# Rue the ROOs:
# Rules of Origin and the Gains (or Losses) from Trade Agreements

**Alan V. Deardorff**

**The University of Michigan**

February 22, 2016

**ABSTRACT**

# Rue the ROOs: Rules of Origin and the Gains (or Losses) from Trade Agreements

This paper provides two simple theoretical examples of trade among three countries in three industries, each of which includes both an intermediate input and a final good. The purpose is to show the adverse effects that rules of origin (ROOs) can have, even in a world where every country has a free trade agreement (FTA) with every other country. In the first, simplest example, ROOs prevent the move to ubiquitous FTAs from attaining the same level of welfare as would be possible with multilateral free trade. In the second example, assumed costs are adjusted to yield an even worse outcome: ubiquitous FTAs yield a level of welfare, for everyone, that is worse than if there were no FTAs at all, and all trade were subject to common nondiscriminatory tariffs. Thus, even if further multilateral trade liberalization is impossible, the move instead to an ever increasing number of FTAs with ROOs may be reducing world welfare.

**Alan V. Deardorff**

**The University of Michigan**

**Keywords:** Rules of Origin
Free trade agreements

**JEL Subject Code:** F1 Trade
F13 Commercial Policies

**Correspondence:**

Alan V. Deardorff
Ford School of Public Policy
University of Michigan
Ann Arbor, MI 48109-3091

Tel. 734-764-6817
Fax. 734-763-9181
E-mail: alandear@umich.edu
http://www-personal.umich.edu/~alandear/

February 22, 2016

# Rue the ROOs: Rules of Origin and the Gains (or Losses) from Trade Agreements*

**Alan V. Deardorff**
**The University of Michigan**

My purpose here is to make both an obvious point and one that may be not so obvious. First, that rules of origin (ROOs) can undermine the gains from free trade agreements (FTAs) when those gains depend on intermediate-input trade, or fragmentation.[1] And second, that ROOs can actually make a world of ubiquitous FTAs worse than a world of no FTAs at all, with positive MFN tariffs applying to all trade. I will present a simple, extremely stylized example in which, depending on the configuration of costs across countries, rules of origin can reduce or even eliminate completely the gains from free trade, if that freedom to trade is sufficiently restricted or discouraged by ROOs. Indeed, and in a second example, a world that is completely covered by ROO-encumbered bilateral FTAs can yield reduce welfare for all compared to the same world with positive tariffs and no FTAs.

The concern with proliferating discriminatory trade agreements is hardly new. Viner (1950) first noted that discriminatory customs unions, by causing what he called trade diversion, could hurt not only the countries outside of an agreement

* I have benefited from feedback on this topic from participants in numerous seminars: at Beijing Institute of Technology, Bocconi University, Erasmus School of Economics, Korea Institute for International Economic Policy, Tsinghua University, University of Wisconsin, and Vanderbilt University, as well as University of Michigan. I have also benefited from discussions with my students, He Yang and Jieun Lee.

[1] This phenomenon seems to be increasingly pervasive and goes under a variety of names. Some of the more common are vertical specialization, offshoring, and trade in tasks.

from which trade was diverted, but also the importing countries themselves. In the years following the negotiation of the North American Free Trade Agreement, and motivated also by the repeated expansion of the European Union, Krugman (1991) and Deardorff and Stern (1994) examined theoretically whether the expanding use of such agreements would cause welfare to rise ultimately toward the level of global free trade. However, these authors considered scenarios in which individual free trade areas grew to include ever more countries within each area, eventually dividing the world into an ever smaller number of ever larger blocs, within each of which there would be free trade. Aside from the expansion of the European Union, however, FTAs until recently have not mostly grown in size, but only in number. Instead, most countries are participants in multiple such arrangements with overlapping memberships. The exceptions are the recent move toward "mega-FTAs," such as the Trans-Pacific Partnership (TPP), which if approved will cover 12 countries that already include a large number of bilateral FTAs among themselves and the Regional Comprehensive Economic Partnership (RCEP) that will intends to do the same for a group of 16 countries.[2] Although the negotiations for the TPP have been completed (in October 2015), it is not at all certain that the agreement will be ratified in the participating countries and go into effect.

The scenario I examine here, therefore, is one in which all countries form FTAs with all other countries, but each individual FTA includes only two countries. Under that assumption, if all traded goods were final goods produced only in a

[2] In addition one might add the Trans-Atlantic Trade and Investment Partnership (TTIP), which if completed will encompass 29 countries. But since these include the 28 countries of the European Union and the United States, it is really only a bilateral FTA, albeit as very large one.

single country, then all trade would be truly free and welfare would be the same as with multilateral free trade. However, if goods can most efficiently be produced with inputs (or "tasks," in the terminology of Grossman and Rossi-Hansberg (2008)) from another country, then the rules of origin that necessarily encumber any FTA mean that traded goods embodying traded inputs may not be subject to zero tariffs, and therefore that a world of all-encompassing FTAs may not even approximate global free trade.

This is hardly the only objection that can be raised to the proliferation of FTAs. Bhagwati (2008) has written eloquently about other reasons why free trade agreements impose costs on countries that multilateral free trade would not. Rules of origin constitute only a small part of his critique, which rests less on the economic effects of preferential tariffs and more on the numerous other provisions of bilateral agreements that are often imposed by large and wealthy countries that dominate their negotiation. But Bhagwati's characterization of the "spaghetti bowl" of emerging FTAs has negative connotation, not because anyone dislikes spaghetti as food, but because each FTA covers a separate strand of trade to which preferential tariffs apply, and it is the ROOs that keep these strands apart.

That FTAs are in fact proliferating is well known, as documented by the World Trade Organization. In its tabulation of active "Regional Trade Arrangements" that have been reported to it, RTAs grew from fewer than 50 in 1990 to over 400 in 2016.[3] Of these, only about ten percent are customs unions, which do not require

[3] Many of these do in fact involve countries that are close geographical neighbors, but a significant minority do not. The "regional" designation is therefore not entirely appropriate. My use of the term "free trade agreement" is also misleading,

rules of origin. With this proliferation of FTAs, it becomes possible for each of a group of countries to have bilateral FTAs with every other member of the group. For example, as of 2012, the countries Peru, Singapore, South Korea, and the United States have six bilateral FTAs, connecting each pair of countries in the group.

Rules of origin need not, in principle, be very restrictive, but in practice they usually are. Most ROOs rest on the concept of "substantial transformation," defined either by a change in tariff heading or by a threshold for the percentage of regional value added. The former will be less restrictive the greater the number of digits in the product code used for the tariff heading, while the latter will be less restrictive the smaller is the required percentage of regional value added. In practice, most ROOs use rather aggregated tariff headings – so that imported inputs are more likely to belong to the same heading as the output – or rather high percentage requirements for regional value added, over 50%. In addition, some FTAs use more technical rules, usually to achieve even more restrictive requirements. The NAFTA's "yarn forward" requirement for textiles and apparel is a notorious example.

In addition to the restrictiveness of a ROO, there is also the question of where a product should originate in order to qualify – the issue of "cumulation." Options here are "bilateral cumulation," where only inputs from within the countries of the FTA count, regardless of other existing FTAs; "diagonal cumulation," where inputs from selected other countries count, such as other FTA partners; and "full

---

however, since these agreements involve free trade only between partners, not externally, and in practice always exempt certain "sensitive sectors" from zero tariffs even internally. But I prefer it here to the more accurate "preferential trade agreement," since the latter term includes such non-reciprocal arrangements as the Generalized System of Preferences.

cumulation," under which once a good satisfies a ROO, it full value is then counted as originating. In practice, a great many FTAs – and all involving the United States that have been approved up to now – use bilateral cumulation, which is the most restrictive. The TPP as negotiated does extend cumulation across the participating group of countries, in spite of US objections, but it does not extend cumulation to countries outside that group even if they have prior FTAs with members.

Therefore, it seems most likely that, as FTAs continue to proliferate, their ROOs will be quite restrictive in both the extent of the required transformation between inputs and outputs and in their cumulation across FTAs. For the analysis here, then, I will assume that inputs must be produced within an FTA in order to qualify for tariff-free trade of the outputs that they help to produce.

The most obvious effect of ROOs is that they reduce the amount of trade that occurs tariff-free within an FTA, and therefore reduce the gains from trade. However, that is not the effect that most concerns me here. My concern, instead, is that producers will alter their behavior in order to satisfy ROOs. That is, if a producer who would have sourced an input from outside an FTA chooses instead to source it inside the FTA in order for its output to satisfy the ROO, that increases costs, in much the same manner as through classical trade diversion. It is that increase in costs that can not just diminish, but reverse, the gains from trade.

**Example 1**

The first example has three countries, A, B, and C, each with a perfectly competitive economy able to produce in three industries, X, Y, and Z, using only labor. Each

industry has two parts, called an input and an output. Production of each part requires a fixed amount of labor per unit, and as in a Ricardian model, these constant labor requirements differ across industries and countries. One unit of each output requires one unit of the corresponding input. Consumers in each country demand the goods in equal proportions, regardless of price, spending all of their (labor) income on an equal number of units of each good. Tariffs, when present and not prohibitive, will generate revenue that is distributed to consumers to be spent like any other income.

Each country is endowed with the same number of units of labor. What that number is matters not at all, but I will set it at 252 for reasons of numerical convenience, as will become clear. Unit labor requirements in a given country are 1, 2, and 3 for the three outputs respectively and also 1, 2, and 3 for the three inputs, but with the order differing across industries and countries as well as across inputs and outputs. The ordering of the labor requirements in Example 1 is shown in Figure 1.

The configuration in Figure 1 is chosen deliberately to assure that each country is the least-cost producer (in terms of labor) of one input and one output, the most-cost producer of one input and one output, and rankings for inputs and outputs are different. Configurations are symmetric, in the sense that countries and/or industries could be renamed without changing anything.

*Autarky*

Note first what each of these countries will do in autarky. In any country, consumers demand an equal number of units of each good. To produce one unit of each good in any of these countries requires 1+2+3 units of labor for the inputs plus 1+2+3 units of labor for the outputs, thus 12 units of labor total. With 252 units of labor available, each country will produce and consume 252/12 = 21 units of each good.

Note that in each country, the cost of one good (X in Country A, for example) is three units of labor, the cost of another is four, and the cost of the third is five. If consumers were willing to substitute among the goods, they would, in each country, consume more of the lowest cost good. This, however, is precluded here by the assumption that they demand the goods in fixed proportions.

*Free Trade and No Fragmentation*

Suppose now that free trade is possible, but that fragmentation is not. That is, assume that input and output to each industry must be produced in the same country. Then the relevant costs determining comparative advantage and trade are the totals just mentioned in autarky and shown in Figure 1. With free trade, countries A, B, and C each specialize completely in goods X, Y, and Z respectively, each exporting that good to both other countries in exchange for the other two goods.

Each country requires 3 units of labor total to produce the input and the output of its favored good. Thus with complete specialization, each country produces 252/3=84 units of its good, keeps 28, and exports 56. Each country now is able to consume 28 units of each good, and the rise from 21 under autarky to 28 constitutes the gain from trade.

*Free Trade and Fragmentation*

Now add the possibility of fragmentation: that is, each country can now produce inputs and outputs separately in each industry and trade them with the other countries at no cost. This means that there are really six goods, not three, entering into international trade, and the symmetry of the example assures that each country will specialize in the two for which it has unit labor requirements of 1.

Because the final goods are demanded in equal proportions and each unit of a final good requires one unit of an input, equilibrium will have each country producing an equal number of units of an input and of an output, even though the input will not be in the same industry as the output. Since it requires just one unit of labor for each, a country will split its 252 units of labor equally between input and output, producing 126 units of each. Each country will export all of the input that it produces to the (single) other country that requires it, and each will keep 1/3 of the units of output that it produces, and export the rest. The result is that each country consumes 126/3 = 42 units of each good. Thus there is further gain from fragmentation and the resulting trade of the input, raising consumption from 28 to 42 units of each good in each country.

This pattern of specialization and trade is suggested by the top panel of Figure 2. The top two panels both reproduce the unit labor requirements from Figure 1. In the topmost panel, with free trade and fragmentation, each country produces only in the shaded cells. Each exports all of its production of the input that it produces (X in Country A) to the only country that produces the corresponding output (Country B in the case of good X), as shown by the heavy arrows near the bottom of the each row. The (only) producer of each output then sells its output equally both to itself and to each other country, as shown by the thinner arrows near the top of each row.

Fragmentation has therefore made its own contribution to gains from trade, and the contribution is large. 2/3 of the ultimate gain from trade in this example would not have been possible without fragmentation.[4]

Given this importance for fragmentation, it should not be surprising that, as we are about to see, ROOs that interfere with fragmentation can be harmful.

*Free Trade Agreements and ROOs*

Suppose now that each of our three countries enters into a bilateral FTA with each of the other countries, keeping prohibitive tariffs on imports from outside. Without trade in intermediate inputs, that second provision would be meaningless, since

[4] This, of course, is specific to the particular configuration of unit labor requirements in this example. If each country had the same requirements for input and output in each industry, then fragmentation would have contributed nothing to the gains from trade. On the other hand, if total labor requirements were the same in each industry but differed between input and output, then the entire gains from trade would require fragmentation. These and other cases can be constructed by rearranging the numbers in this example.

there would be no imports from outside; each country has an FTA with every other country.

But with the potential for intermediate-input trade and ROOs, a country's imports from one partner will be free of that prohibitive tariff only if it satisfies the ROO. Suppose that ROOs require that inputs from inside an FTA have value of at least some percentage, as is often the case in actual ROOs. Suppose, too, that this percentage is more than, say, 50%.[5] Then some of the trade that would have arisen with multilateral free trade will not satisfy that requirement, and producers or consumers will have to make other arrangements.

It is easiest to see the effects of this by starting from multilateral free trade and then moving to FTAs with ROOs, even though one hopes that, if the world were ever to reach free trade, such a change would never be contemplated.

In the example here, each country under free trade is exporting a good the input to which it does not produce itself, and it is exporting that good to both other countries, from only one of which it imported the input under free trade. Exports to the other country will therefore no longer be permitted unless it changes its source of input.

Consider, for example, industry X. Under free trade, Country A produces the input and exports it to Country B. Country B produces the output and exports it to both A and C. But now, with FTAs and prohibitive tariffs on trade that fails to satisfy the ROO, B will be unable to export X to C unless it sources the input either from

[5] Actual ROOs sometimes use such a percentage, while others require a change in tariff classification. Here, to assure that ROOs do bind in such cases, we could assume that the "input" and the "output" within a given industry are part of the same tariff category.

itself or from C. Since C can produce that input more cheaply than B, B's producers of final X will import some of their input from C, enough to produce what they will then export to C. They will continue to import A's X-input for production to consume themselves, as well as for their exports to A. This same argument will apply to each other industry: C's producers of Y-outputs will import Y-inputs from A for export to A; and A's producers of Z-outputs will import Z-inputs from B for export to B.

Thus, for the world to continue producing three units of each final good – so as to sell one unit in each country – each country will now need a total of 7 units of labor instead of 6: 3 for the 3 units of output, 2 for the two units of input not affected by the ROO (for sale of the output to itself and to the partner with the cheapest input), and 2 more for the single unit of input needed to produce the output that it sells to the other partner.[6] This means that each country is able to contribute 252/7=36 units of its final good to each country, so that each country ends up consuming 36 units of each final good in the equilibrium.

All of this is illustrated in Figure 2. As seen before, the top panel shows the patterns of specialization and trade under multilateral free trade. For good X, for example, the input is produced only in Country A, the output only in country B. All of the input is sold from A to B, as shown by the heavy arrow. This input is then

[6] There is actually another equilibrium possible, in which the ROOs divert outputs instead of inputs. That is, for the trade that becomes invalidated by a ROO, instead of the seller of the output switching its input to another source, the buyer may choose to buy the output from a different country – either itself or its other FTA partner, whichever can access the cheapest input with a cost of output of only 2. The result, just as in the equilibrium discussed in the text, is an increase of 1 unit of labor in the cost of providing the final good.

used for producing the output of X (the small dotted vertical line representing that production). And X is sold to all three countries, along the three arrows near the top of the X-row.

Much of this same pattern is repeated under FTAs, in the second panel. However, Country B cannot now continue using input from A to produce output for C, due to the ROO. Therefore it is shown instead as purchasing input from C, with a second arrow along the bottom of the cells from C to B. This input is used in a separate production process (another dotted vertical line) to produce output for sale only to C. Comparing this FTA situation with free trade, the three units of output have now required one more unit of labor, because of the switch of X-input from A to C.

I won't walk through the patterns for goods Y and Z, since they are completely analogous. For each good, the least-cost producer of the output has to source input from a not-least-cost source in order to sell to one country within an FTA. Thus for each of the three goods, production of 3 units of output (enough for one for each country) now requires one more unit of labor than was the case under free trade, raising the total labor requirement for both inputs and outputs from 6 to 7.

Therefore, moving from free trade to FTAs causes consumption of each good to fall from 42 to 36. Or, from the perspective of autarky, ROOs have reduced the gains from trade, which consisted of a doubling of consumption from 21 to 42 with free trade and fragmentation, but which now raise consumption only from 21 to 36.

All of this is summarized in the bottom panel of Figure 2. It shows the units of labor required to provide consumers everywhere with one unit of each of the three goods, X, Y, and Z. As noted earlier, this cost was 12 under autarky, fell to 9 with free trade but without fragmentation, and to 6 with free trade and fragmentation. With FTAs and binding ROOs, as the input for one unit of each good is shifted to a higher cost source within an FTA, this cost rises to 7. The last two rows of the table, mentioning tariffs, will be explained below.

Thus, ROOs are harmful by interfering with fragmentation. Of course, it could be worse. If ROOs had stopped fragmentation completely, then we'd be back in the case seen before of free trade without fragmentation, when the gains from trade were more seriously reduced. Fortunately, ROOs do not prohibit the fragmentation that takes place within an FTA.

On the other hand, one might imagine that producers would be reluctant to source some of their inputs from one country and some from another, and to keep track of and document their sources sufficiently to convince customs officers that their exports do in fact qualify under a ROO. If that is the case, they might opt out of fragmentation entirely and we would indeed be back in the inferior case of free trade without fragmentation.

*Non-prohibitive tariffs, no FTAs*

So far I have assumed for simplicity that, if tariffs are present, they are prohibitive on all trade not covered by an FTA. What if all non-FTA trade is covered instead by

an *ad valorem* tariff $t$, one that may not be high enough to stop trade? What will happen then depends on the size of the tariff.

Suppose that labor costs are as assumed in Example 1 and that there are no FTAs. Consider just good X, with analogous results for Y and Z. Country A would be able to produce X-output using its own X-input for a cost of 3, and A would therefore be able to export X for a price $3(1+t)$. Country B can import the X-input from A for a tariff-inclusive price, $(1+t)$. B will therefore import the X-input from A only if $(1+t)$ is less than B's labor cost of producing the X-input itself, 3, and thus if $t<2=200\%$. Therefore, B's cost of X will be $\{1+\min[(1+t),3]\}\leq 4$, and B will be able to export X for a price of $(1+t)\{1+\min[(1+t),3]\}$. Similarly country C will import the input from A if $t<1=100\%$, so its cost of X will be $\{3+\min[(1+t),2]\}\geq 4$, which is the upper limit on country B's cost of X. Therefore, C will not export good X regardless of the tariff.

From these calculations, it can be verified that if the tariff is not too large, then country B will import the input from A and export the output to both countries, just as under free trade. But if the tariff is somewhat larger – larger than would equate B's export price, $(1+t)\{1+\min[(1+t),3]\}$, to country A's cost, 3 – then B will still import the input from A but it will export only to C, leaving A to produce the output of good X for itself. The cutoff for this change is $\hat{t}$, which solves

$$(1+\hat{t})(2+\hat{t})=3$$

so that $\hat{t}$ is approximately $0.3=30\%$.[7]

Thus if all non-FTA trade is subject to a common tariff that is positive but less than 30%, then the world achieves the same equilibrium as if trade were perfectly

[7] $\hat{t}=0.3028$

free. This is an artifact of the assumption that goods are consumed in fixed proportions, so that there is no substitution in demand and therefore no consumer cost from distorting prices.

If the tariff is above 30%, then Country A produces good X for itself, both input and output. Depending on how high the tariff is, Country B may also use the X-input from itself, and Country C may produce the X-output for itself. Thus as the tariff rises, the labor costs of serving the markets jump up each time one of the countries switches from a least-cost input or a least-cost output.

*Non-prohibitive tariffs, FTAs*

If external tariffs are not prohibitive when FTAs are formed, then a producer has the option of continuing to use an input from outside the FTA, thus potentially violating the ROO, and simply paying the tariff on exports into its FTA-partner country. In the X-industry of Example 1, producers of the output in Country B could continue to use input from Country A even for sales to country C, and will do so if the tariff adds less to cost than sourcing from country C. When will this be the case? B's cost of good X is 2 if it uses input from Country A, and its tariff-inclusive export price to C is therefore $2(1+t)$. Alternatively, Country B can import the more costly X-input from country C at a cost of 2, and thus achieve an export price of (2+1)=3. Only if the tariff $t$ is above 50% will this second option be preferred. For tariffs less than 50% B's producers of X will ignore the FTA with C and the associated ROO. Thus the result found above – in which the ROO causes producers of X to switch to a higher-cost source of input and thus reduces world welfare – occurs within Example 1 only

for tariffs that are large enough to distort trade even without FTAs. This is implied by the last two rows of the bottom panel of Figure 2. That is, if the common tariff is less than 30%, the absence of FTAs achieves the same welfare as free trade. And if the tariff is more than 50%, the presence of all bilateral FTAs reduces welfare below that level.

Exactly what happens for tariffs between these two limits is unclear. Example 1 therefore leaves unanswered the question of whether a world of FTAs and ROOs can be worse than a world of no FTAs and positive tariffs. For that I now turn to another example, similar to Example 1 but with the numbers adjusted to yield a definitive result.

**Example 2**

Example 2 is shown in Figure 3. The configuration of relative costs is very similar to Example 1, but the sizes differ. I will now assume, again for convenience, that each country has 780 units of labor.

The configuration of costs is again such that, if trade were free and fragmentation were possible, then each country would specialize in the same input and the same output as in Example 1. Since each of these uses 10 units of labor, the cost to provide one unit of each good to consumers would be 60 units of labor, and consumers in each country would enjoy 780/60=13 units of each good.

If instead there are ubiquitous 2-country FTAs, as before, and if producers respond to binding ROOs by sourcing inputs within each FTA, then the cost of one input in each country rises from 10 to 15, so that the cost of a bundle with one unit

of each good becomes 65. Consumers in each country now get only 780/65=12 units of each good.

Will this happen? To see, we need to find whether there is a tariff that is both low enough to permit the free-trade equilibrium without FTAs, yet high enough to induce within-FTA input sourcing in the presence of FTAs.

Without FTAs, producers of X in country B will continue to buy the X-input from A and successfully export X back to A only if its cost, inclusive of both the tariff on its input and the tariff on its output, is less that A can produce good X for itself: $(1+t)[10+(1+t)10]<40$. That is, the tariff needs to be less than about 56% in order for production and trade to conform to the free-trade pattern.

With FTAs, producers of X in country B have a choice between 1) buying input from A for 10 and exporting to C for $(1+t)(10+10)$ or 2) buying input from C for 15 and exporting back to C for (10+15). They will do the latter if it is cheaper, that is, if $(1+t)20> 25$, or $t>.25=25\%$. Thus there is indeed a range of tariffs, from 25% to 56%, such that the free trade optimum is achieved in the absence of FTAs while FTAs and ROOs cause inefficient sourcing and a rise in costs from 60 to 65, or more than 8%.

**Is This as Bad as it Gets? Or as Good?**

These have been just examples, very stylized, and they are therefore not likely to be indicative of how costly ROOs may be in the real world. But will actual ROOs be more costly, or less? Two arguments come to mind, one suggesting the latter, the other the former.

In one respect, these examples have overstated the harm that ROOs might do. ROOs often require that some minimum fraction of a good's content originate within the FTA, and I have assumed that if the least-cost input violates that constraint, it will not be used at all. But that need not be the case. Producers might just reduce but not eliminate their use of an input from outside the FTA, reducing it just sufficiently to satisfy the ROO. To allow for that here would be too messy to be worth doing, but it might be very much worth doing for actual producers who want to benefit from the reduced tariffs of an FTA while still buying inputs as cheaply as possible. This is a possibility that has been formalized nicely by Yang (2011) in a partial equilibrium model whose purpose is somewhat different from mine here.

But there are other respects in which the example here may severely understate the harm of ROOs. In the examples here, there are only two stages of production, a single input followed by a single output. In fact however, many modern manufacturing products are the result of multiple stages of production, each of which may be and often is sourced in a different country. This version of fragmentation has been called "slicing up the value chain" by Krugman (1996), suggesting not just two slices, but many. In addition, there were only three countries in the examples here, when in practice there are many more.

To deal first with further slicing of the value chain, suppose we expand our Example 1, still with just 3 goods and 3 countries, but to include three production stages instead of just two, now named S1, S2, and S3 (with S3 the final output). A configuration of unit labor requirements analogous to Example 1 above is shown in Figure 4. Note that now the least-cost producer of any final good (S3) finds its least-

cost inputs in two different countries. Therefore if it has ROO-encumbered FTAs with both countries, sales to either will require that it use an input that is not least-cost. In the earlier example with only two stages of production, we saw that the labor cost of supplying one unit of output to each of the three countries rose from 6 under free trade to 7 under FTAs. Here, by the same reasoning, the labor cost rises from 9 under free trade (higher, because of the extra stage of production) to 11 under FTAs. Thus the cost of FTAs relative to free trade rises from 1/6=17% to 2/9=22% when we go from 2 stages of production to 3 in an example that is otherwise essentially the same. As we expand the number of stages – or slices in the value chain – this cost will only rise.

Now consider instead a greater number of countries. In order to keep the simplicity of each country being least cost for one input and one output, we need to expand the number of industries by the same amount as well. Let there be 4 industries and 4 countries. Figure 5 again shows a pattern for unit labor requirements across these 4 goods and countries analogous to Example 1 above. Look at the least-cost producer of the final good W, Country B. Its least-cost source of input is Country A, but under FTAs with binding ROOs, it won't be able to use that input for sales to either of countries C and D. Nor can it use the second-least-cost source, Country D, for sales to Country C. So under free trade, supplying one unit of a final good to all four countries requires 8 units of labor (4 for inputs and 4 for outputs), but supplying that under FTAs requires 11 (in the case of good W: 1 each for inputs to sales to A and B, 2 for input to sales to D, 3 for input to sales to C, and 4 for outputs). Thus the cost of supplying each good rises from 8 to 11, or 3/8=38%.

Thus, adding goods and countries raises the cost, at least in this particular case, by even more than adding stages of production.

All of this suggests that, in the real world of large numbers of goods, countries, and stages of production, the cost of ROOs may be considerable.

*A Final Caveat*

A final caveat, however. I have assumed either prohibitive tariffs or rather high ones throughout. In practice, most tariffs among industrialized countries are quite small as a result of many rounds of multilateral trade liberalization under the GATT. Even tariffs in most emerging markets are much lower than they once were, mostly for reasons other than multilateral trade liberalization, in which emerging-market countries did not participate. Thus the costs of ROOs and FTAs illustrated in the examples here are considerably overstated. Indeed, I'm told that it is not uncommon for traders not even to try to satisfy the ROOs in some FTAs. They simply pay the rather low tariffs that exist as though there were no FTA. To the extent that this is the case, FTAs are doing no harm, but they also are doing no good.

## References


Bhagwati, Jagdish. 2008. *Termites in the Trading System: How Preferential Agreements Undermine Free Trade*, New York: Oxford University Press.

Deardorff, Alan V. and Robert M. Stern. 1994. "Multilateral Trade Negotiations and Preferential Trading Arrangements," in Alan V. Deardorff and Robert M. Stem (eds.), *Analytical and Negotiating Issues in the Global Trading System*, Ann Arbor: University of Michigan Press, Chapter 2, pp. 27-94.

Grossman, Gene M. and Esteban Rossi-Hansberg. 2008. "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, vol.98, no.5, December.

Krugman, Paul R. 1991. "Is Bilateralism Bad?" in Elhanan Helpman and Assaf Razin (eds.), *International Trade and Trade Policy*. Cambridge: MIT Press.

Krugman, Paul R. 1996. "Does Third World Growth Hurt First World Prosperity?" *Harvard Business Review* 72, pp. 113-121.

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

Yang, He. 2011. "How Free Are Free Trade Agreements?" Senior Honors Thesis, University of Michigan, April 21.

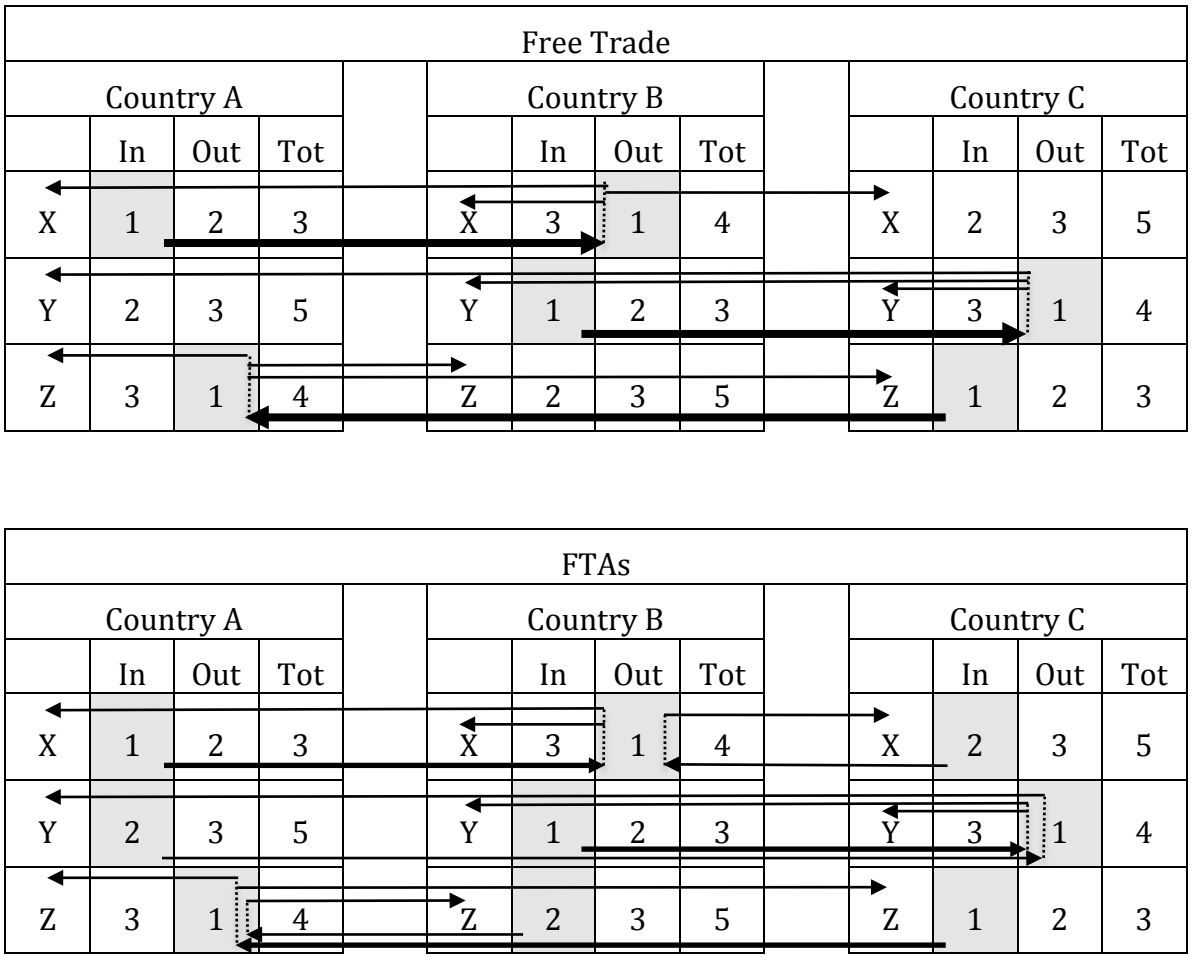
Figure 1:
Unit labor requirements in three countries, Example 1

| Country A | | | |
|---|---|---|---|
| | In | Out | Tot |
| X | 1 | 2 | 3 |
| Y | 2 | 3 | 5 |
| Z | 3 | 1 | 4 |

| Country B | | | |
|---|---|---|---|
| | In | Out | Tot |
| X | 3 | 1 | 4 |
| Y | 1 | 2 | 3 |
| Z | 2 | 3 | 5 |

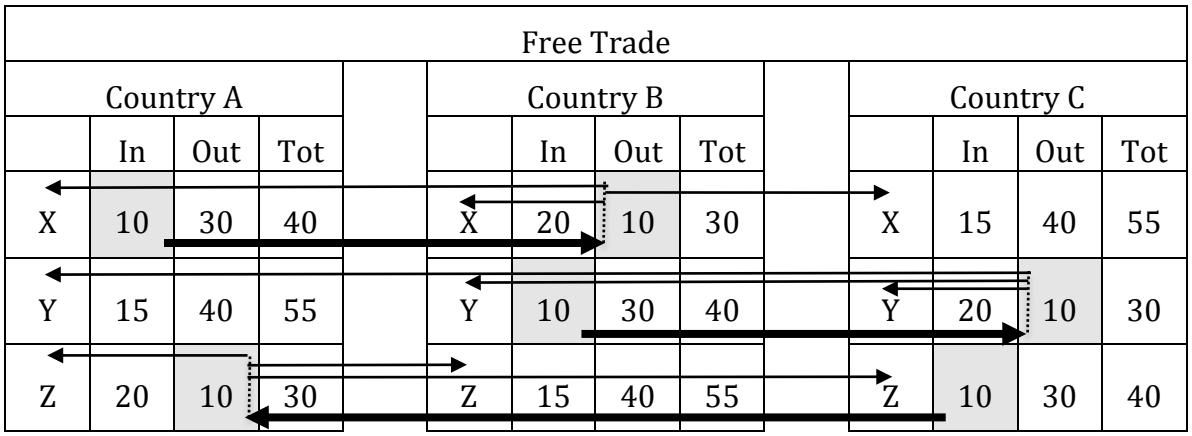
| Country C | | | |
|---|---|---|---|
| | In | Out | Tot |
| X | 2 | 3 | 5 |
| Y | 3 | 1 | 4 |
| Z | 1 | 2 | 3 |

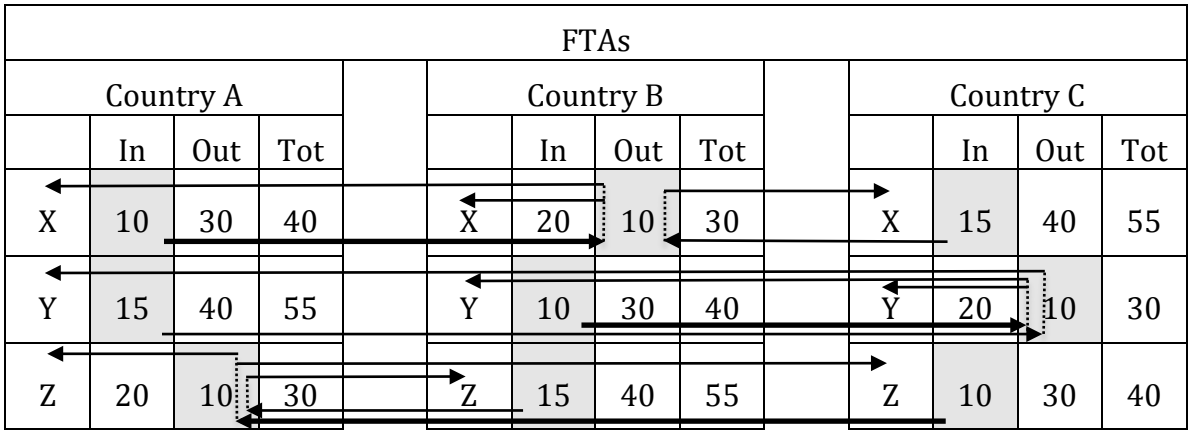
Figure 2:
Patterns of input and output trade in Example 1, Free Trade and FTAs




| | Labor cost of X=Y=Z=1 | Per-country consumption of each good |
|---|---|---|
| Autarky | 12 | 21 |
| Free trade, no fragmentation | 9 | 28 |
| Free trade, fragmentation | 6 | 42 |
| FTAs, binding ROOs | 7 | 36 |
| | | |
| No FTAs, tariffs less than 30% | 6 | 42 |
| FTAs, ROOs, tariffs more than 50% | 7 | 36 |

Figure 3:
Patterns of input and output trade in Example 2, Free Trade and FTAs




| | Labor cost of X=Y=Z=1 | Per-country consumption of each good |
|---|---|---|
| Autarky | 125 | 6.24 |
| Free trade, no fragmentation | 90 | 8.67 |
| Free trade, fragmentation | 60 | 13 |
| FTAs, binding ROOs | 65 | 12 |
| | | |
| No FTAs, tariffs less than 56% | 60 | 13 |
| FTAs, ROOs, tariffs more than 25% | 65 | 12 |

Figure 4
An example with 3 stages of production

| Case 2 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country A | | | | | Country B | | | | Country C | | | |
| | S1 | S2 | S3 | | S1 | S2 | S3 | | S1 | S2 | S3 |
| X | 1 | 2 | 3 | X | 3 | 1 | 2 | X | 2 | 3 | 1 |
| Y | 2 | 3 | 1 | Y | 1 | 2 | 3 | Y | 3 | 1 | 2 |
| Z | 3 | 1 | 2 | Z | 2 | 3 | 1 | Z | 1 | 2 | 3 |

Figure 5
A 4-good, 4-country Example

| Case 2 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country A | | | Country B | | | Country C | | | Country D | | |
| | In | Out | | In | Out | | In | Out | | In | Out |
| W | 1 | 2 | W | 4 | 1 | W | 3 | 4 | W | 2 | 3 |
| X | 2 | 3 | X | 1 | 2 | X | 4 | 1 | X | 3 | 4 |
| Y | 3 | 4 | Y | 2 | 3 | Y | 1 | 2 | Y | 4 | 1 |
| Z | 4 | 1 | Z | 3 | 4 | Z | 2 | 3 | Z | 1 | 2 |