized model developed in section VI.

Let the pre-CU production costs of country B ($P_B^0$) be between $AC_a^d$ – intersection of domestic demand and average costs, denoted by point $v$ – and $P_w$ but above $MC_a$. While we will deal more explicitly with country B in section VI, note that the assumption of MTM tariffs being levied in countries A and B imply that country B’s pre-union price is the ‘break-even’ price $P_B^0$. This corresponds to point $v^1$ in Figure 2. To the extent that we assume that market sizes in countries A and B are similar, this implies that country B is ‘inherently’ more cost competitive in the production of the good vis-a-vis country A. The assumption of differing marginal costs is at variance with
that of C-M, as well as most other contemporary studies on CU, which assume similar cost conditions between the potential CU members. Indeed, Flam [1992, p. 22] has noted that “it is not realistic to assume that all firms in an industry have identical costs, as the quantitative models of (EC) ‘1992’ have typically done. . . . In any given industry and country, marginal costs will typically vary considerably, reflecting different . . . efficiency.” Focusing on country $A$, $MTM$ tariffs ($T_a$) imply that $(P_w + T_a) = AC_a^0 < (P^0 + T_a)$. In such a case, the entire demand of $X_0$ is domestically produced (Figure 1), and, by definition of $MTM$ tariffs, imports, tariff revenues and monopoly profits are all zero.

**IV. Unilateral Tariff Liberalization (UTL)**

If country $A$ undertakes UTL, referring to Figure 1, it is trivial to note that at world price $P_w$ demand rises to $X_w$, all of which is supplied by imports from the rest of the world (ROW). The consumer surplus (CS) is represented by the region $aby$. Following Cooper and Massell [1965], it has generally been accepted that a CU is pareto inferior to a policy of tariff reductions on a most-favored-nation (MFN)-basis. Indeed, Cooper and Massell [1965] drew the seemingly logical conclusion that, insofar as CUs are formed in practice, they are done so primarily for non-economic reasons (see also Berglas [1979] and Wonnacott and Wonnacott [1981]). To their credit, Cooper and Massell did however recognize that the above conclusion might potentially be invalidated under three situations (p. 747), viz. consideration of terms of trade effects; possibility of bargaining for tariff reductions between partner countries in the CU (i.e. market swapping); or the presence of IIIRS.

While Arndt [1968 and 1969], Kemp and Wan [1976] and others (see references in Krauss [1972]) have discussed the possibility of favorable movements in terms of trade resulting in the pareto superiority of a CU over UTL, and Wonnacott and Wonnacott [1981] have formally analyzed the case where a regional trading arrangement provides the framework for mutual tariff reductions, there has, to the author’s knowledge been no systematic study on establishing conditions under which a CU under IIIRS is pareto inferior to UTL. Indeed, while Mead focused on distributional issues of a CU
between the two member countries, as well as the consequences of divergences between social and private costs and benefits; Corden's primary aim was the introduction of the concepts of trade suppression and cost reduction to the CU literature to supplement the Vinerian trade creation and diversion effects. On the other hand, while touching on these issues, this paper takes a welfare-oriented approach implicitly assuming existence of the 'compensation criterion' (Varian [1992] chapter 22) in investigating the effects of CU formation in the spirit of Cooper and Massell [1965].

V. Import Protection as means of Export Promotion (IPEP)

Basevi, Pomfret and Tyler have shown that under the assumption of IIRS and the resultant introduction of an imperfectly competitive market structure, import protection could induce exports ('strategic' trade policy), and could be welfare-improving (Pomfret [1992], p. 8). This is in contrast to conventional trade theory based on perfect competition, which has emphasized that import protection can never be used to stimulate exports. In fact, by inducing a real exchange rate appreciation, as well as possibly raising the price of critical intermediate good, import protection under perfect competition might actually stifle exports.

While not a concern of any of the above-noted papers on the IPEP hypothesis, following Cooper and Massell [1965], it is important for the purpose at hand to establish whether, from the imposing country’s perspective, such a unilateral discriminatory trade policy (which is a non-preferential policy

5. More recently, Krugman [1984] has also formally restated the IPEP hypothesis within the context of a two-country VI, duopolistic market structure. Such a market structure will be taken up in section VI. Further, Frankel [1971] and Bhagwati [1988] have shown that the IPEP hypothesis does not require the assumption of IIRS.

6. While Tyler, which was the earliest of the papers, focused solely on establishing the IPEP hypothesis, the other studies had secondary concerns. Specifically, in the second part of his paper, Basevi discussed the implications of a divergence between social and private costs and benefits in the presence of IIRS; while Pomfret explored the impact of declining average costs due to dynamic learning effects, and provided anecdotal evidence for the IPEP hypothesis in the case of Israel's industrial development. Other early papers on the subject are Corden [1967] and Snape and Pursell [1973], both of which discuss the impact of production subsidies in establishing an industry experiencing IIRS.
with regard to all trading partners), is pareto superior to UTL. To do so, the IPEP phenomenon can be easily captured within our geometric apparatus (Figure 1), which can be supplemented by the minimalist mathematical model developed by Pomfret [1975].

Assume that country A sets a prohibitive tariff, hence allowing its domestic monopolist the opportunity to exercise price discrimination between the domestic market and exports to the ROW. It is assumed that market segmentation is feasible, thus making international price discrimination (dumping) effective. Country A, which under free trade would have been a net importer of the good, could in fact become a net exporter. This occurs if the profits gained by selling $X_2$ at the higher price of $P_m$ domestically (denoted by region $upqx$) exceeds the loss from (dumping) exporting $(X_1 - X_2)$ units at the world price $P_w$ (denoted by region $qefg$). Basevi [1970, p. 332] stated this most succinctly, and it would be insightful to quote him at length:

"(t)he net profit of the firm would be the difference between the monopoly profit gained at home and the loss incurred in the export market. ... Thus ... the existence of a domestic market is a perquisite for being able to export, provided the domestic market can be exploited monopolistically. For this to be possible, it is sufficient for the producer to be protected by a barrier that prevents imports (or reimports). ... The point is that, because of economies of scale, adding one market when the other is given always reduces costs...."

Obviously, the IPEP phenomenon does not depend solely on the presence of 'static' IIRS, dynamic learning effects, competition in R&D, or any other phenomenon that lead to declining average costs sufficing (see Dick [1994], Pomfret [1975] and Krugman [1984]).

Even assuming that it is profitable for the monopolist in country A to be a net exporter, it is readily obvious that given the assumption that $MC_a > P_w$, the consumer surplus from following a liberal trade policy (shown by region $aby$) must exceed the welfare benefits from pursuing the noted prohibitive trade policy; the latter being represented by the summation of the net prof-
its of \(upqx - qefg\) and consumer surplus of \(apu\). Thus, insofar as UTL always pareto dominates a policy of IPEP, following Cooper and Massell [1965], this establishes the justification for using the former as a 'benchmark' for comparisons with the effects of CU formation.\(^7\) This is an important point, as C-M as well as most contemporary studies on the welfare effects of CU under IIIRS (which have been motivated by the formation of the EU – see for instance, Emerson and associates [1990], Flam [1992] and Norman [1989]), use as their point of comparison, the pre-integration status quo (which is characterized by protectionist policies), as opposed to addressing the question of whether a CU is a potentially pareto dominant trade policy per se.

**VI. CU Involving Countries A and B**

Relative neglect of C-M has hitherto precluded a rigorous application of their (similar, though independently developed) conceptual frameworks to conduct in-depth analyses of CU theory. This section aims to fill this void in the literature. It bears emphasizing that it is not the intention here to survey the two papers, this having been done succinctly in Moore [1994] and Sodersten and Reed [1994, chapter 16]. Indeed, certain restrictions imposed by C-M, such as similar cost conditions in both countries have been relaxed in this paper, concomitantly escalating substantially the theoretical complexities that need to be systematically explored.

If country \(A\) forms a CU with country \(B\), there are a number of possible scenarios that could arise. C-M primarily focused on two polar situations viz. the producer in either country \(A\) or \(B\) completely monopolizes the CU-market.\(^8\) However, as Krugman [1984] has illustrated, there is also the interme-

---

\(^7\) Cordon also highlighted two other initial scenarios, \(viz.\) no production by either countries, and production in only one country. We abstract from these possibilities and focus instead on the sole case where there is initial production in both countries, this being the most complex and realistic situation.
diate scenario of a duopolistic situation which gives rise to ‘cross-hauling’, a phenomenon initially introduced by Brander [1981] and extended upon by Brander and Krugman [1983] and Brander and Spencer [1985]. It is shown that the C-M framework can be used to examine all three main cases. It will become apparent that the probability of each case arising turns on the CU tariff rates that are chosen, a point that has been paid scant attention to by the CU literature thus far.

A. Monopolization of CU by Country B's Producer

We assume first the case of where both countries, having initially levied MTM tariffs on imports (section III), set the Common External Tariff or CET ($T_{ca}$) at the minimum of the pre-CU country-specific tariff rates, i.e. $T_{ca} = \min(MTM_a, MTM_b)$. This is consistent with Bhagwati's [1992, p. 23] suggestion that international trading rules on CUs “insist on . . . the requirement that the lowest tariff of any union member on an item before the union must be part of the common external tariff of the union.” Given the assumption of country B being the relatively low cost producer, this necessarily implies that $T_{ca} = MTM_b = T_v$. Thus, the formation of the CU implies that country A has to lower its tariff rate. While not explicitly shown, we make the realistic conjecture that while country B's marginal production costs are greater then those of the ROW (so as to preclude exporting by the CU to the ROW – see footnote 14), after accounting for transport and other trade-related costs, country B is potentially the low cost supplier to country A. Specifically, from country A's perspective, $MC_b$ is below $P_v$ (Figure 2). Consequently, country A's producer is, in the absence of government sup-

---

9. Corden, whose work predates the contemporary literature on 'reciprocal dumping' by James Brander, Barbara Spencer and Paul Krugman, did recognize the possibility of survival of producers in both countries, but only in the presence of product heterogeneity. Smith and Venables [1988a & b and Venables, 1990] have defined economic integration as the absence of such international price discrimination. However, by their own admission (see Smith and Venables [1988a], p. 307), such a definition is more pertinent to a common market (i.e. a "genuinely unified market on a scale greater than the U.S.A."), rather than a CU or FTA.

10. However, insofar as all that is required in our case is for the CET to be between $MTM_a$ and $MTM_b$, it is plausible that the CET might exceed the average pre-CU tariffs.
port, necessarily eradicated. Under these circumstances, the producer in country B will monopolize the entire CU market.

1. Pricing Policies of a CU subject to IIERS

On reviewing C-M, Moore [1994, p. 82], concluded that in the case of “the outcome of a customs union formed between two countries . . . exhibiting economies of scale but lie above the world price, the important feature(s) . . . are the pricing policies pursued.” This important insight, which has hitherto not been explicitly developed, is done so in this section. In particular, two possible broad scenarios need to be considered: case 1 – the CU producer is allowed to price ‘freely to the market’, thus reaping excess profits; case 2 – the producer charges at average costs, hence earning only normal profits.11

Case 1: Pricing to the Market

In this case, the tariff inclusive import price from the ROW \( P_{w} + T_{cu} = AC_{w}^{d} \) sets an upper limit on the price that the CU monopoly can charge. If the producer does price at this level, the producer’s profit is the region \( nih_{s} \), with total production at \( X_{y}, X_{d} \) of which is exported to country A. While trade is diverted in the ‘static’ Vinerian sense, as noted below, there could be trade creation effect in the ‘dynamic’ sense. Further, in both cases, consumers in country B experience a cost reduction effect or a cheapening of an existing source of supply.

11. The assumption of average costs pricing, while seemingly inconsistent with normal profit-maximization theory, is typical of studies involving IIERS (see C-M and Pearson and Ingram [1980], p. 996). While most analysts have attempted to overcome this inconsistency by assuming that a mandatory pricing policy is imposed on the monopolist, such a situation could be fully compatible with a free market environment in which firms play a Bertrand competitive game in the absence of information regarding rivals’ costs (see footnote 17). The latter assumption allows us to maintain focus purely on trade policies, without invoking industrial policies, though this distinction is often clouded in popular, as well as academic discussions on government intervention (for an instance of the latter, see Krugman [1993]). The importance of focusing purely on trade policies is to abstract from the complex issue of pareto optimal government policies, which might encompass an entire milieu of trade, industrial and macroeconomic policies.
Case 2: Average Cost Pricing

In the case of average cost pricing, consumers in country A, faced with price $MC_b$ (Figure 2), will demand $X_5$ (which, by construction equals $1/2X_5$), all of which is imported from the CU partner country. In this case, there is trade creation even in the 'static' sense.

Comparison of the two scenarios with UTL reveals that the welfare-dominance of a CU over UTL does indeed depend on whether the CU monopolist prices according to case 1 or case 2. In the first case, consumers in country A are, in the absence of any intra-union redistribution of profits, necessarily worse-off in comparison to a policy of tariff reductions on an MFN-basis (shown by region sw'by). The second case is a pareto improvement over UTL (represented region ybe'c), and illustrates how a CU could benefit the partner country by enabling it to import from the cheapest source. This is consequently an instance where the generalization by De Melo et al. [1993, p. 169] that in the presence of partner country tariffs, "only if one country is willing to accept a suboptimal position . . . can the other benefit more from an FTA" than from unilateral liberalisation" is invalidated. This possibility was first noticed by Wonnacott and Wonnacott [1981, p. 712] within a general equilibrium, perfectly competitive framework.

2. Bhagwati’s Dynamic Time-Path Issue

The above model also provides us with some interesting insights into what Jagdish Bhagwati [1992] refers to as the time-path issue, i.e. impact of the formation of regional groupings today on future prospects for global economic integration (also see the elaborations by Bhagwati and Panagariya [1996a & b]). As noted, we have made the very plausible supposition that domestic CU producers cannot expand output (and thus lower average costs) for the entire CU instantaneously. Accordingly, insofar as there is a time-lag between CU formation and output expansion by the CU-based producers, there is a need for the CU to maintain a positive CET in the 'initial'

12. Insofar as the analysis by De Melo, et. al. is independent of the level of external tariffs of the regional grouping, we could substitute CU for FTA without loss of generality.
13. We ignore the other interpretation of the dynamic impact effect, viz. whether the regional trading agreement will expand membership until universal free trade is reached, i.e. whether regionalism will be inclusive or open-ended (see Rajan [1995]).
period. In particular, in order for country $B$'s producer to remain solvent in
the short-run following the formation of the CU, the CET must be main-
tained at least country $B$'s pre-CU level.

However, in the longer-run, under average cost pricing (case 2), once the
relatively low cost domestic producer has monopolized the CU market, as
long as $MC_s < P_w$, the union could completely eliminate its external tariffs.
In other words, regionalism need not necessarily act as a 'stumbling bloc' to
global free trade, and could in fact operate as a 'building bloc'. On the
other hand, if the producer is allowed to price to market (case 1), the CU
can eliminate its external tariff in the longer-run, if and only if the average
costs of the firm in country at quantity $X_r$ are less than or equal to $P_w$ (com-
pare for instance, points $i$ and $ii$ in Figure 2). We have the further interesting
point that, insofar as case 1 allows for the domestic monopolist to reap
monopoly profits, this suggests that there is an optimal tariff rate which
maximizes country $B$'s (as well as overall CU) welfare. However, unlike the
optimal tariff argument in the 'traditional' trade literature (see for instance,
Baldwin [1992]), this holds without the 'large' country assumption.

B. Monopolization of CU by Country A's Producer

The previous section assumed that, notwithstanding footnote 10, country
A had to lower its external tariff to country $B$'s pre-union level. In an
increasingly protectionist world, a more likely situation is arguably one in
which average tariff levels were raised under the regional/bilateral trade
agreement. Grossman and Helpman [1993, p. 42] have in fact concluded
that a CU "requires the assent of both governments. We have found that
this outcome is most likely... when the agreement affords enhanced pro-
tection rather than reduced protection to most sectors." Winters [1994]

14. If we assume that minimum production costs of country B were below world price,
formation of the CU could also allow for country A to begin exporting to the ROW,
this depending on the significance of trade-related costs to the ROW. In such a cir-
cumstance, while our analysis of case 1 remains intact, in case 2, abstracting from
mandatory CU sales, the firm might find it relatively more profitable to export at $P_w$,
rather than serve the CU market (in which it earns only normal profits). We have
however not focussed on this 'infant-industry' argument in light of its limited empiri-
cal relevance (see Baldwin [1969] and Johnson [1970] for early discussions).
Figure 3
Post-CU Market Conditions in Country A:
Monopolization of CU by Domestic Producer

draws a largely similar conclusion. Further, De Melo, et. al. [1993, p. 173] have noted that "(i)f lobbies are powerful, the CET may be set at a level high enough to protect producers in the least efficient country in the union." Consistent with the above, assume that the CET $T_{cm} = \max(MTM_a, MTM_b) = T_n$. More generally, all that is necessary is for $T_{cm}$ to be 'adequately high', so as to ensure that $(P_r + T_{cm}) \geq AC_a^0$, hence making it viable for the producer in either country to serve the CU (see Figure 3).^{15}

15. While this is contrary to global trade 'rules' which require that the CET must on average not be higher than the average pre-union levels, three points are of pertinence. First, the very vagueness of GATT/WTO rules allow for the ex-post justification of almost any discriminatory action (see Bhagwati [1992] and De La Torre and Kelly [1992], p.43). Second, one could think of the non-tariff barriers (NTBs) which allow for the desired objective to be attained. Third, to the extent that the CU will lead to the member countries allocating relatively greater negotiating priorities away from the ROW in favor of the union (see Arndt and Willett [1991], p. 1573, Garnaut [1994], Rajan [1995] and Winters [1994]), this could heighten the 'subjective' trade barriers faced by the ROW.
While country $B$ is the inherently low cost producer (defined as per unit cost for a given level of output), if the producer in country $A$ is able to ‘race down’ its average cost curve relatively more rapidly than its competitor in country $B$ (or conversely, the producer in country $B$ experiences relatively greater inertia in the short-term in adjusting to the increased market size), it could potentially monopolize the CU – Upton [1995] provides an interesting discussion about the factors determining the degree of flexibility of a company to changes in market conditions, including market size (also see Bell [1984], p. 111) – On the other hand, rather than assuming differing response rates of firms, one might rationalize the less efficient firm’s monopolization of the CU by appealing to political economy considerations/constraints as alluded to by Wonnacott and Wonnacott [1995] in the case of NAFTA.

As in the previous case, there are two broad scenarios: case 1 – pricing to the market; case 2 – average-cost pricing.

1. Pricing Policies of a CU Subject to IIRS

   (Case 1) Pricing to the Market

   Unlike the preceding section when country $B$’s producer was the CU monopolist, if country $A$’s producer monopolizes the CU, it cannot charge a price of $(P_a + T_{ab})$, as this exceeds the pre-union domestic price of country $B(P^*_b)$, and would thus preclude insolvency of the country $B$ producer (we abstract from international price discrimination in the short-term with the intention of raising prices after driving out the foreign rival). Accordingly, the maximum union price that can be charged by the CU producer is at (or rather, just below) $P^*_b$. We thus have the interesting result that the price charged in the CU might be the same, regardless of whether the CU monopolist is from country $A$ or $B$.

   Referring to Figure 3, at this price, as noted, quantity demanded is $X_d$ (total production is $X_d$), consumer surplus in country $A$ is given by the region $av_1s$, while the producer’s monopoly profit is represented by the region $sudd$. Thus, if the profit made by the domestic producer following CU formation exceeds the loss in domestic consumer surplus (denoted $aby – av_1s$), UTI need not pareto dominate from the perspective of country $A$. This case illustrates that, contrary to Corden, trade suppression – defined as
the replacement of relatively cheaper imports by more expensive domestic production – is not necessarily welfare-reducing. It is however unlikely though theoretically possible that the CU is welfare-superior to UTL from a union-wide perspective, as this requires that the monopoly profits reaped by the producer in country \( A \) offset the losses in consumer surpluses in both the CU members. This provides a contradiction to the generalization by Sodersten and Reed [1994, p. 341], that “(i)f there are fewer producers in the union than previously (i.e. pre-CU), welfare will increase.”

(Case 2) Average Cost Pricing

In this case, the CU price equals \( MC_c \). At this price, while demand in country \( A \) is \( X_r \), the firm in country \( A \) produces a total of \( X_m, (X_m - X_r) \) of which is exported to country \( B \). The assumption of average cost pricing implies that comparisons of the welfare effects of this scenario with one of UTL simplifies to a comparison of the relative sizes of the consumer surplus in either situation. It is readily obvious that UTL is pareto superior to CU from country \( A \)’s perspective. The assumption that \( MC_c < P^*_c \) is critical, failing which, there is, in the absence of an inter-country compensation scheme, no economic incentive for country \( B \) to form a CU with country \( A \), such a policy being welfare-reducing in comparison to an original situation where a MTM was applied. It is interesting to observe that in comparison to an initial position of non-discriminatory but non-prohibitive tariffs, country \( A \) is faced with a welfare-reducing trade suppressing effect, while country \( B \) is faced with a welfare-inferior trade diversion effect, as imports from the ROW are replaced by relatively costlier imports from the CU member. Further, unlike the situation of monopolization by country \( B \)'s producer, in this case, even under average-cost pricing, the CU could not lower its external tariffs to zero, as the CU producer is costlier than the ROW. In particular, tariffs could, in the longer-run, be lowered to at most \( (MC_c - P^*_n) \). Consequently, regionalism might serve as a 'stumbling bloc' to multilateral free trade.

C. Consequences of a Prohibitive CET

In the preceding sections (\( A \) and \( B \)), we have maintained the assumption that the CET is not prohibitive. Concerns have however been expressed
that one of the greatest threat of the present trend towards regionalism, is
the possibility that trade blocs might become exclusionist (see Bhagwati
[1992]), and, at the limit be completely autarkic (fortress Europe?). In such
a case there are three possible scenarios of interest: monopolization of the
CU market by the producer in either A or B, or a situation of co-existence of
both producers (duopoly).

1. Monopoly

In the case of monopolization by a single firm, there are two further possi-
bilities: the producer could treat the entire CU as a single market and
decide its choice variables accordingly, or it could discriminate between the
two markets under the fragmented market hypothesis. However, given our
simplifying assumptions of homogenous demand in, and negligible trans-
port costs between the two member countries, the above two scenarios are
necessarily identical.

If country A is the ‘exporting nation’, its producer’s profits are the sum-
mation of profits from both domestic sales and those from exporting to the
partner country. Specifically, referring to Figure 4, home sales of quantity \( X_p \)
and exports of \((X_p - X_p^e)\) together generate super-normal profits represented
by the region \( 2(wzgjz) = (wwjiz). \) If this exceeds the loss in domestic con-
sumer surplus (shown by region \( abyz - axz) \), it is plausible that such a situ-
tion could be potentially welfare-superior to UTL.\(^{16}\) Conversely, the country
which imports the good following CU formation is made unambiguously
worse-off. Focusing solely on efficiency considerations, insofar as there are
no supra-national income transfer mechanisms, formation of CU under such
circumstances is highly unlikely. Indeed, against this background, it is
revealing to note that Arndt and Willett [1991, p. 1570] have voiced the op-
inion that the success of European product market integration depends to a

---

\(^{16}\) Further, if we were to allow for the possibility of product differentiation and assume
that domestic consumers have Lancastrian-type preferences (thus preferring
domestically produced goods), a policy of UTL might have additional welfare losses
in the form of adverse-selection effects. This occurs as domestic firms, unable to
compete with lower cost imports from ROW in the short-term, would go out of busi-
ness (see Norman [1989], p. 426). Flam [1992, p.11] has highlighted this phenome-
non in the case of the EU.
large extent on the implementation of intra-union redistributionary mechanisms. This, as well as the recognition that political economy realities, which usually dictate that trade policies tend to be heavily biased towards domestic producers (see Frey [1985]), work in tandem to suggest that the most likely scenario is one of duopoly. This outcome is, according to Flam [1992], broadly consistent with the experience of the Western European economies following product market integration.

2. Duopoly

In general, from a single country’s perspective under conditions of duopoly, a CU would be welfare-superior to UTL if, (a) country A’s domestic producer is profitable (necessary condition), and (b) these profits from selling in the CU exceed the potential losses in consumer surplus (sufficient condition). This can be formalized within a minimalist analytical model. Given the
assumptions of quasi-linear utility functions, profit maximization generate simple linear schedules in both countries, which can be written in inverted form as: \( P = a - bX \), where \( b > 0 \); \( P \) = price; \( X \) = output; and \( X = X_a + X_b \) (where the subscripts denote sales by the producer in countries \( A \) and \( B \) respectively). For simplicity, following Meade [1974] and other related contemporary studies, assume that identical fixed costs of \( \alpha \) are incurred by both producers, such that total cost (\( TC \)) faced by the producer in country \( A \) (\( TC_a \)) and \( B \) (\( TC_b \)) are respectively given by the following: 
\[
TC_a = \alpha_a + MC_aX_a \quad \text{and} \quad TC_b = \alpha_b + MC_bX_b.
\]
By assumption, \( MC_a > MC_b \) and \( \alpha_a = \alpha_b = \alpha \). Following much of contemporary literature, assume Cournot behavior by both producers. This seems a reasonable assumption, as it implies that the producers are most interested in relative market shares. Further, given the abstraction from product heterogeneity, Bertrand competition merely recedes into one of marginal cost pricing.\(^{17}\)

Based on the above assumptions, solving for the Cournot-Nash optimal outputs (\( X^* \)), price (\( P^* \)) and profits of the focus country’s (\( A \)) producer (\( \Pi_a^* \)) derives the following:

\[
X_a^* = \frac{(a + MC_b - 2MC_a)}{3b}
\]

\[
X_b^* = \frac{(a + MC_a - 2MC_b)}{3b}
\]

\[
X^* = X_a^* + X_b^* = \frac{(2a - MC_b - MC_a)}{3b}
\]

\[
P^* = \frac{(a + MC_a + MC_b)}{3}.
\]

\[
\Pi_a^* = \frac{2[(a + MC_a + MC_b)(a + MC_b - 2MC_a) - MC_a(a + MC_b - 2MC_b)]}{9b - \alpha}.
\]

---

17. More specifically, in the case of a one-shot, simultaneous competition in prices, if both producers are fully aware of each other’s marginal costs, the low-cost producer (firm in country \( B \)) will price its goods at (or rather, just below) the marginal cost of the high-cost producer (assuming the low-cost producer’s profits are positive). At this price, country \( A \)'s producer will, ceteris paribus, go out of business, thus leaving the entire CU to be monopolized by country \( B \) producer (see Varian [1992], pp. 292-3). In this case the welfare effects on country \( A \) are completely identical to the situation described in section VI.A.1., a policy of UTL necessarily dominating the CU. If however rivals are unaware of one another’s production costs, Bertrand competition could lead to each firm pricing and producing at their minimum possible costs (i.e. average cost pricing), as there is the fear of rivals under-cutting them if they kept prices ‘high’ (see Stykolt and Eastman [1960], pp. 341-2 for an early analysis).
Referring to Figure 4, the graphical equivalent of \( \Pi_a^* \) is the region 2 (l\textsuperscript{mg}k) = (l\textsuperscript{mg}h). This is derived as follows. At output \( X^* \) (i.e. \( X_a^* + X_b^* \)), price charged = \( P^* \). Given symmetrical assumptions, total output produced by the firm in country \( A \) (domestic sales plus exports) equals \( 2X_a^* \) at an average cost of \( AC_a^* \) – note that \( 2X_a^* < X^* \) as \( MC_a > MC_b \) – Consequently, \( \Pi_a^* = 2X_a^*(P^* - AC_a^*) = 2(l\textsuperscript{mg}k) \). Consistent with the graphical representation, we assume that \( \Pi_a^* \) is unambiguously positive, failing which the producer would go out of business (this in turn implies a monopolistic market structure, which, as noted, is a trivial case).

Focusing on the welfare effects on country \( A \), a CU under duopoly would be welfare-superior to UTL. This occurs if the profit made by its domestic producer is greater than the loss in consumer surplus – the latter occurs as price under duopoly exceeds the world price \( P_w \), and referring to Figure 4, is represented by region \( (aby - aol) \) – Thus a CU would be pareto superior to UTL if net welfare \( (L) \) is positive. \( L \), which denotes the difference between the welfare gains from a CU less the welfare gains from UTL, is formally defined as follows:

\[
L = \Pi_a^* + [(2a - MC_a - MC_b)^2]/18b - (a - P_w)^2/2b
\]

The first term (i) is the domestic producer's net profit from selling in the CU market; the second term (ii) refers to the consumer surplus enjoyed by country \( A \) under a situation of duopoly – thus (i) + (ii) are the aggregate welfare benefits from the CU; and the last term (iii) is the consumer surplus under free trade (aggregate welfare benefits under UTL).

3. Comparative Statics

It would be revealing to conduct a simple comparative statics exercise to determine the effect of changes in the variables \( \zeta \), \( P_w \), \( MC_a \) and \( MC_b \) on country \( A \)'s net welfare. We abstract from the limiting, albeit important cases where changes in the variables alter the market structure (see Kabiraj and Marjit [1992]), these having been extensively discussed in the preceding sections.

\[
\frac{\partial L}{\partial \zeta} = -1
\]
\[ \frac{\partial L}{\partial P_w} = -\frac{(a - P_u)}{b} \]  

\[ \frac{\partial L}{\partial MC_a} = 2\left\{ \frac{(a + MC_b - 2MC_a - 2(a + MC_a + MC_b))}{9b} - \frac{2(a + MC_b - 4MC_a)}{9b} \right\} \]  

\[ = \frac{(-6a - 3MC_b + MC_a)}{9b} \]  

\[ \frac{\partial L}{\partial MC_b} = 2\left\{ \frac{(a + MC_b - 2MC_a) - (a + MC_a + MC_b)}{9b} - \frac{2MC_a}{9b} \right\} \]  

\[ = \frac{(2a + 5MC_b - 3MC_a)}{9b} \]  

(1) and (2) are trivial cases. Specifically (1) is unambiguously negative, as would be expected, because ceteris paribus, an increase in the domestic producer’s fixed production cost component implies a reduction in its overall profitability, and hence country A’s welfare. (2) is unambiguously positive, as a rise in world price reduces consumer surplus under a policy of UTL, thus raising overall net welfare (defined as welfare under CU less welfare under UTL). Equations (3) and (4) are far more interesting, and are discussed in some detail below.

(3) describes the changes in country A’s net welfare under UTL for increases in domestic producer’s marginal cost. The first two terms in (3) – (i and ii) – refer to the impact of increases in own marginal costs on profits. Consistent with Brander [1981] and others, under the simplifying assumptions of linear demand and constant marginal costs, these terms are in aggregate negative, i.e. a rise in own marginal costs lead to a decline in profits. This is graphically shown (Figure 5) by a movement to a lower isoprofit curve (from \( P^b \) to \( P^b_{2b} \)), as the reaction function of producer in country A (\( RF^b_a \)), which is negatively related to \( MC_a \), moves left from \( RF^b_{a} \) to \( RF^b_{a} \). However, this result is critically dependent on the goods by both firms being regarded as strategic substitutes (thus ensuring downward sloping quantity reaction functions). Specifically, under specific non-linear demand functions (such as constant elasticity) or intertemporal maximization, the result is reversed (for elaborations, see Rajan [1996] and references therein). The third term in (3) – (iii) – refers to the impact of an increase in domestic pro-
Figure 5

Duopoly under Cournot-Nash Equilibrium

Notes: 1. RF: reaction function
2. \( \pi_1 > \pi_2 > \pi_3 \)

Producer’s marginal costs on domestic consumers’ surplus. As expected, this term is unambiguously negative. Thus, within the context of this model, the overall impact of country A’s welfare for an increase in the domestic producer’s marginal costs will make UTL relatively more attractive to a CU.

(4) describes changes in country A’s net welfare under UTL for increases in the partner country producer’s marginal costs. The first two terms in (4) – (i and ii) – refer to the impact of increases in foreign rival’s marginal costs on domestic producer’s profits. Following conventional reasoning, these terms are in aggregate positive, i.e. a rise in foreign rival’s marginal costs lead to a rise in domestic producer’s profits. This is graphically shown (Figure 5) by a movement to a higher isoprofit curve (from \( \Pi_b \) to \( \Pi_b' \)), as the reaction function of producer in country B (RF\(_{b}\)) which is negatively related to \( MC_b \), moves left from RF\(_{b}^{0}\) to RF\(_{b}^{1}\). This result is however not robust, all the nuances noted in (3) again applying. The third term in (4) – (iii) – refers to the impact of an increase in foreign rival’s marginal costs on domestic consumers’ surplus. This term is obviously negative. However the overall impact cannot be definitely determined. In particular, if \((2a + 5MC_b)/3 >... \)
\[ MC_a, \  \partial L / \partial MC_b > 0, \  \text{else} \  \partial L / \partial MC_b < 0. \] Thus, even in this simplest case, we find that an increase in the partner country’s marginal costs has ambiguous effects on the focus country’s (A) welfare.

**VII. Concluding Observations**

De la Torre and Kelly [1992, p. 24] have noted that scale economies are key to understanding and assessing the impact of the NAFTA and European product market integration, while Wonnacott and Wonnacott [1981, p. 704] have stated that “(i)n many cases, we question how enlightening it is to study a CU without considering economies of scale.” Notwithstanding some recent attempts at analyzing the impact of internal increasing returns to scale (IIRS) on CU theory, research on this issue is still at a highly rudimentary stage (Grilli [1993]), De Melo, et. al. [1994, p. 189] having recently noted that “(t)he literature on imperfect competition and FTAs (CUs) is almost non-existent.” The potentially important implications for CU theory of the introduction IIRS were however independently discussed in seminal, albeit relatively neglected papers by Max Corden [1972] and Donald Mead [1968].

This paper has synthesized, expanded on and applied the basic conceptual framework by Corden, Mead and other early analyses of international trade under imperfect competition, to analyze the welfare effects of forming a CU in the presence of IIRS in the spirit of Cooper and Massell [1965]. The analysis, while static and partial equilibrium in nature, has emphasized that the introduction of production under internal scale could, under some circumstances, lead to a country being better off by forming a CU, rather than undertaking unilateral trade liberalization (UTL). This conclusion adds further weight to the paper by Wonnacott and Wonnacott [1981], in which the invalidity of Cooper’s and Massell’s [1965] thesis regarding the pareto superiority of UTL over CU as a general case was emphasized. On the other hand, this paper has underscored the point that potential welfare effects of forming a CU in the presence of IIRS are not deserving of any *a priori* generalizations. Drawing unabashedly from Paul Krugman [1987, p. 132], we might conclude that the argument regarding the welfare dominance of a policy of unilateral trade liberalization over customs union in the presence
of internal scale economies, "while not passé has irretrievably lost its innocence."

Bibliography


Bhagwati, J. and A. Panagariya [1996a], "The Theory of Preferential Trade Agreements: Historical Evolution and Current Trends," American Eco-
onomic Review, 86; pp. 82-7.


Grilli, E. [1993]. “Regionalism and Multilateralism: Conflict or Coexist-
tence?,” Development and International Cooperation, 9; pp. 3-64.


Krauss, M. [1972], “Recent Development in Customs Union Theory: An Interpretive Survey,” Journal of Economic Literature, 10; pp. 413-36.


eral Reserve Bank of Kansas.


Rajan, R. [1996], “Foreign Direct Investment, Technology Transfer and
Strategic Policy under Uncertainty," mimeo, The University of Michigan, Ann Arbor.
Willig, R. [1976], "Consumer's Surplus without Apology," American Econom-


