



International Center
for Economic Growth
Near East Program

POLICY BRIEF

Economic Policy Initiative Consortium Project
Supported by USAID/Egypt

Brief#0012

Free Trade Agreements and Rules of Origin

Trade agreements traditionally establish preferential terms governing the relationships between two or more countries. The scope of such relationships may vary substantially, but the starting point is invariably centered on the free movements of goods. In a free trade area (FTA) tariffs and quotas are eliminated on goods originating in, and traded between, the member countries. In a custom union, the same principle applies, with the added ingredient of the determination of a Common External Tariff (CET) to be applied to goods originating from non-member countries. The initial step for establishing preferential treatment is to identify which goods originate in member countries, and which do not. This brief introduces the concept of rules of origin and its uses, identifies the criteria for defining those rules, and presents a model that measures the impact of regional trade agreements with and without rules of origin on tariff policy toward non-member countries.*

Rules of Origin: Conceptual Issues

The GATT Agreement defines Rules of Origin as "...those laws, regulations and administrative determinations of general application applied by any Member to determine the country of origin goods, provided such rules of origin are not related to contractual or autonomous trade regimes leading to the granting of tariff preferences going beyond the application of paragraph I of Article I of GATT 1994".

In defining rules of origin (ROOs), one of the main objectives should be uniformity and simplicity in their administration. This is unfortunately not always the case. Currently, developing and developed countries have undertaken the arduous task towards the simplification, harmonization and liberalization of rules of origin. However, until this becomes a reality, the rules of origin remain fragmented. Traditionally these rules take into consideration different components, the main one being the *origin component*, which categorizes products according to where they were obtained. Other components include *consignment standards* and *documentary standards*. Compliance with consignment standards satisfy authorities that products shipped from beneficiary countries are the same at the port of disembarkation, i.e., that no manipulation, exchange, dilution or third country trade of products has taken place. The documentary standards require that adequate documentation of origin and consignment be submitted. Traditionally, this duty is complied with through the presentation of a declaration and/or certificate of origin.

Rules of origin were designed as an uncontroversial, neutral device essential to implementing discriminatory trade policies, compiling economic statistics, and marking a good. Once the origin of a good is known, the importing country can apply any country-specific or trade area-specific trade preferences or restrictions to the imported good (such as duty-free entry for goods originating in a free trade area, quantitative restrictions on goods originating in a country subject to a quota, or anti-dumping duties on goods from the targeted company that originate in the targeted

country). Rules of origin remained an uncontroversial, neutral device as long as the parts of a product were manufactured and assembled primarily in one country, and as long as other mechanisms for implementing protectionism existed.

The rise of multinational corporations and the production of goods in multiple stages using parts produced in different places around the world provided an opportunity to use rules of origin as an effective means of protection. *In a world where goods are produced from different parts from around the world, there is no single, correct definition of origin.* Instead, the origin of a product depends on the formulation and application of the applicable rules of origin. *As such, rules of origin can serve as an extremely effective means of protectionism* in at least two ways. First, overly restrictive definitions or applications of preferential rules of origin may deny trade preferences to products that last underwent substantial processing in a favored country, or trading area, by holding that the product did not originate in the favored country. Second, overly liberal definitions and applications of non-preferential rules of origin will extend country-specific trade restrictive measures to products otherwise exempt from them by holding that the product, even though it last underwent substantial processing in a third country, originated in the disfavored country.

Because most people had the misconception (a proper conception as long as imported goods where produced in a single country with parts and materials from that country) that the formulation and application of rules of origin result from a technical and objective process, few people paid attention to, much less scrutinized, the process of defining and applying rules of origin. This lack of transparency was heightened by the complex, technical nature of rules of origin, which would have made it difficult to realize that they were being used for restrictive purposes. Furthermore, while the GATT increasingly restricted the ability of countries to use tariffs or traditional non-tariff barriers to protect domestic industry from foreign competition, it did not regulate rules of origin. Therefore, the use of rules of origin to insure trade

* This brief is drawn from four separate studies. The first is entitled "Regional Integration and Lobbying for Tariffs against Non-members", by Olivier Cadot, Jaime de Melo, and Marcelo Olarreaga, published in the *International Economic Review*, V. 40, No. 3, August 1999. The second is entitled "Implementing Free Trade Areas: Rules of Origin and Hidden Protection", by Kala Krishna and Anne Krueger, published under the *National Bureau Working Paper Series NO.4893*, January 1995. The third is entitled "An Evaluation of the Uses and Importance of Rules of Origin, and the Effectiveness of the Uruguay Round's Agreement on Rules of Origin in Harmonizing and Regulating Them", by Joseph A. LaNasa III, which appeared under the *Harvard Law School's The Jean Monnet Working Papers Series No. 9601*, 1996. The fourth is an LLM Thesis and is entitled "Rules of Origin in International Trade", by Mariana Silveira. This brief is largely a verbatim of the original texts. The views expressed in this summary are those of the authors, and do not necessarily reflect the views of ICEG.

restrictive effects provided a means for countries to satisfy pressure from domestic industry for protection from foreign competition. By taking advantage of the fact that formulations of the rules and determinations of origin are not technically objective exercises but rather policy-influenced decisions, governments were able to protect domestic industries in a hidden, effective manner. As a result of the growing pressure to find new barriers to trade, the lack of global regulation of rules of origin, and the effectiveness of rules of origin as a protective device, **governments increasingly turned to rules of origin as a mechanism for protectionism.**

Criteria for Defining Rules of Origin

ROOs can be defined in a variety of different ways. **From a legal point of view**, there appears to be four alternative criteria. These are: (a) requirements in terms of domestic content; (b) requirements in terms of a change in tariff heading (CTH); (c) requirements in terms of specified processes that must be performed within the FTA or CU; and (d) requirements that the product has been "substantially transformed".

However, Kala Kirsha and Anne Krueger (1995) argue that these criteria are both too broad and too narrow. They are too narrow, as some of these categories are analytically similar. They are too broad since some of the four methods may differ in the details which could be specified in a number of alternative ways which alter their economic effects. For example, domestic content can be defined in terms of value added or in physical terms. Moreover, the required share of value added can be defined in terms of cost or price. These differences matter in terms of their incentive effects and therefore their effects on resource allocation.

ROOs set in terms of a change in tariff heading, are specified in terms of tariff categories. To satisfy origin requirements a product must change its tariff heading in a specified way. However, by making the changes needed more or less extensive, the origin requirement can be made more or less restrictive. Exceptions to rules on changes in tariff headings can also be used for the same purpose, namely to make the origin requirement more restrictive. When origin is instead conferred by performing particular steps in the production process, restrictiveness depends entirely on the steps prescribed and the nature of the production technology. In the case of American imports of apparel under NAFTA, the rule is one of "triple transformation". Only if each step of the transformation from raw material to finished garment has been undertaken within the FTA will preferential treatment be given.

From an analytical viewpoint the legal classification makes little sense. If ROOs specify minimum requirements for domestic value added, they can be defined in different ways. For example, a ROO can be defined as requiring a minimum proportion relative to unit cost of domestic value added. Alternatively, value added domestically can be defined as the price obtained less the value of imported unit intermediate input requirements, with value added share defined relative to price. A further variation is to put a minimum value share of FTA unit input requirements relative to price, since cost information is harder to get than price information, the price based definition is favored on grounds of implementation. Yet the effects of these three variations differ, and relatively little work has been done on this issue. Especially with the price definition of value added, since the price of the product is itself endogenous, the assumptions made in terms of timing and technology turn out to be vital.

From the above description, **it is evident that ROOs under FTAs bear a resemblance to the domestic content requirements** often imposed by developing countries' governments. Content protection schemes specify requirements on the share of domestic content in production. Failure to meet these requirements results in a penalty tariff on inputs for domestic producers or a penalty tariff on the import of the final good if the final good is imported.

The Use of Rules of Origin as Barriers to Trade

Rules of origin are divided into two categories: **preferential and non-preferential rules of origin.** Preferential rules of origin are used to determine whether certain products originate in a preference-receiving country or trading area and hence qualify for the trade preference. Non-preferential rules of origin are used for all other purposes, including enforcement of product and country specific trade restrictions that increase the cost of entry (i.e., antidumping duties) or restrict or prevent market entry (i.e. quotas). **Both types of rules of origin can be used as a barrier to trade.**

By determining whether a product originates in a preference-receiving country or trading area, and thereby enters the importing country on better terms than products from the rest of the world, preferential rules of origin allow governments to discriminate between products from different countries. That's because instead of a global trading environment ruled by GATT's principle of non-discrimination and its most favored nation clause, countries have created a growing proliferation of trading agreements that give preferential treatment to developing countries and regional trading partners. By varying the severity of the required transformation and by allowing different degrees of cumulation on a regional, donor, or global basis, countries use the rules of origin to control the degree of preference. If the preferential rules of origin are formulated so that they require a greater transformation of the product than the rules of origin otherwise would require, **the rules of origin may be serving as a discriminatory policy device that restricts trade.**

The rapid, recent spread of reciprocal preferential trading agreements that, inter-alia, liberalize trade through the creation of regional free trade areas has focused increasing attention on rules of origin and their importance. **While reciprocal, trade liberalizing preferential agreements in theory should result in net trade creation, their use of more restrictive rules of origin than the non-preferential rules, though nominally designed to prevent trade deflection, may result in trade and investment diversion.**

Trade deflection occurs when companies place a minimal processing or assembly plant in a preference-receiving country to take advantage of those trade preferences. The preferential rules of origin attempt to prevent trade deflection by establishing criteria that ensure an adequate degree of transformation in a preference-receiving country to justify allowing a good to benefit from the preference. However, the rules of origin in reciprocal preferential trading agreements often are more restrictive than necessary to ensure substantial transformation and prevent trade deflection. However, by being more restrictive, they give producers an incentive to increase the amount of intermediate and final good manufacturing, processing and assembly done within the preferential area at the expense of facilities in other countries that would otherwise have a comparative advantage. This distortion of the sourcing and purchasing decisions causes an inefficient allocation of global resources.

Overly restrictive preferential rules of origin are not designed to protect final good producers, as traditional barriers to trade are. Instead, they are designed to increase the amount of investment in production and assembly of intermediate goods and to protect and enhance the position of existing intermediate producers. **This protection of intermediate producers results in inefficient trade diversion,** and is the focus of non-member resentment of the preferential trading agreements. Furthermore, over time, the domestic intermediate producers may be replaced, or crowded out of the market, by foreign producers who relocate their intermediate production facilities to the protected area.

However, rules of origin can also serve as a traditional barrier to trade, i.e., to protect domestic producers of final goods when the rules of origin are so administratively or technically difficult

to comply with that they serve as a non-tariff barrier to trade. If the penalty for non-compliance is severe enough, they will nullify the trade preference, because no firm will seek to take advantage of it. However, if one member's market is much larger than that of other members, firms have an incentive to source factories in that country where most of the final goods are destined to be sold, so as to avoid having to comply with the origin rules. Because rules of origin are only applied to imported goods, i.e., goods crossing a national border, if the good is produced and sold domestically, no origin determination is necessary. Of course, its imported parts will have to comply with non-preferential trade law, including applicable tariffs and quotas.

Rules of Origin as a Factor of Production

Because rules of origin are an essential part of applying country-specific or trading group-specific trade preferences or restrictions, *rules of origin have a significant impact on the strategic planning of profit-maximizing firms*. For this reason, profit-maximizing firms should analyze the different rules of origin, quantify their cost, and treat them as a factor of production in determining where to source their investments, purchase their raw materials, produce or purchase intermediate materials, and assemble their final products.

Given raw or intermediate materials of equal quality, a profit-maximizing firm will purchase the cheapest one, regardless of where it is found. Because trade restrictions and preferences impact the cost of goods and may even restrict the entry of a material or good into a country, a firm should determine the origin of the raw or intermediate part it is considering using and its impact on the origin of the final good to see whether any trade restrictions or preferences apply. If a good from a certain country or trading group is subject to a quota or a measure having an effect equivalent to a quota, a firm will be less likely to use that good because its use may constrain the firm's production capabilities for its final product, since once the good's import limit has been reached, the firm is barred from importing any more of that good. If the good is subject to a tariff or a measure having an equivalent effect, then the firm should include the tariff in its cost calculations for that good in determining where to buy or produce the material.

Trade preferences and restrictions are two of many factors that must be considered in analyzing the relative and absolute costs of production. For example, if a firm produces a good which enjoys large production economies of scale, low transportation costs, and is sold in a number of countries, that firm may wish to produce all of that product in a factory located in the country that has the lowest total costs of production, including the cost of entry into the final markets. In other words, a firm should consider the cost of entry of the product into a market along with the cost of capital, the cost of labor, the skill level of labor, and transportation costs before determining where to source manufacturing and assembly plants and before determining where to buy raw materials or intermediate goods. This type of analysis should be conducted by all firms that produce products for more than one country, regardless of whether the product will be used by another producer or whether the product will be sold to the more traditional consumer.

Given the importance of origin determinations to rational profit-maximizing firms, firms should determine the origin of their final goods in advance, so that they can account for it as a factor of production in their strategic plans. However, the failure to harmonize rules of origin on a global or even national level has impeded efforts to make such determinations, as countries often apply a number of different rules of origin, often resulting in inconsistent determinations of origin, which sometimes appear to have been manipulated to achieve trade-restrictive results.

Regional Integration and Lobbying for Tariffs Against Non-members

Regional integration arrangements (RIAs) and free trade agreements have varying impacts on tariff policy toward non-member countries. Comparing free-trade areas (FTAs) with and without rules of origin, it may be shown how deep integration can lead to rising protection against non-members' imports. To examine the impacts of RIAs *with* rules of origin on protectionist pressures against non-members, Cadot, Melo and Olarreaga utilize a model that assumes a three-country world where each country produces and consumes three homogeneous goods. Each country has a competitive advantage in producing one of the three goods, and thus, it produces that good at minimum cost. Member countries set their trade policy toward third countries independently, and *without* rules of origin. Yet, the ability of members to set external tariffs independently is limited by arbitrage. The model assumes that two of the three countries integrate within a free trade agreement while the third does not join that agreement. It further assumes that each country exports the good for which it is the least-cost producer and imports the other two, but no trading country specializes in production. The aim is to examine the interaction between the external trade policy choices of countries that are members of the FTA, and on the influence that domestic lobbies are able to exert on these choices.

The degree and pattern of protectionism of FTA members against non-member countries is determined by two forces: the elimination of second-best arguments for protection in the good sector, which the non-member country produces at the least cost, and the effect of producer arbitrage on tariff revenue. How these forces interact depends on supply and demand conditions within the area and on whether rules of origin are imposed or not.

Free Trade Areas with Rules of Origin

Rules of origin are common in regional integration agreements. In FTAs, they are used to prevent the transshipment of goods between members with different external tariffs. Since consumers cannot use goods imported from the rest of the world to arbitrage consumer price differences in the area, rules of origin allow prices to differ in the area. However, any such price difference will trigger producer arbitrage. The importance of producer arbitrage stems from the fact that it is a discontinuous phenomenon: arbitrarily small tariff differences trigger discrete shifts in trade flows, affecting tariff-revenue in a discrete way. As a result, the tariff-revenue function facing an individual country is discontinuous in an FTA, even in the presence of rules of origin.

The authors make a three-way classification of FTA with regards the ability of a country's partner output of a good in the trading block to satisfy that country's import demand of that good. Accordingly, a country is a *deficit member*, if its import demand for a good produced at minimum cost in the non-member country cannot be satisfied by its partner's output at the world price, and a *surplus member* if otherwise. A *double-deficit* FTA, therefore, is one in which neither partner can satisfy the other partner's import demand at the world price. The deficit-surplus case is the only asymmetric FTA: it corresponds to a large country entering into an agreement with a small one. In an asymmetric agreement, Whenever a surplus-member country imposes a higher external tariff than the deficit member, (tariffs imposed on goods produced at minimum cost by the non-member country), the latter's output would be large enough (relative to the surplus-member country) to reduce consumer prices at the surplus member country, regardless of whether rules of origins are imposed or not. Accordingly, the consumer price of the good produced by the non-member country in a surplus member cannot exceed the level prevailing in the other member.

The model also suggests a proposition that states that in an FTA with rules of origin, under double-deficit, one member eliminates its external tariff, whereas the other maintains a positive

external tariff. The intuition in this proposition is that since neither member can satisfy its partner's import demand, producers in the low-protection country who can sell all their output in the high protection country have no incentive to contribute. As such, for a small country, tariff has a negative effect on the country's welfare, the low-protection country chooses free trade.

Another proposition suggested by the model is that under a deficit-surplus situation the surplus member eliminates its external tariff, whereas the deficit member maintains a positive external tariff. The intuition of this result is very similar to that of the previous proposition, except that the member country adopting free trade is now necessarily the surplus member, since the domestic price of a surplus member cannot exceed that of its partner.

The last possible case is double-surplus, in which both members meet each other's import demand at the world price. Here producer arbitrage under double-surplus leads to the equalization of consumer prices across the area, making rules of origin redundant.

Free Trade Areas Without Rules of Origin

Although rules of origin are prevalent in FTAs, their effectiveness can be severely undermined by smuggling, particularly in FTAs among developing countries with weak fiscal authorities. Examining the equilibrium outcome of FTAs without rules of origin also highlights the role that the latter perform in other FTAs. In the absence of rules of origin, any price differential within the area induces direct arbitrage through transshipment of imports of good produced by the non-member country at minimum cost. As a result, no price differential can be sustained, and the unique equilibrium is free trade for both members.

This leads to the following proposition: *In an FTA without rules of origin, both members eliminate their external tariff.* This result can be related to an earlier and similar result by Richardson (1995), that with large and symmetric countries, even in the presence of rules of origin, competition for tariff revenue would lead to free trade in both countries. This result appears both with and without rules of origin under a double surplus. It can be explained as follows: Arbitrage (which can be carried out with quantities of the import good produced within the area or imported from the rest of the world) equalizes consumer prices throughout the area, preventing the emergence of asymmetric tariff equilibria. Thus arbitrage works to impose symmetric tariff choices on member governments, but such symmetric solutions are unstable because they are vulnerable to undercutting. The resulting unique solution is free trade. When the source of arbitrage is the ability to bring in unlimited quantities of the import good from the rest of the world, rules of origin can limit its scope in the same way quantity constraints mitigate the force of price competition in oligopoly. However, when the source of arbitrage is the ability of each country to flood the other market's with quantities of the import good produced domestically (this is the double-surplus case), rules of origin are powerless, leading to Richardson's result.

Free Trade Areas and Custom Unions

In double-surplus FTAs (or when rules of origin are not imposed), the domestic prices of the good produced at minimum cost in the non-member country are equalized by arbitrage. In custom Unions (CUs) they are equalized because members agree to harmonize their external tariffs into a common external tariff (CET). How the latter will be determined in turn depends on the institutional setup and on the degree of political integration. One can envision two extreme cases. In the event of "shallow" integration, the CUs trade policy reflects bargaining between member country governments, themselves subject to lobbying pressure. Thus the CET reflects cooperation between member countries but not integration of their economies. By contrast, under "deep" integration, the CU's trade policy is determined by a pan-union agency subject to influence by pan-union lobbies.

Custom Unions Trade Policy under Shallow Integration

Suppose that the governments of the trading block member countries agree to an efficient solution for the CET and then bargain over how they share the benefits of cooperation via monetary transfers as, for instance, in the case of the EU's Common Agricultural Policy, the "disagreement point" being an FTA. Clearly, the final outcome will depend on the type of FTA considered as the disagreement point. Several mechanisms can be envisioned for how the surplus is shared. The model suggests the following proposition: Under shallow integration, the CU's CET reflects the preferences of the member country that is largest in the area's import-competing sector (the sector of good produced at minimum cost by the non-member country) and may be higher than the tariff of high-protection members in double-deficit or deficit-surplus FTAs. Thus the move from an FTA to a CU may imply higher protection levels against non-members. In the asymmetric equilibrium of an FTA with binding rules of origin, the low-protection country faces no lobbying pressure, since producers in that country benefit from the other country's high tariff; moreover, it collects no tariff revenue in equilibrium. In CU, by contrast, not only does it face lobbying pressure (since producers have no arbitrage opportunity left), but it also collects tariff revenue. As for the FTA's high-tariff country, it now collects revenue on all its imports, thereby reducing the marginal welfare cost of the tariff on the good produced at minimum cost by the non-member country. In the case of FTAs without binding rules of origin, consumer arbitrage forces both countries to eliminate their tariffs; thus it is as if lobbies had no voice. By contrast, the institutional setting of a CU ensures that all lobbies make their voice heard irrespective of the potential scope for arbitrage.

Custom Unions Trade Policy under Deep Integration

The model suggests that *deepening integration is likely to work toward reinforcing protectionist pressures against non-members.* In FTAs (the loosest form of regional integration), tariff-revenue competition and consumer arbitrage opportunities within the area tend to reduce the ability of member-country governments to set their trade policies vis-à-vis non-member countries independently. Consequently, the influence of protectionist lobbies on the trade policy of member countries is reduced. By contrast, a CU eliminates any scope for intra-block arbitrage so that all lobbies make their voice heard in the policy making process. Although it is not likely, in the deepest degree of integration, where lobbies merge and the CUs trade policy is directly set by a central body (such as European Commission), the equilibrium tariff may even lie outside the range of individual members' preferred levels. This suggests that a deep-integration CU may end up being even more protectionist than the most protectionist of its members.

While these outcomes formalize some of the concerns expressed recently about the spreading of regionalism on a worldwide basis, they should not be interpreted to imply that deepening regionalism is necessarily biased toward trade diversion. Regional integration helps in internalizing some of the negotiation difficulties encountered in multilateral negotiations with a large number of participants and may lead to "nontraditional" benefits. Moreover, it can promote the creation of institutions increasing the efficiency of markets, institutions that would be too costly to set up in the absence of deep integration. Nonetheless, on balance, the results here give no indication that regional integration would systematically lead to reduction in lobbying pressures against non-members.