Effects of Rules of Origin on FTA Utilization Rates and Country Welfare

Nan Xu*

This Draft: August 2016

Abstract

Free trade agreements are generally under-utilized. Assuming full utilization overstates the benefits and overlooks the substantial distortions generated by free trade preferences. This paper investigates the impacts of rules of origin on FTA utilization rates and on country welfare in the context in which there is a vertical production linkage between FTA members. I develop a general equilibrium model featuring the variable and fixed costs of complying with ROOs. A binding ROO forces exporters to employ more locally produced inputs instead of using the most efficient manufacturing process. Additionally, firms face fixed documentation costs due to the administrative process of obtaining certificates of origin. Whether firms utilize an FTA depends on the tariff benefits net of the extra costs. Both local analytical solutions and global numerical results show that: (1) as the ROO becomes stricter, the FTA utilization rate decreases and the dominant extensive margin drives down the wage rate in the downstream country; (2) a slightly binding ROO encourages regional production, and the welfare effects of ROOs are ambiguous depending on the elasticity of substitution between varieties.

Keywords: free trade agreement, rules of origin, general equilibrium, welfare effect
JEL Classification: C63, F12, F13, F16

*Department of Economics, University of Colorado at Boulder, 256 UCB, Boulder, CO 80309-0256. Email: nan.xu@colorado.edu

1 Introduction

Free Trade Agreements (FTAs) are one of the most popular forms of trade liberalization in the past 10 years. The World Trade Organization records more than 200 active regional trade agreements by January 2012. FTAs are designed to promote bilateral trade between participating countries by reducing tariffs and other trade barriers. The entry into these agreements is expected to be beneficial for exporters, since preferential market access can help firms export more and earn more profits. Most analyses so far on the impacts of FTAs assume that these agreements are used by all exporters. However, recent statistics reveal that the average FTA utilization rate is far below 100 percent, meaning that only a proportion of exporters actually take advantage of FTA preferences. Why are many exporting firms unwilling to engage in the FTA framework? What are the actual effects of FTAs on bilateral trade and country welfare?

This paper documents three stylized facts on FTA utilization and the associated costs generated by rules of origin, which motivates a re-examination of FTA welfare effects in the context of partial utilization. First, the utilization rates of FTAs are low in general. Most trade within an FTA region takes place under the MFN tariff rates, rather than under the preferential tariff rates. For instance, the US Generalized System of Preference (GSP) provides duty-free market access to developing countries, but 40 percent of imports qualifying for GSP entered the US market without claiming the tariff benefits in 2008 Similarly, even though the ASEAN-Korea FTA permits 5 percent tariff reduction, only one fifth of Korean exporters utilize the preference (Cheong and Cho, 2009). This suggests that assuming full utilization would overestimate the benefits of FTAs on economic outcomes. Second, strict rules of origin could be a reason for low FTA utilization rates, as it prevents goods that contain high non-regional value content from being eligible for preferential tariffs, especially in regions where countries have close vertical linkages in production with FTA outside countries. The methods of origin determination vary across products and across FTAs, which further complicates the use of FTA preferences. Third, administrative compliance and documentation costs offset tariff margin attractiveness. Rules of origin are often expensive to document. Exporters must obtain a certificate of origin from its national government and present it to the customs authority of the importing government. Herin (1986) shows that the cost of proving origin leads over a quarter of European FTA exports to pay the MFN tariff, even if the products satisfy origin. These three facts together suggest the necessity of incorporating the variable and fixed costs of complying with ROOs into FTA welfare analysis.

In this paper, I develop a general equilibrium model in which utilizing FTA preferences incurs extra costs and heterogeneous exporting firms endogenously choose whether to be FTA users. ROOs intend to limit preferences to member parties. Free trade agreements are conditional policies, in that they require exporters to use at least some level of locally produced inputs. Products are eligible for zero tariffs only if they are actually produced in the FTA region. Otherwise, the Most Favored Nation (MFN) tariffs [1] apply. When the ROO is binding, firms have to deviate from their

[1] MFN tariffs are what countries promise to impose on imports from other members of the WTO, unless the country is part of a preferential trade agreement (such as a free trade area or customs union). In practice, MFN rates are the

optimal production strategies and employ more domestic inputs, which generates additional costs, in order to meet the ROO criterion. Additionally, to claim the trade preference, exporters must obtain a certificate from its national government attesting that the good has met the ROO. The certificate should be presented to the customs authority of the importing government to qualify for the preferential tariff rate. Going through this administrative procedure incurs additional costs, called fixed documentation costs. Thus, exporting firms face a tradeoff between enjoying lower tariffs and paying more variable and fixed costs due to complying with ROOs. Heterogeneous firms in terms of productivity and intensity of imported inputs would respond differently to the implementation of FTAs, resulting in the general under-utilization of FTAs.

My model differs from previous studies in that it focuses on the case where there is a vertical linkage in production between FTA participating countries. This is an interesting case for study for two reasons. First, ROOs can provide hidden protection for producers of intermediate inputs within the FTA region. The degree of restrictiveness of the ROO could affect compliance costs and the value chain in the most fundamental way. The downstream country could source intermediate inputs worldwide, but turn to import from regional suppliers if the agreement sets a strict rule of origin. For instance, the trade value of automobile parts and intermediate inputs from the US to Mexico increases dramatically after the implementation of NAFTA (Oliver et al., 2002). AFTA also promotes trade in intermediate goods between ASEAN countries (Hiratsuka et al., 2008). Second, such a production connection between countries is important for welfare analysis, since it rationalizes welfare gains following a slightly stricter rule of origin. The conventional studies argue that a stricter rule of origin reduces the FTA utilization rate and mitigates the potential benefits of the preferential trade arrangement. These conclusions are based on the assumption that the FTA countries trade horizontally differentiated goods and there is no vertical production linkage between them. However, as we see in the real world, a large amount of FTAs are signed between developed and developing countries which are good at different stages of production. In such cases, ROOs are not always bad. A binding ROO could not only effectively protect regional producers of intermediate goods from import competition, but also raise country welfare when the elasticity of substitution between varieties are relatively small.

In order to explore the implications of heterogeneous utilization of a free trade agreement, I show global numerical solutions to the model by using GAMS. The further decomposition of welfare effects of ROOs is conducted by local comparative static exercise in the neighborhood of full utilization. The model yields three main results. First, within an industry, more productive firms choose to comply with ROOs and utilize the preferential tariffs granted by FTAs, and less productive firms choose to not comply with ROOs and pay the most favored nation (MFN) tariffs. This is due to the fact that it is more costly for a firm with a lower productivity to switch to more expensive intermediate inputs produced locally, and the firm, as a result, is less likely to comply with the rule and claim preferential tariffs. Also, an industry that heavily relies on imported inputs tends to have a low FTA utilization rate if the agreement sets a small tariff margin, a strict ROO,

highest (most restrictive) that WTO members charge one another.

and complicated documentation procedures.

Second, the intensive use of locally produced inputs due to ROOs generates distortions in the labor market. As ROOs become stricter, there are two opposing factors that influences labor demand. On the extensive margin, the trade preferences will be utilized by fewer exporters. For those who exported under preferential tariffs before, an exit from the FTA framework leads to a reduction in the demand for local inputs, since they no longer need to comply with the origin rule. On the intensive margin, for those who are still using FTA, more domestic inputs are demanded due to the tightening local content requirement. The overall effect on wage depends on the distribution of sales of FTA users and nonusers. The local comparative static analysis and the numerical results show that the extensive margin dominates. The total demand for labor decreases in the restrictiveness of ROOs. So does the equilibrium wage rate.

Third, regarding the welfare effects of ROOs on FTA participating countries, the elasticity of substitution between varieties plays an essential role. The welfare level is measured as real income. The overall effects of ROOs on country welfare can be decomposed into three channels. One is income channel, in that a stricter ROO generates large distortions in manufacturing process and forces more exporters to be FTA non-users and pay the MFN tariffs. Thus, the importing country earns more tariff revenues. Also, labor income changes with ROOs in a general equilibrium framework. The second channel is through terms of trade. The exporting prices charged by the firms which are affected ROOs are jointly determined by the restrictiveness of ROOs, domestic labor cost, and the MFN tariff rates. Higher individual prices benefit the exporting country but hurt the importing country. The last one is variety channel. The price index of differentiated varieties decreases with the number of varieties available in the market. Thus, firm entry raises country welfare. The magnitude of the variety effect depends on how substitutable the varieties are.

The overall effects of ROOs on country welfare are considered in two seperate cases. If the elasticity of substitution between varieties is relatively small, both FTA participating countries experience welfare increase, which is mainly driven by the significant positive variety effect. In such a case, consumers benefit more from the introduction of new varieties. Resources are allocated to the more efficient producers and a tighter ROO induces the production of intermediate goods to move from the outside country to within the FTA area. ROOs help protect regional producers and promote regional trade in intermediate and final goods. If the elasticity of substitution is relatively large, the welfare of the upstream country, which is more efficient in producing intermediate inputs, develops an inverted-U shape relationship with ROOs. When the ROO is slightly binding, the positive income effect and variety effect outweigh the negative price effect, and as a result, country welfare increases with ROOs. As the ROO rises further, the distortions in firm production generated by a strict ROO have a larger impact on product prices and the negative price effect (terms of trade) is a more important determinant of welfare, implying a decrease of welfare with ROOs. The welfare of the downstream country declines following a stricter ROO, since variety effect is negatively correlated to the elasticity of substitution between varieties and the negative terms of trade effect

dominates. Both local analytical solutions and simulation results confirm these conclusions.

There is usually a hot debate over a suitable rule of origin for each industry before a free trade agreement is signed. The underlying policy goal of setting ROO is to protect domestic producers and improve returns to local factors of production. But the actual effects of origin rules are not well understood. Demidova et al. (2006) and Ju and Krishna (2005) incorporate binding ROOs which allow firms to escape paying tariffs but involve an additional per-unit cost in order to satisfy local content requirements. They examine the impacts of FTA on wage in a general equilibrium framework, and conclude that stricter ROOs reduce the wage while raising the cutoff for firms invoking ROOs. But their work does not take into account domestic input producers whom ROOs primarily aim to protect. There is also no attention given to the magnitudes of the increase in unit cost caused by rules of origin. The other type of extra costs that may prevent exporters from utilizing FTA preferences is fixed documentation costs. In the literature, Brenton and Manchin (2003) points out that FTAs are not fully exploited because of the costs of proving origin and difficulties in passing though customs. A simpler and less demanding system would make it easier for small companies in developing countries to obtain preferential access to foreign markets. Cherkashin et al. (2009) argue that firm heterogeneity and higher fixed costs of exporting due to using FTAs can rationalize partial utilization. But this strand of literature does not consider firms' responses to local content requirements in order to be qualified for trade preferences. Neither the additional variable costs nor fixed costs accompanied with FTAs are negligible in the analysis of firms' decisions on the use of FTAs.

The rest of the paper is organized as follows. Section 2 documents three stylized facts on FTA utilizations and Rules of Origin. In section 3, I derive a partial equilibrium to analyze firms' decision on the use of conditional trade preferences and investigate the determinants of FTA utilization rates. I then turn to a general equilibrium model to examine the role of ROOs in determining wage rate and country welfare in Section 4. In section 5, I conduct comparative statics analysis to explore the local effects of ROOs on FTA utilization and welfare. Section 6 reports the simulation results. Section 7 concludes.

2 Stylized Facts

2.1 Utilization rates of free trade agreements are low

Free trade agreements are designed to promote regional exports and imports by granting traded goods preferential or even zero tariff rates. Such tariff margins are obviously beneficial to exporting firms, and therefore most studies on the effects of FTAs on trade and welfare assume all exporters in the region are claiming preferential tariffs. However, in fact, the utilization rates of FTAs are far below 100 percent. Utilization rate is defined as the ratio of value of imports granted under FTA preference to value of imports eligible for FTA preference. Table 1 lists the actual rates of utilization of several major FTAs. The utilization rates vary considerably from 5.49% to 71.97%, and most of them are below 50%. For instance, only 20 percent of Korean firms import goods from ASEAN

countries under Korea-ASEAN FTA, and 13.45% of exports from Singapore to China utilize the preferential tariff rates granted by China-Singapore FTA. NAFTA and China-Pakistan FTA are relatively intensively utilized and the utilization rates are 55% and 71.79% respectively. But none of them are close to full utilization. In other words, within an FTA area, only a proportion of exporting firms choose to export under lower tariffs and there are a large amount of products are not engaged in the free trade arrangement.

Moreover, FTA utilization rates also vary considerably across industries. Under the US Generalized System of Preference (GSP), the average utilization rate across all industries is 60% in 2008. The industrial level data reveal that the utilization rate in the Electrical Machinery and Equipment industry is barely 22%, while the rate in the Iron and Steel industry is as high as 85%. The cross-industry variations suggest that factor intensity of production and procurement of intermediate inputs could be determining factors of firms' decisions in FTA participation.

In sum, the utilization rates of FTAs are low in general. Most trade within an FTA region takes place under the MFN tariff rates, rather than under the preferential tariff rates. Assuming full utilization would overestimate the benefits of FTAs on economic outcomes to a large extent. It is worth investigating the reasons for low utilization rates and bringing partial utilization to the related analysis of free trade agreements.

2.2 Rules of Origin hurle is high and varies across FTAs and across products

Free trade agreements are conditional trade policies. The common low utilization rates of FTAs around the world suggest that in addition to tariff benefits, the associated costs of FTA implementation are substantial. It is the additional costs that prevent eligible firms to utilize the preferential tariff rates. Rules of origin could be one of them. Rules of origin require that only the products which are actually produced in the FTA region are qualified for preferential tariff rates. Each free trade agreement assigns specific rules of origins for goods in each category. Table 2 lists the product specific rules stated in ASEAN (Annex 3) and NAFTA (Annex 401). There are three main methods of origin determination. The most strict one requires the products are wholly obtained or produced in the exporting member state. The second sets a minimal percent of local value content in the products. For instance, ASEAN FTA testifies the origin of a product if it contains 60 percent or more regional value content, and NAFTA requires a regional value content of not less than 60 percent where the transaction value method is used or 50 percent where the net cost method is used. The third method to confer origin is to show that the production taking place within the region makes substantial transformation to the products so that its tariff heading or subheading is changed. Considering the fact that most ASEAN countries undertake a great volume of production along the global value chain, the rules of origin in ASEAN FTA is relatively strict.

All in all, rules of origin are an important part of an FTA and vary across products and across FTAs. Strict rules of origin could be a reason for low FTA utilization rates, as it prevents goods that contain high non-regional value content from being eligible for preferential tariffs, especially in regions where countries have close vertical linkages in production with outside countries.

2.3 Administrative compliance and documentation costs offset tariff margin attractiveness

In addition to the restrictiveness of ROOs, the costs of complying with the procedures of origin certification are another reason for low FTA utilization rates. Rules of origin are often expensive to document. Exporters must obtain a certificate from its national government and present it to the customs authority of the importing government. Going through such an administrative procedure makes the FTA preferences less attractive. Medalla and Rosellon (2012) conduct a survey on the typical process for acquiring a certificate of origin. The pre-export verification requires documentations like company registration, business license, organization code, etc. Processing time from pre-export verification to issuance of certificate of origin ranges widely, from one working day (as for Australia and New Zealand) to not more than 30 working days (as for China and Brunei). Part of the survey results are reported in Table 3. It is the least costly to obtain a certificate for exporters producing in Australia and New Zealand. The pre-export verification can be done within one working day or automatically for electronic application. The entire processing time is within a day. In the case of Korea and Japan, it takes about three working days to complete the administrative procedure. For other countries, like Laos, China, and Malaysia, exporters have to wait five to thirty working days to get a certificate.

Large fixed costs are another hurdle for FTA utilization, including learning about FTA provisions and obtaining certificates of origin. Processing time for origin certification varies across countries. Small and medium size exporters are less able to muster the requisite financial and human resources than large firms, and therefore are less likely to utilize trade preferences. The tariff margin attractiveness granted by an FTA is, at least partially, offset by documentation costs.

3 The Model

3.1 Set-up

Consider a free trade area consists of two countries: country A and country B. There are many downstream firms producing final differentiated goods in both countries, and firm productivities follow the same distribution. The only difference between the final good producers in the two countries lies in the sources of intermediate inputs. Country A is assumed not to produce intermediate goods, and final good producers import intermediate goods from either country B or an outside country, depending on import prices. In country B, there are upstream firms supplying intermediate goods, and final good producers in that country choose to employ domestically produced inputs. Differentiated final products can be sold to the other country at a zero tariff rate if the products meet ROOs. As such, ROOs would affect the productions of exporters in country A only. Since this paper focuses on the use of FTA, I assume away fixed costs of producing and exporting, meaning that all firms are active exporters but export to the other country at different tariffs. Labor is inelastically supplied and perfectly mobile between upstream and downstream firms. The model

builds on Melitz (2003), to which I add intermediate goods, tariff preferences and fixed costs of using FTAs.

3.1.1 Preference

On the demand side, the preference of a representative consumer in country j ($j = A,\ B$) is given by a nested Cobb-Douglas utility function:

$$U_j = X_j^{\alpha} H_j^{1-\alpha} \tag{1}$$

where

$$X_j = \left[\int_{\Omega_j} x_{ij}(\varphi)^{1-1/\sigma} \, d\varphi \right]^{1/(1-1/\sigma)}.$$

Ω denotes the endogenous set of differentiated varieties sold in country j, $x_{ij}(\varphi)$ is quantity of variety φ produced in i and consumed in j, and the elasticity of substitution between any two varieties within this industry is constant and equals $\sigma > 1$. H_j represents the consumption of homogeneous goods in country j. Homogeneous goods can be traded freely across countries and its price is normalized to be 1. α is the standard Cobb-Douglas parameter which stands for the expenditure share on differentiated varieties. As in Dixit and Stiglitz (1977), the demand for variety φ is given by

$$x_{ij}(\varphi) = \frac{\alpha I_j}{P_{xj}^{1-\sigma}} p_{ij}(\varphi)^{-\sigma}, \tag{2}$$

where $p_{ij}(\varphi)$ is the price of variety φ produced in i and sold in j, I_j is the representative consumer's income in country j, and P_{xj} is a price index of differentiated goods in j such that $P_{xj} = [\int_{\varphi \in \Omega_j} p_{ij}(\varphi) d\varphi]^{1/(1-\sigma)}$.

Under monopolistic competition, the optimal price for each variety is a constant mark-up over unit cost. Hence, I have

$$p_{ij}(\varphi) = \frac{\sigma}{\sigma - 1} \tau_{ij} c_i(\varphi), \tag{3}$$

where $c_i(\varphi)$ denotes variety φ's variable production costs in country i. $\tau_{ij} > 1$ is gross tariff rate imposed by importing country j, and $\tau_{ii} = 1$.

3.1.2 Production in country A

Final goods X are produced using labor (L_a) and intermediate inputs. Since there are no upstream firms in country A, final good producers import intermediate inputs from country B (M_b) and/or an outside country C (M_c). The intermediate inputs originated from different countries are assumed to be perfect substitutes, and how intensively the firm uses the imported inputs depends on the ROOs and import prices. The production function takes a Cobb-Douglas form with constant returns to scale:

$$x_a(\varphi) = \varphi \left(\frac{M_b + M_c}{\beta} \right)^{\beta} \left(\frac{L_a}{1-\beta} \right)^{1-\beta},$$

where φ denotes a firm's total factor productivity. Final good producers are heterogeneous in the sense that they have different productivity draws from the Pareto distribution $G(\varphi) = 1 - \varphi^{-\varepsilon}$, where ε measures the inverse of dispersion. A large value of ε implies that firms are less heterogeneous and the market structure is more competitive. Firms produce horizontally differentiated final goods and each variety is indexed by $\varphi \in [1, +\infty]$. To ensure there exists a closed solution, I assume $\sigma < \varepsilon + 1$. β stands for the cost share of imported inputs in the production, which is independent of wage rate and prices of intermediate inputs.

Firms take the prices of inputs as given. Without any additional restrictions, each firm chooses labor and intermediate goods to minimize unit production costs:

$$\min_{\{M_b, M_c, L_a\}} p_{mb} M_b + p_{mc} M_c + w_a L_a$$

$$s.t.\ x_a(\varphi) \geqslant 1$$

Thus, the unconstrained unit production cost can be expressed as

$$c_a(\varphi) = \frac{p_{mc}^{\beta} w_a^{1-\beta}}{\varphi}. \tag{4}$$

For simplicity, I assume the outside country C is more efficient than country B in producing intermediate goods, so that $p_{mb} > p_{mc}$. The final good producers in country A initially choose to import inputs from country C. The unconstrained unit cost is decreasing in productivity and increasing in prices of inputs. Also, the cost share of foreign contents is β. Firms within one industry differ in productivity and unit cost, but spend the same percentage of production costs on intermediate inputs produced abroad.

Accordingly, the unconstrained unit demand for labor and the unconstrained unit demand for intermediate goods are:

$$l_a = \frac{1-\beta}{\varphi} \left(\frac{p_{mc}}{w_a} \right)^{\beta}, \tag{5}$$

$$M_c = \frac{\beta}{\varphi} \left(\frac{w_a}{p_{mc}} \right)^{1-\beta}. \tag{6}$$

3.1.3 Production in country B

In the intermediate good sector, labor is the only factor of production, and the products are sold in a perfectly competitive market. Thus, I have the price of intermediate goods produced in country B equals costs:

$$p_{mb} = \frac{w_b}{a_{mb}}, \tag{7}$$

where w_b is the wage rate in country B and a_{mb} is the labor efficiency in the production of intermediate goods in country B.

Following the same production function as in country A, firms use labor and domestically produced intermediate inputs to produce final products. The variable cost of producing 1 unit of output is given by

$$c_b(\varphi) = \frac{p_{mb}^{\beta} w_b^{1-\beta}}{\varphi}. \tag{8}$$

As shown in the cost equation, all inputs are sourced from domestic suppliers. Thus, the final goods produced in country B automatically satisfy the rule of origin, no matter how strict it is. All the exporters producing in country B would utilize the FTA preferences and export to country A under zero tariff rates.

Moreover, country B also produces homogeneous goods, which are the numeraire goods in the model. Producing 1 unit of homogeneous good requires1 unit of labor, and homogeneous goods can be traded freely across countries. Hence, the wage rate in country B is normalized to be 1.

3.2 Utilizing FTAs

The preferences granted by the free trade agreement are conditional. To be qualified for zero tariffs, final varieties must satisfy the ROO. Suppose the ROO requires the cost share of local contents to be no less than γ and $0 < \gamma < 1$. As such, the cost minimization problem for the firm with productivity φ is subject to an additional constraint regarding the source of the inputs:

$$\min_{\{M_b, M_c, L_a\}} p_{mb} M_b + p_{mc} M_c + w_a L_a$$

$$s.t.\ x_a(\varphi) \geqslant 1$$

and

$$\frac{p_{mc} M_c}{p_{mb} M_b + p_{mc} M_c + w_a L_a} \leq 1 - \gamma.$$

Then, the constrained unit production cost can be derived as

$$c_{ac}(\varphi) = \lambda c_a(\varphi) \tag{9}$$

where

$$\lambda = \begin{cases} 1, & if\ \beta \leqslant 1 - \gamma \\ \left[\frac{\beta}{(\beta+\gamma-1)\frac{p_{mc}}{p_{mb}} + 1 - \gamma} \right]^{\beta}, & if\ \beta > 1 - \gamma. \end{cases} \tag{10}$$

λ measures the rise in unit production costs due to complying with ROO. If $\beta \leqslant 1-\gamma$, the ROO is not binding. In this case, firms' productions involves more domestic contents than the requirement, therefore the second constraint of the optimization problem is automatically satisfied and firms will stay with their optimal production plans. Comply with ROO does not generate additional production costs, so $\lambda = 1$. If $\beta > 1-\gamma$, the ROO is binding. That is, firms use more foreign intermediate goods than the requirement. In this case, if firms intend to utilize the FTA preferences, they need to deviates from their optimal production strategies by using more local inputs to meet the ROO. Hence, $\lambda > 1$, and it increases with β and γ. The stricter the origin rule is, the more the unit cost will increase by.

The corresponding constrained demands for labor and intermediate inputs are given by

$$l_{ac}(\varphi) = \lambda l_a(\varphi), \tag{11}$$

and

$$M_{cc} = M_c C \quad and \quad M_b = \frac{\beta}{\varphi}\left(\frac{w_a}{p_{mb}}\right)^{1-\beta} B, \tag{12}$$

where $C = \left(\frac{1-\gamma}{\beta}\right)^{1-\beta}\left[\frac{1}{\left(\frac{\beta+\gamma-1}{1-\gamma}\right)\left(\frac{p_{mc}}{p_{mb}}\right)+1}\right]^{\beta} < 1$ and $B = \frac{\beta+\gamma-1}{\beta}\left[\frac{\beta}{\beta+\gamma-1+(1-\gamma)\left(\frac{p_{mb}}{p_{mc}}\right)}\right]^{\beta} > 1$.

Since the ROO requires more intensive use of locally produced inputs, the constrained demands for labor and the intermediate goods from country B are greater than in the unconstrained case, and they increase in the restrictiveness of ROO (γ). In contrast, the demand for the intermediate goods originating from country C decreases as a binding ROO takes effect on firm production.

3.3 Partial equilibrium

Whether a firm chooses to utilize the preferential tariffs or not depends on the associated benefits and costs. On the one hand, a firm would charge a lower price and make more sales in the foreign market due to the tariff reductions. On the other hand, the distorted sourcing of inputs generates extra production costs. Moreover, firms have to pay fixed documentation costs (f_d) in order to obtain a certificate of origin and claim the FTA preferences.

Assume there are no fixed costs of production and exporting. If the firm exports under the MFN tariffs, it needs to pay a per unit tariff $\tau_{ij} > 1$ to access the foreign market. The profits are revenues less costs in the domestic and the foreign markets. In particular, the profits of an FTA non-user can be expressed as:

$$\pi_1 = \pi_{ad} + \pi_{an} = \frac{1}{\sigma} p_{ad} x_{ad} + \frac{1}{\sigma \tau_{ij}} p_{an} x_{an}, \tag{13}$$

where π_{ad}, p_{ad}, and x_{ad} are firm's profits, price, and quantity sold in the domestic market, and π_{an}, p_{an}, and x_{an} are firm's profits, price, and quantity sold in the foreign market as an FTA non-user.

If a firm chooses to utilize the FTA preferences, it pays zero tariff, constrained unit costs, and a positive fixed documentation cost. Hence, the corresponding profits of an FTA user is given by

$$\pi_2 = \pi_{ad} + \pi_{af} = \frac{1}{\sigma} p_{ad} q_{ad} + \frac{1}{\sigma} p_{af} q_{af} - f_d w_a, \tag{14}$$

where fixed documentation costs are paid in units of labor, and π_{af}, p_{af}, and x_{af} are the firm's profits, price, and quantity sold in the foreign market as an FTA user. Comparing the profits of FTA users and no-users, there is a tradeoff between optimizing the source of intermediate inputs and having access to preferential tariffs, given that $\lambda \geqslant 1$ and $\tau_{ij} > 1$. Exporters would enjoy a zero tariff at the expense of using more expensive inputs produced domestically. I assume $\lambda < \tau$, the increase in unit cost due to ROO is smaller than the tariff margin. Otherwise, no one would utilize the preferential tariffs offered by FTA.

From the individual firms' perspective, it is the increased profits that motivate them to join the FTA. Firms are willing to take advantage of zero tariffs if and only if $\pi_2 \geqslant \pi_1$. The productivity threshold for utilizing FTA preferences should be the value which makes the firm indifferent between being and not being an FTA user. That is, $\pi_1(\varphi^*) = \pi_2(\varphi^*)$. Combining the conditions I have derived above, the critical value is determined by

$$\varphi^* = \left[\frac{\kappa f (w_a)^{\sigma-1} f_d}{\lambda^{1-\sigma} - \tau^{1-\sigma}} \right]^{\frac{1}{\sigma-1}} \tag{15}$$

where $\kappa = \frac{\sigma^\sigma}{(\sigma-1)^{\sigma-1}} \frac{1}{P_{xb}^{\sigma-1} I_b}$ and $f(w_a) = p_{mc}^{\beta} w_a^{1-\beta}$. κ represents the market demand condition in the importing country, which is held constant in partial equilibrium analysis. A small value of κ is associated with a large foreign market size (I_b) and a low price index (P_{xb}), both of which indicate a strong market demand for variety φ. Similar to Melitz (2003), firms with different productivities end up with different sales patterns. Only firms with productivities above the cutoff φ^* invoke the preferential tariffs while exporting. If $\lambda^{1-\sigma} - \tau^{1-\sigma} \geqslant \kappa f(w_a)^{\sigma-1} f_d$, the FTA will be fully utilized. All firms export under zero tariffs and make more revenues due to a lower trade barrier: $r_2(\varphi) - r_1(\varphi) \geqslant (\frac{\sigma}{\sigma-1})^{1-\sigma} \frac{I_b}{P_{xb}^{1-\sigma}} c_a(\varphi)^{1-\sigma} \kappa f_d > 0$, where $r_1(\varphi)$ and $r_2(\varphi)$ refer to export revenues by not using and using FTA. If $\lambda^{1-\sigma} - \tau^{1-\sigma} < \kappa f_d$, $\varphi^* > 1$ follows. The FTA is partially utilized. Firms with productivities smaller than the cutoff would export under the MFN tariffs. The increase in exporting revenues induced by claiming the zero tariffs is still positive, but it is much less than the revenue gains in the case with full utilization, since it can be shown: $0 < r_2(\varphi) - r_1(\varphi) < \frac{I_b}{P_{xb}^{1-\sigma}} c_a(\varphi)^{1-\sigma} \kappa f_d$. Therefore, assuming full utilization of trade preferences may overstate the role of FTAs in promoting bilateral trade.

The FTA utilization rate is the fraction of export value that takes advantage of the trade agreement and pays lower tariffs. In this model, it is determined by the productivity threshold of using FTA (φ^*):

$$
\begin{aligned}
u &= \frac{\int_{\varphi^*}^{\infty} r_2(\varphi) dG(\varphi)}{\int_1^{\varphi^*} r_1(\varphi) dG(\varphi) + \int_{\varphi^*}^{\infty} r_2(\varphi) dG(\varphi)} \\
&= \frac{1}{\left(\frac{\tau}{\lambda}\right)^{1-\sigma} \left[(\varphi^*)^{\varepsilon-\sigma+1} - 1\right]}
\end{aligned} \tag{16}
$$

As shown in (16), the larger the threshold is, the lower the utilization rate is. Thus, if an FTA sets a large tariff margin, a loose ROO and simple administrative procedures, there will be a low productivity threshold of utilizing FTA preferences and a high utilization rate. Also, the utilization rate is relatively high in the industries where imported inputs account for a small cost share and it is less costly to comply with ROOs.

4 General Equilibrium

Welfare analysis is approachable in a general equilibrium framework. In this section, I present a general equilibrium model to investigate the welfare effects of ROOs and use GAMS to simulate the results. Now consider a world comprised of three countries (A, B, and C). Two of them, country A and country B, sign a free trade agreement in which countries permit zero import tariffs to exports from the partner country if the products satisfy the ROO. Exporters producing in country A are affected by ROOs. They have to reduce the use of inputs imported from country C in order to be qualified for tariff benefits when the ROO is restrictive. Firms have different productivities in producing differentiated final goods and they select themselves into profitable sales patterns. Homogeneous goods are traded costlessly across countries in order to keep trade balanced. Since the firm heterogeneity complicates the computation, I adopt the representation of the average firm operating in each sale pattern, which is proposed by Balistreri and Rutherford (2011) and simplifies the model effectively.

4.1 Country A

Country A is one of the participants of the free trade agreement. The labor endowment is denoted as L_a. Workers in country A produce final differentiated goods. The documentation costs and the firm entry costs are paid in units of labor. Workers are also consumers, and the utility function is given by (1).

4.1.1 Production of final goods

As shown in the previous section, more productive firms choose to comply with the ROO and export under zero tariffs. I assume the number of FTA non-users is n_1 who stick with their original production strategies and pay tariffs while exporting, and the rest n_2 firms are FTA users. The

average productivity of FTA non-users is denoted by $\tilde{\varphi_1}$ and the average productivity of FTA users is $\tilde{\varphi_2}$. Thus, the unit production costs for two types of firms are

$$c_1(\tilde{\varphi_1}) = \frac{p_{mc}^{\beta} w_a^{1-\beta}}{\tilde{\varphi_1}} \tag{17}$$

and

$$c_2(\tilde{\varphi_2}) = \frac{\lambda p_{mc}^{\beta} w_a^{1-\beta}}{\tilde{\varphi_2}}. \tag{18}$$

Under monopolistic competition, heterogeneous firms set a constant markup over the unit costs and the optimal prices for two groups of firms are given by

$$\tilde{p_{1j}} = \frac{\sigma \tau_{ij}}{\sigma - 1} c_1(\tilde{\varphi_1}) \tag{19}$$

and

$$\tilde{p_{2j}} = \frac{\sigma}{\sigma - 1} c_2(\tilde{\varphi_2}). \tag{20}$$

Accordingly, the demands for the average varieties in the two countries are

$$\tilde{q_{1j}} = \frac{\alpha I_j \tilde{p_{1j}}^{-\sigma}}{P_{xj}^{1-\sigma}}, \tag{21}$$

and

$$\tilde{q_{2j}} = \frac{\alpha I_j \tilde{p_{2j}}^{-\sigma}}{P_{xj}^{1-\sigma}}, \tag{22}$$

where I_j represents consumers' total income in country j and P_{xj} stands for the price index of differentiated goods in country j. The corresponding profits earned by firms can be expressed by

$$\tilde{\pi_{1j}} = \frac{1}{\sigma} \tilde{p_{1j}} \tilde{q_{1j}} \tag{23}$$

and

$$\tilde{\pi_{2j}} = \frac{1}{\sigma} \tilde{p_{2j}} \tilde{q_{2j}} - f_d w_a. \tag{24}$$

The indifference condition for the marginal firm can be derived in terms of the average FTA user's revenues and the parameters by linking the average FTA user's and nonuser's productivities and revenues to the marginal firm through the Pareto distribution. The equivalent condition to the cutoff condition (15) is given by the equation below:

$$\frac{1}{\sigma} \frac{\varepsilon + 1 - \sigma}{\varepsilon} \left[1 - \left(\frac{\tau}{\lambda} \right)^{1-\sigma} \right] \tilde{p_2} \tilde{q_2} = f_d w_a. \tag{25}$$

Moreover, the free entry condition drives firms' profits down to zero:

$$\sum_{j=A,B} \tilde{\pi_{1j}} = \sum_{j=A,B} \tilde{\pi_{2j}} = w_a f_e \tag{26}$$

The productivity of the marginal exporter and the fraction of FTA users can be linked by $n_2 = 1 - G(\varphi^*)$. With this relationship, I can express the average productivities of nonusers and users in terms of the number of exporters using FTA as

$$\tilde{\varphi_1} = \left[\frac{\varepsilon}{\varepsilon + 1 - \sigma} \frac{1 - n_2^{\frac{\varepsilon+1-\sigma}{\varepsilon}}}{1 - n_2} \right]^{\frac{1}{\sigma-1}} \tag{27}$$

and

$$\tilde{\varphi_2} = \left(\frac{\varepsilon}{\varepsilon + 1 - \sigma} \right)^{\frac{1}{\sigma-1}} n_2^{-\frac{1}{\varepsilon}}. \tag{28}$$

4.1.2 Labor market and income balance

The labor market clearing condition is

$$L_a = n_1 l_a \sum_{j=A,B} \tilde{q_{1j}} + n_2 \left(\tilde{q_{2a}} l_a + \tilde{q_{2b}} l_{ac} \right) + n_a f_e + n_2 f_d \tag{29}$$

where $n_a = n_1 + n_2$ is the total number of firms producing in country A, and n_1 and n_2 are the number of FTA non-users and users respectively. f_e is the units of labor required to hire to enter the market. Labor demand consists of four parts: the demand by FTA non-users and users to produce outputs, the demand by all active firms to pay entry costs, and the demand by FTA users to pay documentation costs.

Following the free entry condition, consumer's total income is wage payment. Therefore, the income balance condition is

$$I_a = w_a \bar{L_a}. \tag{30}$$

4.2 Country B

Country B, the second participant in the FTA, produces and exports to country A differentiated goods and homogeneous goods H ($p_H = 1$). The labor endowment is L_b. Consumers' utility is derived over a continuum of differentiated varieties and homogeneous goods.

All exports originating from country B satisfy the ROO and claim preferential tariffs. The free entry condition for firms producing in country B implies that

$$\tilde{\pi_b} = \frac{1}{\sigma} \tilde{p_b} \sum_{j=A,B} \tilde{q_{bj}} = f_b w_b, \tag{31}$$

where $\tilde{\pi_b}$, $\tilde{p_b}$, and $\tilde{q_{bj}}$ are the profits, price, and quantity of the average exporter in country B. f_b is the entry costs in country B and is paid by labor. Since all exports from country B are qualified for zero tariffs, firms charge the same price in the domestic and the foreign markets.

The labor market clearing condition equates labor supply and demand:

$$L_b = n_b l_b \sum_{j=A,B} \tilde{q_{bj}} + n_b f_b + \frac{n_2 \tilde{q}_{af} M_b}{a_{mb}} + h_{bp}, \tag{32}$$

where n_b is total number of exporters from country B, l_b is the labor demand by producing 1 unit of differentiated good, and h_{bp} is the quantity of homogeneous goods produced in country B. Country B conducts three production activities. The first two terms on the right hand side of (32) indicate the labor demand by firms producing final goods. The third term is the labor working in the intermediate good sector, and the last term is the labor demand in the production of homogeneous goods.

Country B's total income comes from two sources. One is labor income by producing differentiated and homogeneous goods, and the other is import tariff revenues imposed on country A's FTA non-users. The income balance condition is

$$I_b = w_b \bar{L_b} + \frac{\tau - 1}{\tau} n_1 \tilde{p_{1b}} \tilde{q_{1b}}. \tag{33}$$

4.3 Country C

Country C is the outside country. The labor endowment is L_c. People living country C produce intermediate inputs (M_c) and export them to country A. The intermediate inputs are produced using labor only and sold in a perfectly competitive market at the price of $p_{mc} = \frac{w_c}{a_{mc}}$. Country C also produces homogeneous goods, and therefore the wage rate is normalized to be one. There are no consumption and production of differentiated goods in country C. Homogeneous goods are the only consumption goods.

The labor supply-equal-demand condition yields

$$L_c = n_1 M_c \sum_{j=A,B} \tilde{q_{1j}} + n_2 M_c \tilde{q_{2a}} + n_2 M_{cc} \tilde{q_{2b}} + h_{cp}, \tag{34}$$

where $\tilde{q_{2a}}$ and $\tilde{q_{2b}}$ are the quantities of outputs sold by the average FTA user in country A and country B, and h_{cp} is the amount of homogeneous goods produced in country C. A binding ROO forces FTA users to switch from foreign to local intermediate inputs so that the unit demand for intermediate goods originating from country C by FTA users is smaller than by non-users $M_{cc} < M_c$.

The total income can be expressed as

$$I_c = L_c w_c. \tag{35}$$

4.4 Closing the model

The trade and payment systems are closed by trade balance conditions. Country A produces and exports differentiated goods to country B under either zero or the MFN tariff rates. It imports intermediate inputs from country B and country C. Country A's imports of final products from

country B are under zero tariffs. To keep trade balanced, country A also imports homogeneous goods. The trade balance condition equates value of exports with value of imports:

$$n_1 \frac{\tilde{p_{1b}} \tilde{q_{1b}}}{\tau} + n_2 \tilde{p_{2b}} \tilde{q_{2b}} = n_b \tilde{p_b} \tilde{q_{ba}} + IM_{mc} + IM_{mb} + p_H h_{aim} \tag{36}$$

where

$$IM_{mc} = p_{mc} \left(n_1 M_c \sum_{j=A,B} \tilde{q_{1j}} + n_2 M_c \tilde{q_{2a}} + n_2 M_{cc} \tilde{q_{2b}} \right)$$

$$IM_{mb} = p_{mb} n_2 \tilde{q_{2b}} M_b.$$

IM_{mc} and IM_{mb} represent the value of imported intermediate inputs from country C and country B respectively. h_a is the consumption of homogeneous goods in country A. The left hand side of (36) is the exporting revenues earned by FTA non-users and users, and the right hand side is the sum of import values.

Country B makes export revenues by selling differentiated goods and intermediate inputs to country A under zero tariffs as well as homogeneous goods. The tariff payments from country A's FTA non-users are redistributed to country B's consumers. The trade balance condition for country B is

$$n_b \tilde{p_b} \tilde{q_{ba}} + \frac{\tau - 1}{\tau} n_1 \tilde{p_{1b}} \tilde{q_{1b}} + EX_{mb} + p_H h_{bex} = n_1 \tilde{p_{1b}} \tilde{q_{1b}} + n_2 \tilde{p_{2b}} \tilde{q_{2b}}, \tag{37}$$

where $EX_{mb} = IM_{mb}$ is the export value of intermediate inputs from country B to country A. h_{bex} is quantity of homogeneous goods exported from country B, which could be negative if country B imports homogeneous goods.

Country C consumes homogeneous goods only and conduct productions of intermediate inputs and homogeneous goods. The trade balance conditions comes as

$$EX_{mc} = p_H h_{cim}, \tag{38}$$

where $EX_{mc} = IM_{mc}$ stands for the value of intermediate goods exported from country C to country A. h_{cim} is the amount of homogeneous goods imported by country C to keep trade balanced. Also, h_{cim} should be equal to the difference between country C's consumption and production of homogeneous goods.

Lastly, the world demand and supply of homogeneous goods are equalized. The market clearing condition for the homogeneous goods is

$$h_{cim} + h_{aim} = h_{bex}. \tag{39}$$

4.5 Equilibrium

So far, every dollar paid is a dollar of revenue earned. Given σ, α, β, τ, a_{mc}, a_{mb} and productivity distribution $G(\varphi)$, an equilibrium is a set of numbers of firms ($\dot{n_1}$, $\dot{n_2}$, $\dot{n_b}$), aggregate price statistics (P_{xa}, P_{xb}), wage rate w_a, consumer allocations $\tilde{q_{ij}}$ ($i = 1, 2, b$ and $j = a, b$), firm pricing rules $\tilde{p_b}$ and $\tilde{p_{ib}}(i = 1, 2)$, firm profits $\tilde{\pi_{ij}}$ ($i = 1, 2, b$ and $j = a, b$), and homogeneous good trade value (h_{aim}, h_{bex}, h_{cim}), such that: (i) $\tilde{q_{ij}}$ is given by (21) (22) and solves the representative consumer's problem; (ii) $\tilde{p_b}$ and $\tilde{p_{ib}}$ are given by (19) (20) and solves the firm's problem; (iii) $\tilde{\pi_{ij}}$ is given by (23)(24); (iv) n_1, n_2, n_b, P_{xa}, P_{xb}, w_a, $\tilde{q_{ij}}$, h_{aim}, h_{bex}, h_{cim} jointly satisfy (25)-(39).

There are no analytical solutions to this complicated system of equations. I adopt two approaches to make a progress. First, I conduct comparative static exercise in the neighborhood of full utilization and show the local analytical results. Second, I take advantage of GAMS program to provide global numerical solutions.

5 Local Analytical Solutions and Welfare Analysis

In this section, I conduct comparative static exercise in the neighborhood of full utilization. The results show the local effects of a slightly binding rule of origin on market outcomes and welfare.

Consider the equilibrium in which the ROO is exactly satisfied by the Cobb-Douglas cost share of imported inputs, that is $\beta = 1 - \gamma$. Now if there is a slight increase in the local content requirement (ROO), the ROO starts being binding. The producers in country A are faced with choices of whether to utilize the trade preferences. The endogenous variables of the model would change accordingly.

5.1 Number of producers

The percentage change of number of FTA users in country A can be expressed as

$$\hat{n_2} = -\frac{\sigma - 1}{x} \underbrace{\left[\frac{1}{(\beta + \sigma(1-\beta))(1 + y/\tau^{\sigma}) s_f} + \frac{\tau^{\sigma}(1 + 1/y)}{\tau^{\sigma} - 1} \frac{1}{s_f} - 1 \right]}_{(+)} \hat{\lambda}, \tag{40}$$

where $x = \left(\frac{f_d}{f_e} + 1\right)\left[\alpha\tau^{-\sigma}(\tau - 1) - \tau^{1-\sigma} + 1\right] > 0$ and $y = \frac{f_e}{f_d}(\tau^{\sigma} - 1) - 1 > 0$. $s_f = \frac{n_2 \tilde{p}_{2b}^{1-\sigma}}{P_{xb}^{1-\sigma}}$ is the market share of exports by FTA users in country B. $\hat{n_2} = \frac{dn_2}{n_2}$ and $\hat{\lambda} = \frac{d\lambda}{\lambda}$ represent the percentage changes of number of FTA users and ROO. The effect of a slightly binding ROO on firm production is fully reflected in the change of variable costs λ. Thus, (40) states a negative relationship between number of FTA users and the restrictiveness of ROO. As λ increases, n_2 decreases. The total number of exporters originating from country A is negatively correlated with the number of FTA users, since I have

$$\hat{n_a} = -\frac{f_d}{f_e} \hat{n_2}. \tag{41}$$

Considering that the number of FTA non-users is equal to the total number less the number of users ($n_1 = n_a - n_2$), a stricter ROO leads to an increase in total number of exporters producing in country A, an increase in number of FTA non-users, and a decrease in number of FTA users. This is consistent with the intuition. When the ROO requires more local inputs, exporters in country A have to alter their productions and produce at a higher cost. The trade preferences are costly to utilize, which results in an exit from being FTA users.

par As for the change in number of exporters producing in country B, it can be seen from

$$\hat{n_b} = \frac{\sigma - 1}{1 - s_f} \left[\frac{\sigma + (1-\sigma)s_f}{(1 + y/\tau^\sigma)(\sigma + \beta/(1-\beta))} + 1 - \frac{\tau^\sigma}{\tau^\sigma - 1} \right] \hat{\lambda}. \tag{42}$$

Given that $\sigma > 1$, $s_f < 1$, and the sign of the bracket is positive, $\hat{n_b}$ and $\hat{\lambda}$ are positively correlated. As the ROO becomes tighter, more producers in country B start exporting to country A. This is due to the fact that a higher requirement of local content disadvantages producers in country A. More exporters from country B benefit from the preferential tariffs and make sales in the foreign market.

5.2 Wage rate

In such a framework, country A can be thought of Mexico in NAFTA. An interesting and important question is how the implementation of NAFTA and ROOs affects the wage rate in Mexico. The local comparative static analysis shows that

$$\hat{w_a} = -\frac{\sigma - 1}{(1 + y/\tau^\sigma)(\beta + \sigma(1-\beta))} \hat{\lambda}. \tag{43}$$

When the ROO increases, there are two opposing factors that influences labor demand. On the intensive margin, FTA users demand more labor and local inputs to meet the ROO. On the extensive margin, the trade preferences will be utilized by fewer exporters. For those who exported under preferential tariffs before, an exit from the FTA framework leads to a reduction in the demand for local inputs, since they no longer need to comply with the rule. (43) shows that the overall effect is negative. An increase in ROO (λ) leads to a decrease in the wage rate of the downstream country (w_a).

5.3 Price index

The change in ROO reallocates sales of heterogeneous firms. As a result, the price indexes in the FTA participating countries change. Both domestic firms and exporters from country B sell in country A. The price index in country A is given by

$$P_{xa}^{1-\sigma} = n_a \tilde{p}_a^{1-\sigma} + n_b \tilde{p}_b^{1-\sigma}.$$

Accordingly, the percentage change in price index caused by the change in ROO is

$$\hat{P_{xa}} = \underbrace{-\frac{1-s_b}{\sigma-1}\hat{n_a} - \frac{s_b}{\sigma-1}\hat{n_b}}_{variaty\,effect(-)} \underbrace{+(1-s_b)\,\hat{\tilde{p}_a}}_{price\,effect(-)}, \tag{44}$$

where $s_b = \frac{n_b \tilde{p}_b^{1-\sigma}}{P_{xa}^{1-\sigma}}$ denotes the market share of exports from country B in country A. The effect of ROO on price index can be decomposed into two parts. One is variety effect in that a stricter ROO leads more firms to start producing in and exporting to both countries, which drives down the price index and intensifies market competition. The other effect is price index. The domestic producers in country A would charge lower prices following a binding ROO, because the wage rate goes down, as discussed in the previous section. Hence, both variety effect and price effect are negative. An increase in local content requirement reduces the price index in country A.

Similarly, in country B, there are three types of producers supplying differentiated goods. The associated price index is

$$P_{xb}^{1-\sigma} = n_1 \tilde{p}_{1b}^{1-\sigma} + n_2 \tilde{p}_{2b}^{1-\sigma} + n_b \tilde{p}_b^{1-\sigma}.$$

Total differentiating with respect to rules of origin yields

$$\hat{P_{xb}} = \underbrace{-\frac{1-(1+f_d/f_e)\,\tau^{1-\sigma}}{\sigma-1} s_f \hat{n_2}}_{(+)} + (1-s_f)\underbrace{\frac{\hat{n_b}}{1-\sigma}}_{(-)} + \underbrace{s_f \hat{\tilde{p}_{2b}}}_{(+)}. \tag{45}$$

The sign of the percentage change in price index in country B is ambiguous, depending on the market share of country A's exporters in country B. If the export sales of FTA users account for a large market share in country B, the positive variety effect driven by fewer FTA users and the positive price effect dominate. Thus, country B's price index P_{xb} increases in ROO. If the export sales of FTA users account for a small market share in country B, the negative variety effect due to the firm entry in country B outweighs the other two effects, which decreases the price index P_{xb} as the ROO rises.

5.4 Welfare

Country welfare is jointly determined by income and price index. A country with more income and a lower price index has a higher welfare level. Specifically, country A's welfare change due to a binding ROO is

$$\hat{Welfare_a} = \underbrace{\left(\frac{1}{1-\beta} - (1-s_b)\right)\hat{\tilde{p}_a}}_{TOT\,(-)} + \underbrace{\frac{1-s_b}{\sigma-1}(\hat{n_b} + \hat{n_a})}_{variety\,effect(+)}. \tag{46}$$

As shown in (46), the terms of trade effect and the variety effect run in the opposite directions. On the one hand, country A's exports become cheaper relative to its imports from country B, because a binding ROO generates distortions in the labor market, which reduces the wage rate. On

the other hand, there are more varieties sold in country A, since an increase in ROO encourages firm entry. It can be shown that if $\frac{\tau^\sigma}{\lambda^\sigma - 1} > 1 + \frac{f_d}{f_e}$ is satisfied and the elasticity of substitution between varieties is small, the variety effect dominates. In other words, if the FTA is under-utilized and varieties are less substitutable, country A's welfare rises with the restrictiveness of ROO, as a wider range of varieties are sold in the market. If the elasticity of substitution between varieties is large, the magnitude of variety effect is small. In such a case, the overall effect of ROO on country A's welfare is negative, with the terms of trade effect more than offsetting the variety effect.

The percentage change in country B's welfare due to a tighter ROO can be expressed as

$$\hat{Welfare_b} = \frac{(\tau - 1)\, y}{\tau^\sigma s_f} \hat{n_1} + (1 - s_f) \frac{\hat{n_b}}{\sigma - 1} - s_f \hat{\tilde{p}_{2b}} + \frac{1 - (1 + f_d/f_e)\, \tau^{1-\sigma}}{\sigma - 1} s_f \hat{n_2}. \tag{47}$$

The analysis of country B's welfare involves three channels. In terms of income, as the ROO become tighter, country B's tariff revenues increase, since more of exporters from country A choose to be FTA non-users and pay import tariffs. In the meantime, country's B wage rate is normalized to be 1. Hence, the income effect is positive. The second channel is through the terms of trade effect. Firms producing in country B are not affected by ROO and the wage rate is fixed, therefore, the prices of exports remain the same. But country A's production costs drop, which results in a decrease in the average price of its exports. Facing a higher relative price of exports to imports, country B's welfare rises. Lastly, the sign and the magnitude of variety effect vary with the elasticity of substitution between varieties as well as the market share of country A's exports in country B. Thus, the overall effect of ROO on country B's welfare is ambiguous.

As we can see in (47), total differentiation yields a complicated expression for the percentage change in welfare. There are three types of firms supplying differentiated goods in country B, and the prices and the numbers of these firms change with ROO. Based on the analytical solution, I can hardly conclude the direction towards which welfare moves following a stricter ROO. I resort to numerical approach to pin down the relationship between country welfare and ROO.

6 Simulation results

In this section, I take advantage of GAMS program to show global numerical solutions to the model. The results present the effects on key endogenous variables over a full range of ROOs. In the benchmark, I assume the parameters take on the following values: the shape parameter of Pareto distribution $\varepsilon = 5$, the elasticity of substitution between varieties $\sigma = 4$, the expenditure share on differentiated varieties $\alpha = 0.8$, the cost share parameters of Cobb-Douglas production function $\beta = 0.8$, the local content requirement (ROO) $\gamma = 0.2$, the gross import tariffs $\tau = 1.3$, the benchmark wage rates in three countries are equal $w_A = w_B = w_c = 1$, and the efficiency parameters of producing intermediate inputs in countries B and C $a_{mb} = 0.9$ and $a_{mc} = 1$. The fixed documentation costs, fixed entry costs in countries A and B, and the labor endowments in three countries can be calibrated from the model.

Table 4 reports the results for the benchmark. The free trade agreement offers a 30% tariff

reduction to final varieties whose cost shares of domestically produced inputs are no less than 80%. In the benchmark, since $\beta = 1 - \gamma$, the rule of origin is not binding. There are 20 firms exporting from A to B, and all of them utilize the preferential tariffs. The FTA utilization rate is 100 percent. The number of exporters originating from country B is 20. The outside country C is more efficient in producing intermediate inputs. When the ROO is loose, firms in country A source inputs from country C only. The import value of intermediate inputs from C to A is 11.62, and the import value of intermediate inputs from B to A is 0. The benchmark welfares in countries A and B are 10.35 and 26.72 respectively.

Next, I conduct counterfactual exercises to investigate the impacts of ROOs on country's welfare over the range from 1% to 95%. Figure 1 shows the simulation results on the use of FTA by country A exporters. In panel (a), the number of FTA non-users is 0 when the ROO is 20% or below. It rises from 0 to 37.63 gradually when the ROO increases from 20% to 59%. Since all firms choose to be non-users if the local content requirement is 59% or even higher, the number of FTA non-users remain the constant when the ROO is greater than 59%. In contrast, the number of FTA users moves in the opposite direction. There is a monotonic decreasing trend over the range of ROO from 20% to 59%, as shown in panel (b). As it is more and more costly for firms to comply with the ROO, fewer exporters find profitable to utilize the trade preferences. Panel (c) displays the relationship between FTA utilization rate and ROO. When the ROO is loose, the utilization rate is 100 percent. As the ROO increases, the utilization rate decreases monotonically. Once the local content requirement is too restrict (beyond 59%), none of the exporters take advantage of the preferential tariffs and the utilization rate is zero.

Figure 2 shows how ROOs affect the trade value of intermediate goods between countries. In panel (a), there is a non-monotonic relationship between imports of intermediate inputs from country B to country A. Since country B's intermediate inputs are more expensive than those produced in country C, firms in country A choose to import from country C when the ROO is not binding. As the ROO becomes stricter, country A producers have to import intermediate inputs from country B in order to satisfy the ROO criterion. Thus, the trade value between country A and B goes up. However, a further increase in ROO generates large distortions in manufacturing process and therefore makes the conditional trade preferences less attractive. As a result, more and more firms choose to be FTA non-users, and the trade in intermediate goods within the FTA area declines. When the ROO is greater than 59%, no firms use inputs produced in country B, and the trade value goes back to zero. Panel (b) shows how the trade value between countries A and C changes with ROO. The highest value occurs when the ROO is smaller than or equal to 20%. This is the case when all firms in country A use the cheaper inputs produced in country C. As the ROO increases, firms demand more locally produced inputs and the import from C to A goes down. The trade value between A and C start to rise when the ROO is too costly to comply with, and eventually it stays at a constant level which is lower than the unbinding case. This is because exporters from A make less sales in country B as they pay import tariffs and charge higher prices. Consequently, firms demand less inputs and produce fewer outputs.

Figure 3 gives the relationship between the wage rate of the downstream country and ROOs. A stricter rule of origin generates more demand for locally produced intermediate goods, but less demand for labor in the downstream country. As discussed in the section of local analytical solutions, there are two opposing factors that affect labor demand. The decrease of the wage rate with ROOs implies that the negative extensive margin dominates. The labor released from user to non-user firms is more than the labor demanded by a stricter ROO compliers.

Figures 4 and 5 report how country welfare changes with ROOs. The results differ by the value of elasticity of substitution between varieties. Figure 4 corresponds to the case where $\sigma = 4$, and Figure 5 refers to the case where $\sigma = 8$. If the substitution elasticity is relatively small, as shown in Figure 4, both countries experience welfare increases as the ROO increases. As the FTA utilization rate ranges from 0 to 100 percent, country A's welfare increases by 4.83% (Panel (a)) and country B's welfare increases by 16.79% (Panel (b)) . When the varieties are less substitutable, consumers benefit more from the introduction of new varieties. Resources are allocated to the more efficient producers in country B and a tighter ROO induces the production of intermediate goods to move from country C to country B. As such, the variety effect for both countries are positive and significant, which is more than offsetting the negative terms of trade effect. Hence, a stricter ROO effectively protects regional producers and promotes within FTA area trade. Such conditional trade preferences benefit both participating countries.

Country welfare exhibits different patterns when the elasticity of substitution is large ($\sigma = 8$). In Figure 5, when the ROO starts binding (greater than 20%), country A's welfare decreases, while country B's welfare increases slightly. Following a stricter ROO, the wage rate in country A drops and therefore the income effect is negative. The price index depends on the prices of individual varieties as well as the number of varieties available in the market. A lower wage rate reduces the production costs of firms producing in country A. The prices of domestically produced differentiated goods become cheaper. Considering that the prices of imported products from country B are not affected by ROOs, the average price decreases with ROOs. In the meantime, firm entry introduces new varieties, which would further reduce the price index. But this variety effect turns to be small when the elasticity of substitution between varieties is large. The simulation results show that the income effect is more negative and the overall effects of ROOs on country A's welfare are negative.

In Figure 5 Panel (b), country B's welfare increases and then decreases with the restrictiveness of ROOs. When the ROO is binding and the FTA is under-utilized, country B obtains more tariff revenues and the income effect is positive. Country A's producers have to employ more expensive intermediate goods to comply with ROOs and their production costs rise accordingly. But as another factor of production, domestic labor is cheaper as the ROO increases. Also, more firms choose pay the MFN tariffs when the ROO is too costly to satisfy, which in turn raises the prices of imports in country B. These three elements together determine the prices of imported differentiated goods from country A. Consumers in country B may face a higher price index of differentiated varieties and a lower real income. Moreover, similar to the analysis on country A, a large elasticity of substitution between varieties mitigates the variety effect on welfare. The

numerical solutions imply that when the ROO is slightly binding (between 20% and 40%), the positive income effect and the variety effect outweigh the negative price effect, and as a result, the country welfare increases with ROOs. As the ROO rises further (greater than 40%), the distortions in production generated by a strict ROO have a larger impact on product prices and the negative price effect is a more important determinant of welfare. Hence, ROOs reduce country B's welfare.

7 Conclusions

Motivated by the fact that free trade agreements are actually under-utilized and the associated costs accompanied with trade preferences are substantial, this paper investigates how rules of origin reallocate profits and incomes among countries when not all firms export by invoking preferential tariffs. I develop a general equilibrium model in the context where there is a vertical linkage between FTA member countries, featured the effects of ROOs on firms' decisions on the use of trade preferences. Tariff margin, rules of origin, and fixed documentation costs are identified as three key determinants of FTA utilization rates. Exporters are more likely to benefit from claiming the preferential tariffs offered by an FTA if the agreement sets a large tariff margin, a loose origin rule, and a simple administrative procedure.

The welfare patterns across countries differ by elasticity of substitution between varieties. When differentiated goods are highly substitutable, a slightly binding origin rule serves as a protection device to regional producers. Final good exporters may have an incentive to source intermediate inputs from even a higher cost local producer in order to satisfy the ROO criterion. However, a further stricter rule of origin would hurt exporters and domestic workers because of the low utilization rate of FTAs and the large distortions in manufacturing process. As such, there is an inverted-U shape relationship between the welfare of the upstream country, and the downstream country's welfare decreases with ROOs. When differentiated goods are less substitutable, variety effect dominates and both countries experience welfare increase as the ROO becomes stricter.

The model consists of a number of non-linear equations and lacks a closed form solution. I adopt two methods to proceed. First, I conduct comparative statics analysis in the neighborhood of full utilization and show the local effects of ROOs on endogenous variables. Second, I employ the numerical approach and provide a visualized example and intuitive explanations. Both the analytical solutions and counterfactual simulations yield consistent results.

References

[1] Balistreri, Edward, Russell H. Hillberry, and Thomas F. Rutherford. "Structural estimation and solution of international trade models with heterogeneous firms." *Journal of International Economics*, 2011, 83(2), pp. 95-108

[2] Brenton, Paul and Miriam Manchin. "Making EU Trade Agreements Work: The Role of Rules of Origin." *The World Economy*, 2003, 26(5), pp. 755-769

[3] Cadot, Oliver, Jamie de Melo, Akiki Suwa Eisenmann and Bolormaa Tumurchudur. "Assessing the Effect of NAFTAs Rules of Origin. " *Mimeo*, 2002

[4] Cadot, Oliver, Jaime De Melo, and Alberto Portugal-Perez. "Rules of origin for preferential trading arrangements : implications for the ASEAN Free Trade Area of EU and U.S. experience." *Journal of Economic Integration*, 2008, 22, pp. 256-287

[5] Cheong, Inkyo and Jungran Cho. "An Empirical Study on the Utilization Ratio of FTAs by Korean Firms." *Journal of Korea Trade*, 2009, 13(2), pp. 109-126

[6] Cherkashin, Ivan, Svetlana Demidova, Hiau Looi Kee, and Kala Krishna. "Firm Heterogeneity and Costly Trade: A New Estimation Strategy and Policy Experiments." *NBER Working Papers 16557*, 2009

[7] Demidova, Svetlana, Hiau Looi Kee, and Kala Krishna. "Do Trade Policy Differences Induce Sorting? Theory and Evidence from Bangladeshi Apparel Exporters." *Journal of International Economics*, 2012, 87(2), pp. 247-261.

[8] Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. "Dissecting Trade: Firms, Industries, and Export Destinations." *American Economic Review*, 2004, 94(2), pp. 150-154

[9] Foster, Lucia, John Haltiwanger, and Chad Syverson. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 2008, 98(1), pp. 394-425.

[10] Hayakawa, Kazunobu. "Impacts of FTA Utilization on Firm Performance." *IDE Discussion Paper No. 366*, 2012

[11] Hayakawa, Kazunobu, Hansung Kim, and Hyun-Hoon Lee. "Determinants on Utilization of Korea-ASEAN Free Trade Agreement: Margin Effect, Scale Effect, and ROO Effect." *BRC Research Report No.9*, 2010

[12] Hayakawa, Kazunobu, Daisuke Hiratsuka, Kohei Shiino, and Seiya Sukegawa. "Who Uses Free Trade Agreements?" *Asian Economic Journal*, 2013, 27(3), pp. 245-264

[13] Herin, Jan. "" *Rules of Origin and Differences between Tariff Levels in EFTA and in the EC.*, 1986, EFTA Occasional Paper No. 13

[14] Hiratsuka, Daisue, Kazunobu Hayakawa, Kohei Shino, and Seiya Sukegawa. "Maximizing Benefits from FTAs in ASEAN." *ERIA Research Project Report*, 2008(1), pp. 407-545

[15] Ju, Jiandong and Kala Krishna. "Firm Behavior and Market Access in a Free Trade Area with Rules of Origin." *Canadian Journal of Economics*, 2005, 38(1), pp. 290-308

[16] Keck, Alexander and Andreas Lendle. "New Evidence on Preference Utilization." *The WTO Working Paper, No.ERSD-2012-12*, 2012

[17] Lopez-Silanes, Florencio, James R. Markusen, and Thomas F. Rutherford. "Trade policy subtleties with multinational firms." *European Economic Review*, 1996, 40, pp.1605-1627

[18] Medalla, Erlinda and Maureen Rosellon. "Rules of Origin in ASEAN+1 FTAs and the Value Chain in East Asia." *Discussion paper series No. 2012-37*, Philippine Institute for Development Studies, 2012

[19] Melitz, Marc J. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 2003, 71(6), pp. 1695-1725.

[20] Melitz, Marc J., and Gianmarco Ottaviano. "Market Size, Trade, and Productivity." *Review of Economic Studies*, 2008, 75(1), pp. 295-316.

[21] Menon, Jayant. "Preferential and Non-Preferential Approaches to Trade Liberalization in East Asia: What Differences Do Utilization Rates and Reciprocity Make?" *ADB Working Paper Series on Regional Economic Integration, No. 109*, 2013

[22] Oyamada, Kazuhiko. "Parameterization of Applied General Equilibrium Models with Flexible Trade Specifications Based on the Armington, Krugman, and Melitz Models." *IDE Discussion Papers 380*, 2013, Institute of Developing Economies, Japan External Trade Organization(JETRO)

[23] Roberts, Mark, and James Tybout. "Sunk Costs and the Decision to Export in Colombia." *American Economic Review*, 1997, 87(4), pp. 545-564.

[24] Tybout, James. "Plant and Firm Level Evidence on "New" Trade Theories." *In J.Harrigan Ed. Handbook of International Economics, Basil Blackwell.*, 2002.

Table 1: Utilization Rates of Free Trade Agreements

FTA	**Year**	**Utilization Rate**
NAFTA	2005	55%
China-ASEAN FTA	2013	34.95%
Asia-Pacific Trade Agreement	2013	14.34%
China-Singapore FTA	2013	13.45%
China-New Zealand FTA	2013	31.67%
China-Chile FTA	2013	6.55%
China-Peru FTA	2013	5.49%
China-Pakistan FTA	2013	71.97%
Korea-ASEAN FTA	2005	20.11%
US Generalized System of Preference (GSP)	2008	60%
GSP-Electrical Machinery and Equipment	2008	22%
GSP-Iron and Steel	2008	85%

Data Sources: General Administration of China Customs, Office of the United States Trade Representative, and http://akfta.asean.org/.
Utilization rate is caclulated as the ratio of value of imports granted under FTA preference to value of imports eligible for FTA preference.

Table 2: Free Trade Agreements and Rules of Origin

FTA	Chapter (HS2007)	Rules of Origin
ASEAN FTA	41 - 48	A regional value content of not less than 40 percent; or A significant change in heading/subheading.
	51 - 63	A regional value content of not less than 40 percent; or A significant change in heading/subheading; or Process Rules for Textile and Textile Products as set out in Attachment 1.
	79, 80	Wholly obtained or produced in the exporting Member State.
	84, 85, 87, 90	A regional value content of not less than 40 percent; or A significant change in heading/subheading.
NAFTA	41 - 60	A significant change in heading /subheading.
	61 - 63	A significant change in heading /subheading; and (a) the good is both cut (or knit to shape) and sewn or otherwise assembled in the territory of one or more of the Parties; and (b) the visible lining fabric listed in Note 1 to Chapter 61 satisfies the tariff change requirements provided therein.
	64 - 72	A significant change in heading /subheading.
	73, 74, 79	A significant change in heading /subheading; or a regional value content of not less than 60 percent where the transaction value method is used or 50 percent where the net cost method is used.
	80 - 82	A significant change in heading /subheading.
	84, 85, 87, 90	A significant change in heading /subheading; or a regional value content of not less than 60 percent where the transaction value method is used or 50 percent where the net cost method is used.

Data Sources: ASEAN ANNEX 3 Product Specific Rules and NAFTA ANNEX 401.

Table 3: Processing time of origin certification across countries

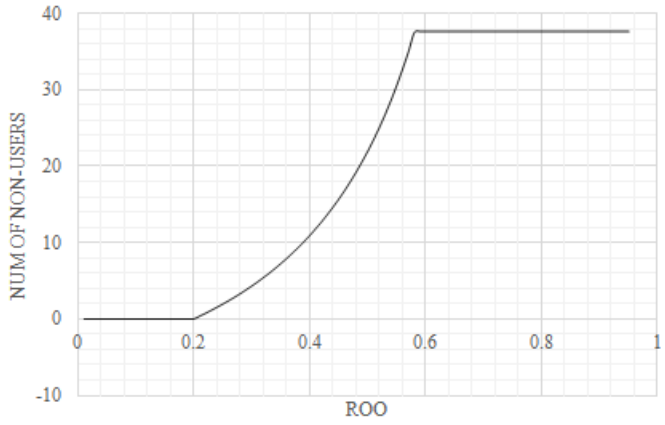
Country	**Issuing Authority**	**Examination of Origin**	**Issuance of Certificate**
Australia	Australian Industry Group	Automatic	Within 1 working day
Brunei	Department of Trade Development	30 days	1-2 working days
Cambodia	Ministry of Commerce	Within 7 working days	10 hours
China	Entry-Exit Inspection Bureau	20-30 working days	Within 1 day
Japan	Japan Chamber (JCCI)	Within 3 working days	Within 2 working days
Korea	Korean Customs Service	3-10 working days	Within 1 day
Laos	Ministry of Commerce and Industry	3-7 days	3 days
Malaysia	Ministry of International Trade and Industry	5 working days for online, 7 working days for manual application	1 working day for online, 2 days for manual
New Zealand	New Zealand Chamber	1 working day	1 working day
Philippines	Bureau of Customs	Within 5 working days	Within same day
Singapore	Singapore Customs	Factory registration - 1 week Manufacturing cost statement - at least 7 days	2-3 working hours
Thailand	Thailand Ministry of Commerce	3 working days	Within 1 day

Sources: ERIA Project (2011): Towards Accessible FTA: The Role of ROO Documentation in FTA Utilization; Medalla and Rosellon (2012).

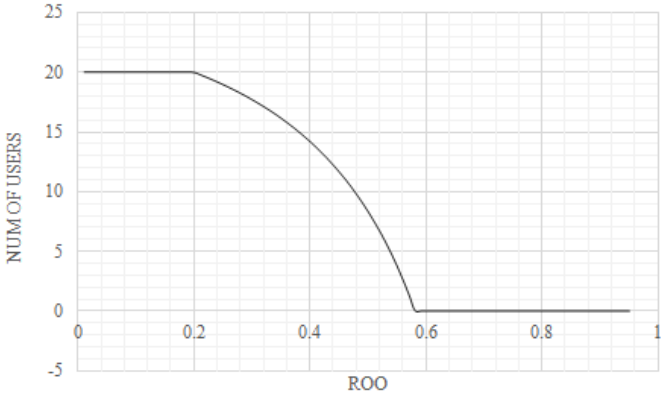
Table 4: Simulation results: the benchmark

Variable name	Value
Number of exporters from country A: n_a	20
Number of FTA non-users from country A: n_1	0
Number of FTA users from country A: n_2	20
FTA Utilization Rate: u	100%
Number of exporters from country B: n_b	20
Import value of intermediate inputs from B to A: EX_{mb}	0
Import value of intermediate inputs from C to A: EX_{mc}	11.62
Wage rate in country A: w_a	1
Welfare of country A	10.35
Welfare of country B	26.72

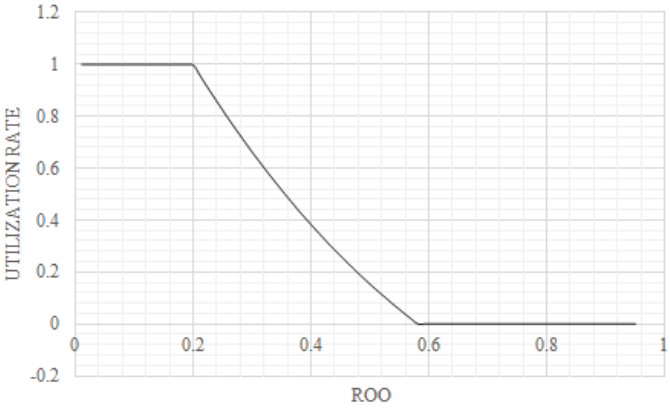
Figure 1: Numbers of FTA non-users and users and FTA utilization rate



(a)

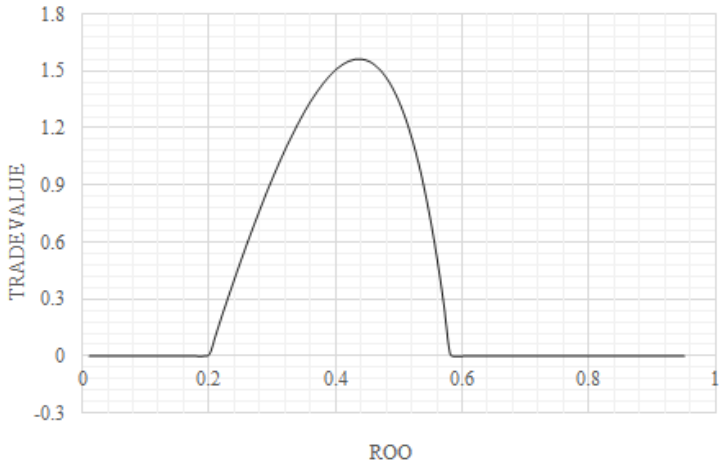


(b)

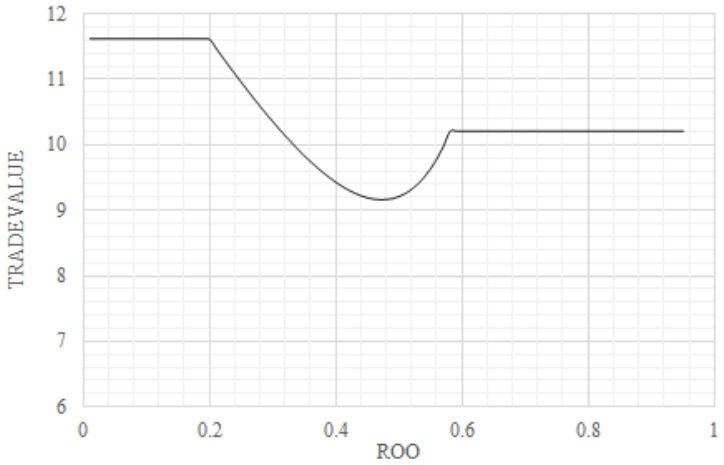


(c)

Figure 2: Import value of intermediate inputs between countries



(a)



(b)

Figure 3: Wage rate and ROOs

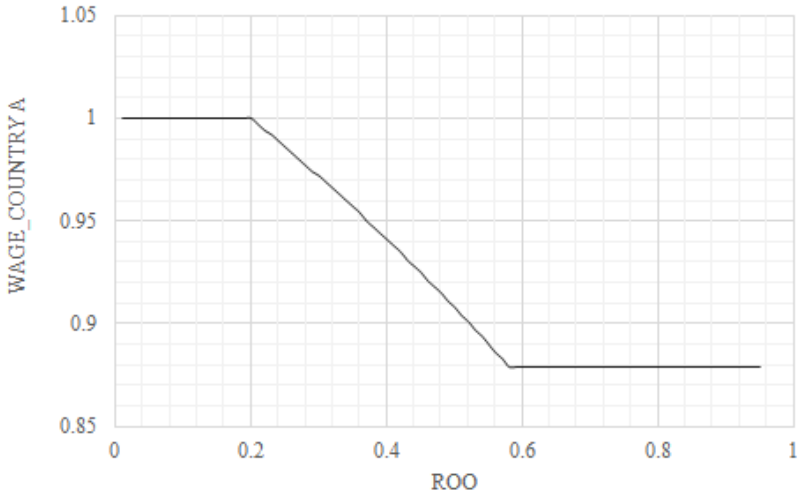
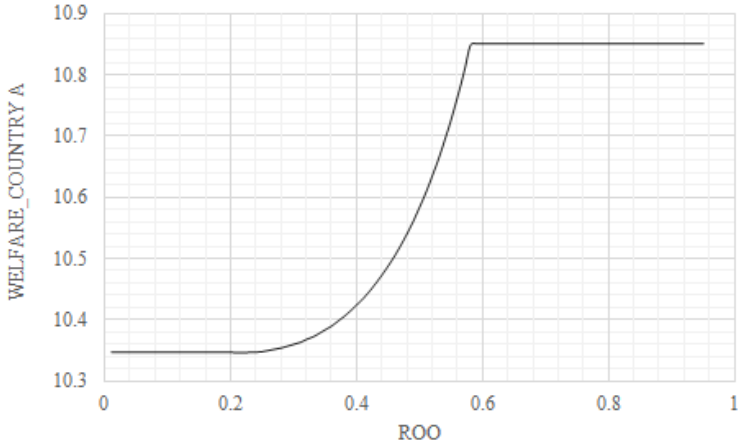
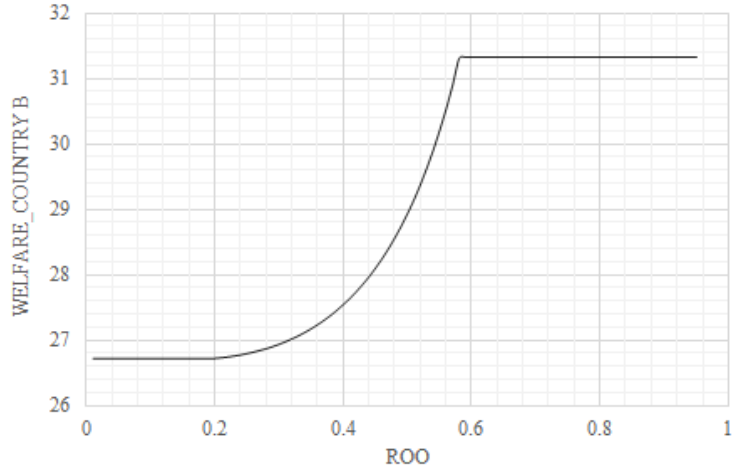


Figure 4: Country welfare and ROOs: $\sigma = 4$

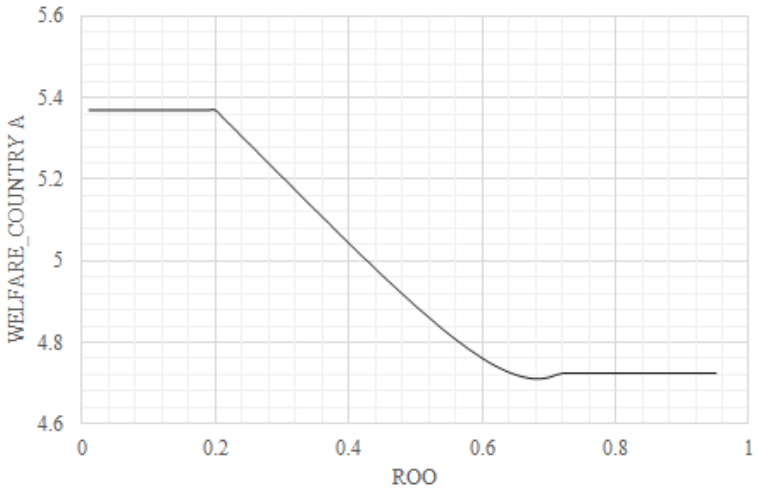


(a)

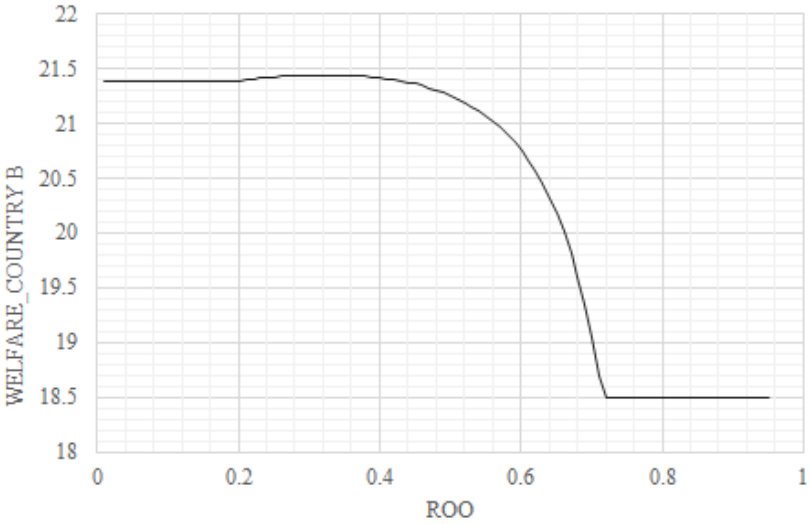


(b)

Figure 5: Country welfare and ROOs: $\sigma = 8$



(a)



(b)