Rules of origin and consumer-hurting free trade agreements*

Hiroshi Mukunoki[a] Hirofumi Okoshi[b,c]

Abstract

This study examines how the rules of origin (RoO) of a free trade agreement (FTA) affect firms' pricing strategies. The value-added criterion (VAC) of the RoO requires firms to add more than a certain level of value within an FTA when firms use inputs originating from outside the FTA. The VAC may work as a price floor, and the resulting increases in prices can benefit all firms if it induces an offshoring firm to manipulate its output price. Meanwhile, a consumer-hurting FTA formation is possible, even if all firms make tariff-free exports. Furthermore, such an FTA formation may worsen total welfare.

Keywords: Rules of origin; Free trade agreement; Pricing strategy

JEL classification: F13; F15; L11

*We appreciate Carsten Eckel and Andreas Haufler for their comments and suggestions. Hiroshi Mukunoki acknowledges the financial support of the JSPS KAKENHI (Grant Numbers JP16H03620 and JP17K03706). Hirofumi Okoshi acknowledges financial support from Deutsche Forschungsgemeinschaft (German Science Foundation, GRK1928).

[a]Gakushuin University, Faculty of Economics, Mejiro 1-5-1, Toshima-ku, Tokyo 171-8588, Japan

[b]Corresponding author. University of Munich, Munich Graduate School of Economics, Seminar for Economic Policy, Akademiestr. 1/II, D-80799 Munich, Germany

[c]Okayama University, Faculty of Economics, 3-1-1 Tsushima-naka, Kita-ku, Okayama, 700-8530, Japan

1 Introduction

Economic integration has boosted worldwide fragmentation and resulted in increased international trade over the last few decades. One of the main vehicles of economic integration is the proliferation of regional trade agreements (RTAs). As of September 2020, there were 306 RTAs in force worldwide.[1] Although trade liberalization through RTAs seems to reduce trade costs, decrease consumer prices, and benefit consumers, the effect is not as simple as it seems.[2] Among others factors, it has been suggested that rules of origin (RoO), which are indispensable in implementing free trade agreements (FTAs), complicate the effects of trade liberalization. If firms export their products by utilising preferential tariffs in an FTA, then firms must comply with RoO and prove that the exported products are produced within the FTA.

For instance, some empirical studies investigating the impact of FTAs have cast a sceptical eye on the fruitfulness of trade liberalisation because of RoO. Conconi et al. (2018) showed that the RoO used in NAFTA significantly reduce imports of intermediate products from non-member countries relative to member ones, which implies that, in order to meet the RoO, there is input relocation from more efficient input production countries to less efficient ones. Takahashi and Urata (2010) and Hayakawa et al. (2013) found that not all firms utilise FTA tariffs owing to RoO, which indicates that the impact of forming an FTA is heterogeneous across firms and may be overestimated. Unfortunately, despite the practical importance, little empirical investigation has been undertaken because the complexity of RoO makes it difficult to conduct empirical research on the impact of FTA formation with RoO.[3] Thus, identifying the impact of FTAs with a particular focus on RoO remains an important research topic that could provide useful policy implications. This paper theoretically tackles this issue and provides new insights that have been overlooked in the literature.

[1] The cumulative number of RTAs in force was 83 in 2000, 137 in 2005, and 214 in 2010. See http://rtais.wto.org/UI/charts.aspx

[2] Goldberg and Pavcnik (2016, p. 164) pointed out that 'the world may not be as liberalized as it seems and that failure to document significant effects of trade policies may instead be due to measurement and identification challenges rather than the absence of such effects.

[3] The existing research analysing the impact of FTA formations with RoO on prices include Cadot et al. (2005) and Hayakawa et al. (2019).

Among the several methods for checking the origin of products, the valued-added criterion (VAC) is the focus of this study. Let p denote the export value of the product and c denote the value of input materials not originating in the FTA; then, the VAC typically requires that the value-added ratio, $\frac{p-c}{p}$, be larger than a specified level. This method of calculating the value-added content is called the 'transaction value method.' Estevadeordal and Suominen (2003) reported that among the 87 FTAs they analysed, 68 employed this method, at least in a particular product category. Because the value-added ratio is associated with an exporting firm's pricing as well as its input sourcing, the VAC may induce exporting firms to increase the value-added ratio by increasing their export prices. These price increases, in turn, soften market competition and lead to increases in the prices of other firms. In other words, the VAC can be a commitment device to increase output prices.

Krishna and Krueger (1995) mentioned the possibility of the RoO-induced manipulation of output prices. They compared the transaction value method of the VAC with the cost-based method, by which the value-added content is calculated as the proportion of the value of input materials originating in the FTA relative to the total value of input materials. They suggested that firms prefer the transaction value method because a higher price lowers the restrictiveness of the VAC. Some policy papers have also pointed out that firms are incentivised to manipulate output prices under the VAC. For instance, Falvey and Reed (1998, p. 213) indicated that the VAC '. . . allows room for [the] manipulation of prices,' and Reuter (2012, p. 214) also pointed out that 'most rules of origin are on a percent-of-value basis. . . . By overinvoicing the value added, the MNE can more easily meet a rule-of-origin test and qualify for duty-free entry for its products into another country in the free trade area'. These studies, however, did not analyse the welfare effects of FTAs when the VAC of the RoO leads to the manipulation of output prices.[4]

Against this backdrop, this study examines how RoO affect firms' incentives to set the

[4]Although there is no direct empirical evidence that supports our arguments, some empirical studies have found that firms in an FTA use significantly increase trade prices, and this increase is not necessarily attributed to an increase in variable costs to comply with RoO. See, for example, Cadot et al. (2005) and Hayakawa et al. (2019). These studies do not identify how the price-increasing effect of FTAs is attributed to the manipulation of export prices to meet the VAC of RoO, but at least the price-increasing nature of FTAs is consistent with our theoretical prediction.

prices of final goods in an oligopoly model with one input-offshoring firm and one firm using the inputs produced inside the FTA. We demonstrate that there is a case in which an FTA formation with RoO increases the prices of all goods, even when all firms export tariff-free products within the FTA. This is because a firm that uses inputs outside the FTA has an incentive to increase its product price to comply with the VAC of the RoO. In other words, the firm can credibly commit to setting high export prices with RoO. Thus, the VAC of RoO may work as a price floor, below which the firm using non-FTA inputs cannot comply with the RoO. In response, the other exporting firm, which is a rival in the product market, also increases its export prices. Because of this price-increasing effect, more stringent RoO can benefit all firms. If the induced increases in export prices outweigh the tariff-elimination effect that should reduce the consumer price, the FTA formation will hurt consumers even though tariffs against all firms are eliminated in the equilibrium. In other words, an FTA with stringent RoO can be 'consumer-hurting FTA', rather than 'consumer-benefiting FTA', in the sense that its formation decreases the consumer surplus of the liberalising countries. Furthermore, such stringent RoO may make an FTA formation welfare reducing for those countries inside it, even compared to their welfare before the FTA.

Extant theoretical studies have focused on how RoO change input sourcing (Ju and Krishna, 2005), the degree of market integration inside an FTA (Ishikawa et al., 2007), the productivity cut-off of exporting for users and non-users of FTA tariffs (Demidova and Krishna, 2008), and the patterns of foreign direct investment (Mukunoki, 2017). Jinji and Mizoguchi (2016a,b) have analysed the optimal choice of RoO. However, these studies have not considered price manipulation to comply with RoO. The closest research to the current study is another paper of ours, Mukunoki and Okoshi (2020), which focuses on transfer-price manipulation of a multinational firm to comply with the VAC of an RoO. This work complements the current study in that both focus on different levels of price manipulation.

The rest of the paper is organised as follows. Section 2 introduces our model and derives the optimal strategy for firms. Section 3 explores the effects of FTA formation on consumer surplus and the profits of firms. Section 4 examines the changes in total welfare. Section

5 discusses the robustness and extensions of the baseline model. Section 6 concludes the paper.

2 Model

Two firms (I and O) produce differentiated products only in a home country (H) and sell them in country H and a foreign country (F). Countries H and F are potential FTA partners. The reason why the two firms are located in country H is that we assume the cost of production in country H is substantially low, enough to attract both firms to produce their products in that location, because of low wages or the availability of labour with the specific skills needed to produce these firms' goods.[5] For simplicity, we abstract from the market in H as the primary focus of this paper is on the firms' export-pricing strategies. Even if we explicitly consider the market in country H, the qualitative nature of our results remains unchanged. See Section 5.3 for further details.

The production of final products requires inputs made either in country H or in countries outside the two countries. One unit of input is transformed into one unit of output at a constant marginal cost, which is normalised to 0.[6] The input market is under perfect competition, and the prices of the 'inside inputs' produced in country H are given by c, while the prices of the 'outside inputs' produced outside the two countries are lower, given by $c - \Delta$ ($\Delta \in (0, c)$). We assume that only firm O is an offshoring firm that can use inputs imported from outside the two countries. Firm I always uses the inputs produced in country H because the firm lacks knowledge about foreign input markets and cannot cover the cost of searching for appropriate suppliers.[7]

The indirect utility of the representative consumer in country F is $V(p_I, p_O) = \bar{V} - a(p_I + p_O) + \frac{(p_I)^2 + (p_O)^2}{2} - bp_I p_O + Y$, where $\bar{V}$ is a positive constant, p_i is the price of product i manufactured by firm $i \in \{I, O\}$, b is the degree of product substitutability, and

[5] If the market size in country H is sufficiently large, the two firms will also choose to produce in country H, though the market in H is not explicitly considered in our model.

[6] In Section 5.1, we discuss the robustness of our results when firms incur a positive marginal cost.

[7] Considering firm I's procurement of the outside inputs complicates the model without changing the main results of the baseline model. See Section 5.4 for details.

Y is the consumption of the numéraire good. The utility maximisation yields:

$$x_i = a - p_i + bp_j, \tag{1}$$

where x_i is the demand for product i in country F, and $j \in \{I, O\}$ $(j \neq i)$.

A specific import tariff, τ, is imposed on both products by country F. There are no tariffs on the inputs. The FTA formation removes τ, but zero tariff is applied only if firms comply with RoO. Firm O may use the local inputs to meet the RoO, and thus, its marginal cost is either c or $c - \Delta$. Firm i's profit is given by

$$\Pi_i = (p_i - c_i - \lambda_i \tau) x_i, \tag{2}$$

where $c_I = c$ and $c_O \in \{c, c - \Delta\}$. λ_i is an index that takes zero if firm i meets the RoO after an FTA formation and takes one otherwise.

For the RoO, we consider a VAC that requires the firms to add at least $\underline{\alpha}$ fraction of the value of the exported products within the FTA countries. Since firm I never uses the imported inputs, it always meets the VAC upon the FTA formation and enjoys tariff-free access to country F.[8] However, firm O needs to use either (i) the local inputs or (ii) the outside inputs, and it sets its export price such that the following is satisfied:

$$\alpha(p_O) \equiv \frac{p_O - (c - \Delta)}{p_O} \geq \underline{\alpha}. \tag{3}$$

Because $\alpha(p_O)$ increases with p_O, even if firm O procures the outside inputs, $\alpha(p_O)$ exceeds the required level ($\underline{\alpha}$) if p_O is high enough.

Equilibrium

Given that (3) is not binding, the equilibrium price of good i is obtained by maximising (2):

$$\tilde{p}_i = \frac{(2+b)a + 2c_i + bc_j + \left(2\lambda_i + b\lambda_j\right)\tau}{4 - b^2} \quad (j \neq i), \tag{4}$$

[8] We abstract from the fixed costs of the RoO. The effect of introducing fixed costs on the results is discussed in Section 5.2.

The equilibrium sales of good i are $\widetilde{x}_i$, and the equilibrium profit and consumer surplus become $\widetilde{\Pi}_i = (\widetilde{x}_i)^2 = (\widetilde{p}_i - c_i)^2$ and $\widetilde{CS} = V(\widetilde{p}_I, \widetilde{p}_O) - Y$, respectively. By (4), we can characterise the equilibrium in the following regimes.

- No Agreement (N): Before the FTA formation, a tariff is imposed on both goods ($\lambda_I = \lambda_O = 1$). Firm O uses outside inputs $c_O = c - \Delta$. By substituting these parameters into (4), we obtain the equilibrium prices, denoted as $p_O^N = \frac{(2+b)(a+c+\tau)-2\Delta}{4-b^2}$ and $p_I^N = \frac{(2+b)(a+c+\tau)-b\Delta}{4-b^2}$.

- Non-compliance (NC): After the FTA formation, firm O uses the outside input, and the tariff is imposed only on firm O's product. By substituting $\lambda_I = 0$, $\lambda_O = 1$, and $c_O = c - \Delta$ into (4), the equilibrium price becomes $p_O^{NC} = \frac{(2+b)(a+c+\tau)-2(\Delta-\tau)}{4-b^2}$ and $p_I^{NC} = \frac{(2+b)(a+c)-b(\Delta-\tau)}{4-b^2}$. If τ satisfies $\tau \geq \tau^{exit} \equiv \frac{(2+b)(a-c+bc)}{2-b^2}$, firm O exits the market.

- Input relocation (IR): After the FTA formation, firm O uses inside inputs to comply with the RoO. With $\lambda_I = \lambda_O = 0$ and $c_O = c$ in (4), the equilibrium price is $p_O^{IR} = \frac{(2+b)(a+c)}{4-b^2}$ and $p_I^{IR} = \frac{(2+b)(a+c)}{4-b^2}$.

Alternatively, suppose that firm O uses the outside inputs ($c_O = c - \Delta$) after the FTA formation and still complies with the RoO. There are two sub-cases.

- Non-binding RoO (NB): Both firms set the unconstrained, optimal prices, which are given by substituting $c_O = c - \Delta$ and $\lambda_I = \lambda_O = 0$ into (4), yielding $p_O^{NB} = \frac{(2+b)(a+c)-2\Delta}{4-b^2}$ and $p_I^{NB} = \frac{(2+b)(a+c)-b\Delta}{4-b^2}$. The VAC is non-binding if

$$\alpha(p_O^{NB}) = \frac{p_O^{NB} - (c-\Delta)}{p_O^{NB}} = 1 - \frac{(4-b^2)(c-\Delta)}{(2+b)(a+c)-2\Delta} \equiv \alpha^{NB} \geq \underline{\alpha} \tag{5}$$

holds.

- Binding RoO (B): If $\alpha^{NB} < \underline{\alpha}$ holds, the RoO are binding and firm O sets the price, p_O^B, such that $\alpha\left(p_O^B\right) = \underline{\alpha}$ is satisfied. The equilibrium prices become

$$p_O^B = \frac{(c-\Delta)}{(1-\underline{\alpha})} \text{ and } p_I^B = \frac{(a+c)}{2} + \frac{b(c-\Delta)}{2(1-\underline{\alpha})}. \tag{6}$$

If $\underline{\alpha}$ is high enough to satisfy $\alpha^{exit} \equiv \frac{(2+b)(a-c+bc)+(2-b^2)\Delta}{(2+b)a+bc}$, then firm O exits the market, where $\alpha^{NB} < \alpha^{exit}$ holds.

Because (3) is binding, the prices in Regime B are different from the prices given by (4). The profits of firm O in Regime B are given by $\Pi_O^B = \frac{\underline{\alpha}(c-\Delta)}{2(1-\underline{\alpha})}\left((2+b)a+bc-\frac{(c-\Delta)(2-b^2)}{1-\underline{\alpha}}\right)$. The profits of firm I are given by $\Pi_I^B = (p_I^B - c)^2$. The consumer surplus in Regime B is given by $CS^B = V(p_I^B, p_O^B) - Y$.

The profits and consumer surplus under Regime $\omega \in \{N, NC, IR, NB, B\}$ are given by Π_i^ω and CS^ω, respectively. The total welfare of member countries is given by $W^\omega \equiv CS^\omega + \Pi_O^\omega + \Pi_I^\omega + TR^\omega$, where TR^ω is the tariff revenue under Regime ω. We have $TR^{IR} = TR^{NB} = TR^B = 0$.

Firms' choices of FTA use and input relocation

When the RoO are non-binding ($\underline{\alpha} \leq \alpha^{NB}$), $\Pi_O^{NB} > \max[\Pi_O^{NC}, \Pi_O^{IR}]$ holds. Thus, firm O always chooses Regime NB, and both firms use the FTA without changing their pricing and sourcing strategies. The FTA formation lowers the prices of all goods and benefits consumers and all firms.

However, when the RoO are binding ($\underline{\alpha} > \alpha^{NB}$), firm O chooses among Regimes NC, IR, and B. We confirm that $\Pi_O^{NC} \geq \Pi_O^{IR}$ holds if $\tau \leq \Delta$, whereas $\Pi_O^{IR} > \Pi_O^{NC}$ holds if $\tau > \Delta$.[9] When the tariff that is eliminated by complying with the RoO is lower than the additional marginal cost from input relocation, firm O prefers Regime NC to Regime IR, and vice versa.

We confirm that Π_O^B takes an inverse U-shaped form in $\underline{\alpha} \in [\alpha^{NB}, \alpha^{exit}]$. Specifically, $\Pi_O^B = \Pi_O^{NB}$ at $\underline{\alpha} = \alpha^{NB}$, Π_O^B takes the maximum at $\underline{\alpha} = \alpha^{\max} \equiv \frac{(2+b)a+bc-\left(2-b^2\right)(c-\Delta)}{(2+b)a+bc+(2-b^2)(c-\Delta)}$, and $\Pi_O^B = 0$ at $\underline{\alpha} = \alpha^{exit}$. This result implies that there is a unique cut-off level of the VAC, α^{NC}, such that $\Pi_O^B > \Pi_O^{NC}$ holds with $\alpha^{NB} < \underline{\alpha} < \alpha^{NC}$ and $\Pi_O^{NC} \geq \Pi_O^B$, with $\alpha^{NC} \leq \underline{\alpha}$. Similarly, there is a unique cut-off level, α^{IR} such that $\Pi_O^B > \Pi_O^{IR}$ holds with $\alpha^{NB} < \underline{\alpha} < \alpha^{IR}$ and $\Pi_O^{IR} \geq \Pi_O^B$ with $\alpha^{IR} \leq \underline{\alpha}$. Note that $\alpha^{IR} < \alpha^{NC}$ holds if $\tau > \Delta$ holds and $\alpha^{NC} \leq \alpha^{IR}$

[9]Given $\tau < \tau^{exit}$ such that $\Pi_O^{NC} > 0$, we have $\Pi_O^{NC} - \Pi_O^{IR} = (\Delta - \tau)\left(2 - b^2\right)^2 (2\tau^{exit} - \tau + \Delta)$. This means that $\Pi_O^{NC} \gtreqless \Pi_O^{IR} \iff \Delta \gtreqless \tau$.

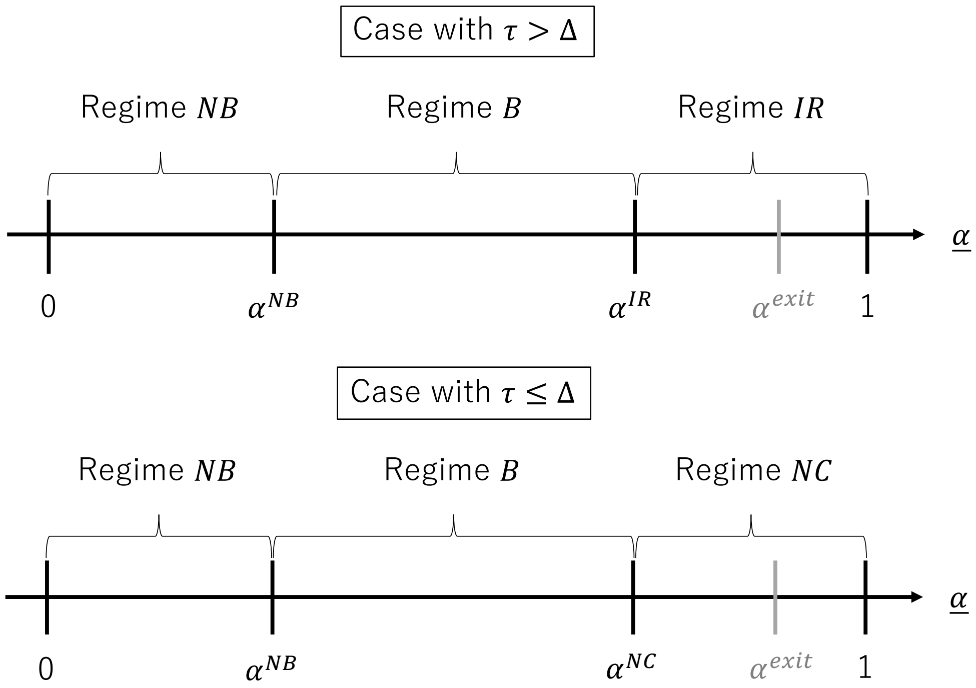

Figure 1: RoO and Equilibrium Regime

otherwise. Furthermore, as Π_O^{NC} and Π_O^{IR} are positive and independent of α, $\alpha^{NC} < \alpha^{exit}$ and $\alpha^{IR} < \alpha^{exit}$ hold. This means that firm O never exits the market of country F. The following proposition summarises the choice of firm O.

Proposition 1. After an FTA formation, (i) the RoO are non-binding (Regime NB) if $\underline{\alpha} \leq \alpha^{NB}$ holds; (ii) firm O complies with the RoO with price manipulation (Regime B) if $\alpha^{NB} < \underline{\alpha} < \min[\alpha^{NC}, \alpha^{IR}]$ holds; (iii) firm O complies with the RoO with input relocation (Regime IR) if $\Delta < \tau$ and $\alpha^{IR} \leq \underline{\alpha}$ hold; and (iv) firm O does not comply with the RoO (Regime NC) if $\Delta \geq \tau$ and $\alpha^{NC} \leq \underline{\alpha}$ hold.

The equilibrium choice presented in Proposition 1 is depicted in Figure 1.

3 Profit-enhancing RoO and consumer-hurting FTA

In the previous section, we characterise the regime that becomes the equilibrium outcome. In this section, we examine how a formation of FTA affects firms and consumers. Under Regimes NB, IR, and NC in the post-FTA equilibrium, the prices set by the firms and their profits are independent of the stringency of the VAC of the RoO, either because the VAC is

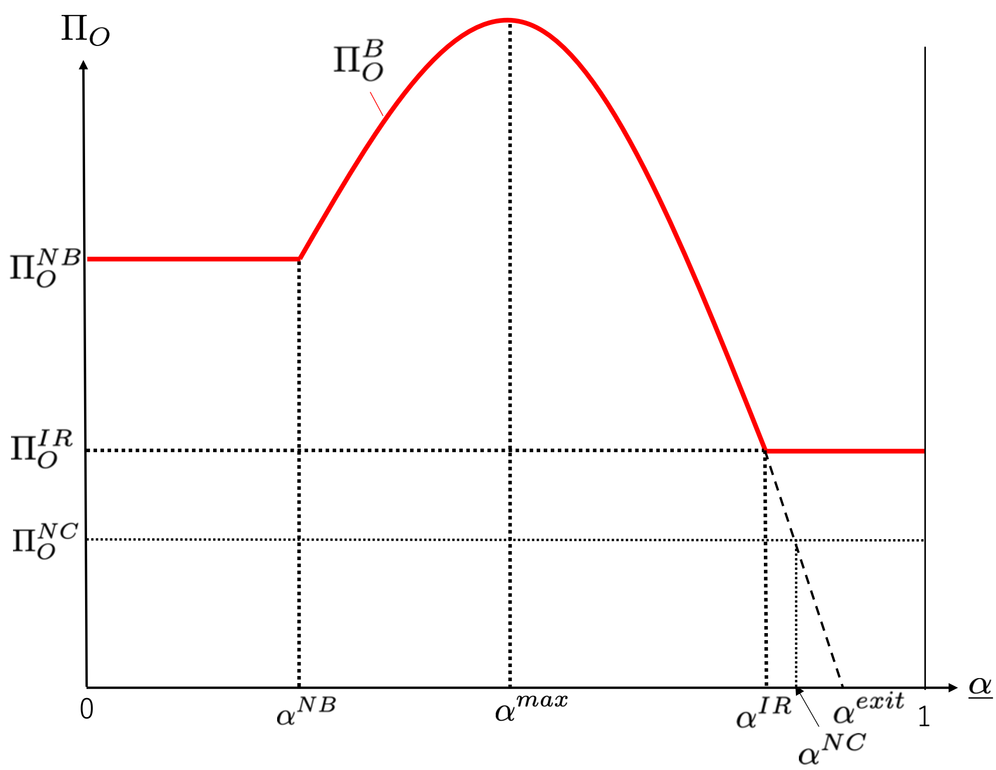

Figure 2: RoO and Profit of Firm O

non-binding or because firm O does not use the FTA tariff.

Under Regime B in Figure 1, however, firm O sets the price to satisfy the VAC, regardless of how firm I sets the price of its product, and the price is higher than p_O^{NB}. Then, firm I's optimal reaction is to raise its own price. If $\underline{\alpha}$ is not very high, these increases in prices raise the profits of both firms. We have the following proposition.

Proposition 2. Under Regime B, the profits of all firms increase as the VAC becomes more stringent for $\underline{\alpha} \in [\alpha^{NB}, \alpha^{\max})$.

In other words, the RoO function as a price floor and become a commitment device for firms to weaken price competition. If $\underline{\alpha}$ is very high, however, the price increase is too large for firm O, and its profit will fall below Π_O^{NB}. That is, firm O's profits have an inverse U-shape, which is shown in Figure 2 in the case of $\Delta \leq \tau$. Since the pre-FTA profit of firm O, Π_O^N, lies somewhere between Π_O^{NC} and Π_O^{NB}, an FTA formation increases firm O's profit unless $\underline{\alpha}$ is very high. Firm I always gains from an FTA formation because it benefits not only from tariff elimination but also from a price increase in the rival's product in Regimes B and IR. If firm O chooses Regime NC, only firm I benefits from tariff elimination.

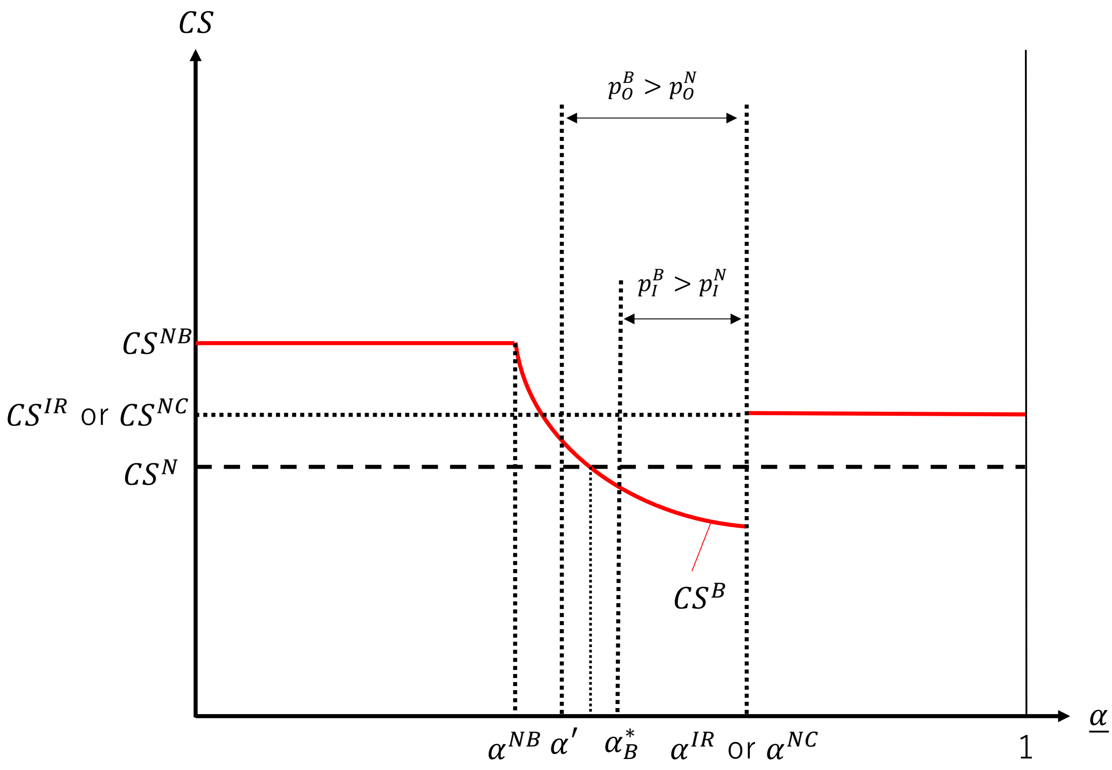

Figure 3: RoO and Consumer Surplus

Moreover, the price-increasing effect of binding RoO substantially changes the effects of FTA formation on consumers. Under non-binding RoO (NB), input relocation (IR), or non-compliance (NC), an FTA formation is consumer-benefiting because it decreases prices. It also benefits firms that use the FTA. However, the price-increasing effect of the RoO might overturn the positive effect for consumers, as the following proposition states.

Proposition 3. When the post-FTA equilibrium regime is Regime B, there exists a unique cut-off level for the VAC, α_B^*, such that an FTA formation raises the prices of all goods, benefits all firms, and hurts consumers for $\underline{\alpha} > \alpha_B^*$, even though all imports become tariff free with the FTA. As long as the initial tariff is not very large, we have $\alpha_B^* \in (\alpha^{NB}, \min[\alpha^{NC}, \alpha^{IR}])$ and a profit-increasing, consumer-hurting FTA becomes the equilibrium outcome if $\alpha_B^* < \underline{\alpha} < \min[\alpha^{NC}, \alpha^{IR}])$ holds.

Figure 3 shows the relationship between $\underline{\alpha}$ and consumer surplus, where CS^ω is the consumer surplus under Regime $\omega \in \{N, NC, IR, NB, B\}$. In Regime B, the equilibrium prices increase in $\underline{\alpha}$. The price of good O is higher than the pre-FTA level if $\underline{\alpha} > \alpha'$ holds (α' is defined in the proof of Proposition 3). The price increase of good I is smaller than that of good O because it is the second-order effect induced by the price increase of the other good. The price of good I also exceeds the pre-FTA level if $\underline{\alpha} > \alpha_B^*$ holds. The cut-off level of $\underline{\alpha}$, above which the FTA becomes consumer-hurting, lies somewhere between α' and α_B^*. Since the degree of the increase in the price of good O is larger than that of good I, consumers can become worse off, even with $p_O^B > p_O^N$ and $p_I^B < p_I^N$.

4 Welfare effects of FTA with RoO

Here, we explore the welfare effect of forming an FTA and how it is related to the stringency of the RoO. In line with traditional models, it is easily confirmed that the total welfare of member countries is improved when the RoO change neither firm O's pricing nor its input sourcing. Thus, an FTA formation always improves total welfare inside the FTA if the post-FTA equilibrium regime is Regime NB. However, as analysed, a stricter VAC of the RoO can influence firm O's pricing or input sourcing.

From Proposition 3, the overall impact of the FTA formation that leads to Regime B on total welfare is not obvious, and it reduces the total welfare of member countries with $\underline{\alpha} > \alpha_B^*$, even though all imports become tariff free. Even if $\underline{\alpha} < \alpha_B^*$ holds and the FTA formation decreases the price of good I, it can reduce total welfare because the negative effect of an increase in the price of good O can dominate the positive effect of a decrease in the price of good I. Because we confirm that $\frac{\partial W^B}{\partial \underline{\alpha}} < 0$ holds, $W^B < W^N$ at $\underline{\alpha} = \alpha_B^*$ means that there is a cut-off level of the VAC, α_B^W ($< \alpha_B^*$), such that $W^B < W^N$ if and only if $\underline{\alpha} > \alpha_B^W$ holds under Regime B. The level of α_B^W is depicted in Figure 3 in the case of $\Delta \leq \tau$.[10]

In an extreme case with $b = 0$, there are no interactions between firms. In this case, an increase in p_O does not affect the equilibrium level of p_I, and an FTA formation worsens welfare only if $\underline{\alpha}$ is sufficiently large (shaded area of Regime B in Figure 4). As b becomes higher, an increase in p_O increases the equilibrium level of p_I, and an FTA formation is more likely to be welfare reducing. The following proposition summarises the welfare effect when the post-FTA regime is Regime B.

[10] From (5), $\frac{\partial \alpha^{NB}}{\partial b} > 0$ is obtained. Furthermore, $\frac{\partial \tilde{\Pi}_O^{\omega}}{\partial b} > 0$ and $\frac{\partial^2 \tilde{\Pi}_O^B}{\partial \underline{\alpha} \partial b} > 0$ implies α^* is increasing in b. We confirm that $\frac{\partial^2 \Pi_O^B}{(\partial \underline{\alpha})^2} < 0$, holds for $\underline{\alpha} \in [\alpha^{NB}, \alpha^{exit}]$, where $\frac{\partial \Pi_O^B}{\partial \underline{\alpha}} > 0$ at $\underline{\alpha} = \alpha^{NB}$ and $\frac{\partial \Pi_O^B}{\partial \underline{\alpha}} < 0$ at $\underline{\alpha} = \alpha^{exit}$. The detailed calculation is available upon request.

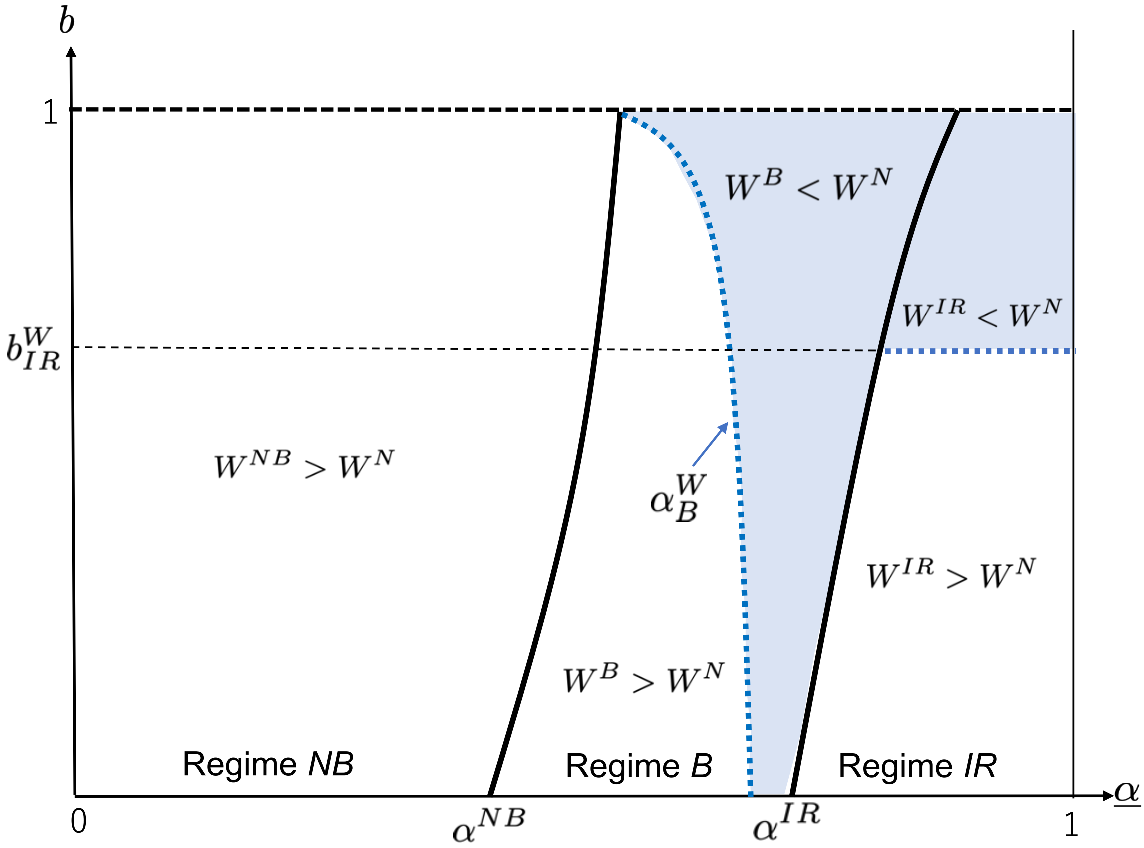

Figure 4: RoO and Welfare

Proposition 4. When the post-FTA equilibrium regime is Regime B, there exists a unique cut-off level of the VAC, α_B^W, such that an FTA formation worsens total welfare if $\alpha_B^W < \underline{\alpha} < \min[\alpha^{NC}, \alpha^{IR}]$ holds.

We should also note that even if the post-FTA equilibrium regime is either Regime IR or Regime NC, an FTA formation may worsen welfare. Regime IR has a similar threshold b_{IR}^W such that the FTA worsens the total welfare if $b \in (b_{IR}^W, 1]$. This area is depicted in Figure 4 as the shaded area in Regime IR.[11] If τ is only slightly greater than Δ, then the FTA formation is always harmful for the member countries (see the proof of Proposition 5 for details). This is because the input relocation increases the marginal cost of production of firm O with limited declines in consumer prices. Therefore, an FTA that leads to Regime IR results in welfare reduction in this case.

Proposition 5. When the post-FTA equilibrium regime is Regime IR, if the initial tariff level

[11]We always have $\frac{\partial^2(W^{IR}-W^N)}{(\partial b)^2} < 0$ and $(W^{IR} - W^N)|_{b=1} < 0$, and $(W^{IR} - W^N)|_{b=0} > 0$ also holds if $\tau \geq \sqrt{\{2(a-c)+\Delta\}\{2(a-c)+7\Delta\}} - \{2(a-c)+\Delta\} > 0$.

is large, there exists a unique cut-off level of b, $b^W_{IR}(\in (0,1))$, such that an FTA formation worsens total welfare if $b \in (b^W_{IR}, 1)$ holds. If the initial tariff level is small, an FTA formation always worsens total welfare.

Similarly, if the post-FTA equilibrium outcome is Regime NC, then there also exists a unique level of product substitutability, b^W_{NC}, such that an FTA formation reduces total welfare for $b \in (b^W_{NC}, 1]$.[12] In Regime NC, the production cost of firm O is lower than that of firm I. Because the tariff is eliminated only for good I in this regime, the production and exports of good O decrease while those of good I increase. This 'trade diversion' effect causes inefficiency, so an FTA can be welfare reducing. As the two goods become closer substitutes, the substitution effect between the two goods becomes larger, and an FTA that leads to Regime NC is more likely to become welfare reducing.

Proposition 6. When the post-FTA equilibrium regime is Regime NC, there exists a unique cut-off level of b, $b^W_{NC}(\in (0,1))$, such that an FTA formation worsens total welfare if $b \in (b^W_{NC}, 1)$ holds.

5 Discussion

Our benchmark model has employed some key assumptions to simplify the model. First, we assume that only one of the two firms (firm O) is able to access cheaper offshore inputs, even though firms have the same technology to produce final goods. As both existing theoretical and empirical studies provide evidence that offshoring firms are more productive than onshoring firms, it is natural to introduce cost differences among them. Second, we assume that the marginal cost of transforming inputs into final goods is zero. One may question the assumption as it means that firm O re-exports the product at price $p^\omega_O > c_O$ without adding any value. Third, we abstract from the fixed costs of the RoO. It is believed that one of the reasons why not all firms use FTA tariffs is the bothersome documentation process.[13] Fourth, we did not consider the market in country H. Finally, we

[12]Since $\frac{\partial^2(W^{NC}-W^N)}{(\partial b)^2} < 0$, $(W^{NC} - W^N)|_{b=0} > 0$, and $(W^{NC} - W^N)|_{b=1} < 0$ hold, b^W_{NC} always exists in this range.

[13]For example, Francois et al. (2006) estimated that the tariff-equivalent costs of using preferences under the ACP-EU Partnership Agreement ranged between 4% and 4.5%.

assume that only one of the two firms can use the inputs imported from countries outside the FTA.

Here, we show that qualitative nature of our results remains unchanged if we relax these assumptions.

5.1 Labour cost and productivity difference

Let us modify the model by introducing labour to transform inputs into final goods. Both firms can hire labour from a perfectly competitive labour market at wage w, but the unit-labour requirement is different. Let l_i be the unit-labour requirement of firm i, and $l_I > l_O$ holds as we suppose that firm O is more productive. Then, the marginal cost for firm i is rewritten as $C_i = c_i + wl_i$. The term wl_i can also be interpreted as the input costs of the domestic materials.

By replacing c_i with C_i in the benchmark model, we can obtain the optimal prices. As the VAC is calculated as $\alpha\left(p_O^{NB}\right) = \frac{p_O^{NB}-(c-\Delta)}{p_O^{NB}}$,

$$\alpha^{NB} = 1 - \frac{(4-b^2)(c-\Delta)}{(2+b)(a+c) - 2\Delta + (2+b)wl_O + bw(l_I - l_O)} \;(<1) \tag{7}$$

is derived. Because including the labour cost and productivity difference increases p_O^{NB}, it raises the level of α^{NB} and expands the range of $\underline{\alpha}$ that realises Regime NB. Intuitively, since the labour costs contribute to piling up the regional value-added as well as an increase in the price–cost margin, it becomes easier for firm O to meet the VAC of the RoO, and it widens the range of $\underline{\alpha}$ in which the RoO is non-binding. For the same reason, including labour costs should increase α^{IR} and α^{NC} because the adjustment of the export price needed to meet the VAC of the RoO is smaller with the domestic labour costs, and it becomes easier for firm O to comply with the RoO in Regime B.

For $\underline{\alpha} \in (\alpha^{NB}, \min[\alpha^{IR}, \alpha^{NC}])$, Regime B becomes the equilibrium outcome and changes in $\underline{\alpha}$ affect firm O's pricing and profits as those in the benchmark model do. Therefore, the main results remain unchanged with this alternative setup. The benchmark analysis corresponds to the case with $wl_O = wl_I = 0$.

Furthermore, as the difference in productivity $(l_I - l_O)$ increases, α^{NB}, α^{IR}, and α^{NC}

increase, and it becomes easier for firm O to comply with the RoO. This property is consistent with Demidova and Krishna (2008), which suggests that only firms whose productivity is above a cut-off level invoke RoO and use FTA tariffs.

5.2 Fixed costs of meeting RoO

Let us consider a case where firms must bear the fixed costs of meeting RoO. Firms incur fixed costs only when they apply for FTA tariffs. In Figure 2, the curves representing the profits of firm O under Regimes NB, B, and IR are shifted downward by the fixed costs, but the curve under Regime NC remains unchanged.

The presence of fixed costs makes the use of FTA tariffs less attractive. For instance, $\Pi_O^{NC} > \Pi_O^{IR}$ can hold even if $\tau > \Delta$ holds and firms are more likely to choose non-compliance with the RoO. In addition, an FTA formation is more likely to be welfare-reducing because the fixed costs shrink the profit gained from the FTA formation. Furthermore, if we consider the productivity difference described above, only firm O complies with the RoO and uses the FTA tariff in equilibrium. However, Regime B will remain as an equilibrium outcome, and the qualitative nature of our results remains unchanged as long as the fixed costs are not very large.

5.3 Market in country H

For simplicity, we ignore the market in H, as the primary focus of this paper is on firms' export pricing strategies. When the post-FTA equilibrium is Regime NB, B, or NC, the effects of an FTA formation are the same as the baseline model, even if we consider the market in country H. This is because firm O's marginal cost remains $c - \Delta$, and the prices, consumer surplus, and profits earned in country H are independent of the FTA formation as long as the two markets are segmented and firms can make independent decisions in each market.

When the post-FTA equilibrium is Regime IR, however, an FTA formation changes the market outcomes in country H because it changes firm O's marginal cost from $c - \Delta$ to c. The cost increase of firm O decreases firm O's profits earned in country H, decreases consumer surplus in country H by increasing the equilibrium prices, and reduces the welfare in country H, although the profits of firm I in country H increase. These effects make a consumer-hurting, welfare-reducing FTA more likely, and they make firm O's input relocation less likely. Even if we take these additional effects into account, the qualitative nature of our results remains unchanged.

5.4 Both firms can procure inputs outside the FTA

We have assumed that only firm O can procure outside inputs. Given that the outside inputs are supplied in a perfectly competitive market, there will be a case in which both firms can procure outside inputs. In this case, firm I chooses whether to comply with the RoO. If firm I does, it either manipulates its output price or changes the sources of inputs from countries outside the FTA to countries inside the FTA when the VAC of the RoO is the binding constraint.

The modified model should share the basic properties with the baseline model, such as the effects of an increase in $\underline{\alpha}$ on the prices, profits, and total welfare. However, each firm's choice of regimes will be more complicated. For instance, suppose initially, the VAC is non-binding for both firms and each firm operates under Regime NB. Then, suppose that an increase in $\underline{\alpha}$ forces one of the two firms to manipulate its output price to meet the VAC. Because the price commitment of one firm increases the price set by the other firm, it becomes easier for the other firm to meet the VAC, and it may remain in the NB regime.

This implies that the cut-off levels for firm i ($i \in \{I, O\}$), $\alpha_i^{NB}(s_j)$, $\alpha_i^{IR}(s_j)$, and $\alpha_i^{NC}(s_j)$, depend on the rival's choice of regime, $s_j \in \{NB, B, IR, NC\}$ ($j \in \{I, O\}, j \neq i$). For instance, the above discussion suggests that $\alpha_i^{NB}(NB) < \alpha_i^{NB}(B)$ holds. If we have $\alpha_i^{NB}(NB) < \underline{\alpha} < \alpha_i^{NB}(B)$, the two firms make heterogeneous choices and one firm chooses Regime B and the other firm chooses Regime NB. Irrespective of whether one firm or both firms are under Regime B. However, an increase in $\underline{\alpha}$ will have qualitatively the same effects on the prices, profits, and total welfare. Therefore, the main results remain unchanged.

6 Conclusion

This study has revisited the welfare effects of an FTA when exporting firms must meet a VAC to comply with RoO. When the value-added threshold is neither very low nor very high, a firm manipulates its output price to satisfy the VAC. In this case, the RoO work as a price floor and soften product market competition. The resulting increase in prices can hurt consumers while benefiting firms. Because of this price-increasing effect, an FTA formation with RoO might hurt consumers, even though all trade within the FTA is tariff free. Moreover, the total impact of the conflicting effects may be negative if the required threshold is sufficiently high. Furthermore, even if all firms keep the input procurement from efficient countries and enjoy tariff-free trade, the similarity of goods and the degree of restriction of the VAC play core roles in explaining whether the FTA is harmful or not.

These results suggest that a VAC of RoO can transform a consumer-benefiting (or welfare-improving) FTA into a consumer-hurting (or welfare-reducing) FTA. The RoO of FTAs should be designed such that they do not induce firm manipulations of output prices. One policy option is to employ the net-cost method, by which a calculation of the value-added ratio is unrelated to output prices.

There is room for further research. This study has assumed that the prices of inputs are exogenously given, but it would be interesting to consider how the VAC changes input prices. Another possible extension is to consider firms' location choices of inside or outside the FTA. The development of empirical analysis is left as the most important future task.

Appendixes

Proof of Proposition 3

By (6), both p_O^B and p_I^B are increasing in $\underline{\alpha}$, while p_O^N and p_I^N are independent of $\underline{\alpha}$. We have

$$\begin{aligned} p_O^B - p_O^N &= \frac{\left(4-b^2\right)(c-\Delta) - (1-\underline{\alpha})\left\{(2+b)(a+\tau) + 2(c-\Delta) + bc\right\}}{(4-b^2)(1-\alpha)}, \\ p_I^B - p_I^N &= \frac{b\left(4-b^2\right)(c-\Delta) - (1-\underline{\alpha})\left\{(2+b)(ab+2\tau) + 2b(c-\Delta) + b^2c\right\}}{2\left(4-b^2\right)(1-\alpha)}. \end{aligned}$$

By these equations, $p_O^B > p_O^N$ and $p_I^B > p_I^N$ hold if and only if

$$\alpha > \alpha' \equiv 1 - \frac{\left(4-b^2\right)(c-\Delta)}{(2+b)(a+\tau)+2(c-\Delta)+bc} \text{ and}$$
$$\alpha > \alpha_B^* \equiv 1 - \frac{b\left(4-b^2\right)(c-\Delta)}{(2+b)(ab+2\tau)+2b(c-\Delta)+b^2c}$$

hold, respectively.

Because we have $\alpha_B^* - \alpha' = \frac{\left(4-b^2\right)(c-\Delta)\tau}{\{(2+b)(a+\tau)+2(c-\Delta)+bc\}\{(2+b)(ab+2\tau)+2b(c-\Delta)+b^2c\}} > 0$, both $p_O^B > p_O^N$ and $p_I^B > p_I^N$ hold if and only if $\alpha > \alpha_B^*$ holds. There is a case where this cut-off level, α_B^*, satisfies $\alpha_B^* < \min[\alpha^{NC}, \alpha^{IR}]$, implying a price-increasing FTA when Regime B becomes the equilibrium outcome. For instance, we know that $\alpha^{\max}$ always satisfies $\alpha^{NB} < \alpha^{\max} < \min[\alpha^{NC}, \alpha^{IR}]$. If we compare α_B^* and $\alpha^{\max}$, we have

$$\alpha^{\max} - \alpha_B^* = \frac{b^3\{(2+b)a+bc-(2-b^2)(c-\Delta)\} - 4(2+b)(2-b^2)\tau}{\{(2+b)(ab+2\tau)+2b(c-\Delta)+b^2c\}\{(2+b)a+bc+(2-b^2)(c-\Delta)\}}.$$

Thus, as long as the tariff satisfies $\tau < \frac{b^3\{(2+b)a+bc-(2-b^2)(c-\Delta)\}}{4(2+b)(2-b^2)}$, $\alpha^{\max} > \alpha_B^*$ holds, and we have $\alpha_B^* < \min[\alpha^{NC}, \alpha^{IR}]$ because $\alpha^{\max} < \min[\alpha^{NC}, \alpha^{IR}]$ always holds. ■

Proof of Proposition 5

The welfare comparison between Regime IR and N yields:

$$W^{IR} - W^N = \frac{2(1-b)(2+b)^2\Gamma - \Delta\{2a(2+b)^2(3-2b) + (12-9b^2+2b^4)\Delta\}}{2(2-b)^2(2+b)^2}.$$

where $\Gamma \equiv \tau^2 + \{2(a-c+bc) + \Delta(1-b)\}\tau + \Delta(3-2b)c$. We can confirm that

$$(W^{IR} - W^N)|_{b=0} = \frac{2(\tau-\Delta)\{2(a-c)+\tau+\Delta\} - \Delta\{2(a-c)+\Delta-2\tau\}}{8},$$
$$(W^{IR} - W^N)|_{b=1} = \frac{-\Delta(18a+5\Delta)}{18} < 0, \text{and}$$
$$\frac{\partial^2(W^{IR} - W^N)}{(\partial b)^2} = \frac{-2(1+b)(2+b)^4\tau^2 + \Phi}{(2-b)^4(2+b)^4} < 0,$$

where $\Phi \equiv -2\tau\{2b^2(2b^3+15b^2+40b+40)c + 32(a-c) + (a-\Delta)b^2(2b^3+15b^2+20b+20)\} - 2\Delta\{3c(16+16b-8b^2-16b^3-7b^4-b^5) + 6b^2(7b^2+8) + 16(a-\Delta)\} + \tau\{-12\Delta(\tau-$

$\Delta) - 20\Delta\tau - 6ab^2\tau(b^2+8b+24)\} - 2ab\{96\tau - (40b+40b^2+15b^3+2b^4)\Delta\} < 0 \quad (\because \tau > \Delta)$.

The sign of $(W^{IR} - W^N)|_{b=0}$ is ambiguous, and it is increasing in τ. $(W^{IR} - W^N)|_{b=0} > 0$ holds if and only if $\tilde{\tau} \equiv \sqrt{\{2(a-c)+\Delta\}\{2(a-c)+7\Delta\}} - \{2(a-c)+\Delta\} < \tau$ holds. In this case, there exists a unique level of b, $b_{IR}^W \in (0,1)$, such that $W^{IR} > W^N$ holds for $b < b_{IR}^W$, $W^{IR} = W^N$ holds at $b = b_{IR}^W$, and $W^{IR} < W^N$ holds for $b > b_{IR}^W$.

Moreover, $\frac{\partial W^{IR} - W^N}{\partial b}|_{b=0} \geq 0$ holds if and only if

$$\hat{\tau} \equiv \frac{\Delta(a+2c)}{2c-\Delta} < \tau.$$

By substituting $\tau = \hat{\tau}$ into $(W^{IR} - W^N)|_{b=0}$, we get

$$(W^{IR} - W^N)|_{\tau=\hat{\tau},\, b=0} = \frac{\Delta\{2(4c-\Delta)a^2 - 8(2c^2 - 4\Delta + \Delta^2)a + 8c^2 - 12\Delta c^2 + 14\Delta^2 c - 3\Delta^3\}}{8(2c-\Delta)^2},$$

which takes a minimum at $a = \tilde{a} \equiv \frac{2(c^2-4\Delta c+\Delta^2)}{4c-\Delta}$, and the minimum level is given by $(W^{IR} - W^N)|_{\tau=\hat{\tau},\, b=0,\, a=\tilde{a}} = \frac{\Delta(2c-\Delta)^2(18c-5\Delta)}{4c-\Delta} > 0$. This implies that $\tilde{\tau} < \hat{\tau}$ and $\frac{\partial W^{IR} - W^N}{\partial b}|_{b=0} < 0$ whenever $(W^{IR} - W^N)|_{b=0} < 0$ holds. Therefore, if $\tau \leq \tilde{\tau}$ holds, we have $W^{IR} \leq W^N$ for any $b \in [0,1)$. ■

Proof of Proposition 6

The welfare gains from the FTA formation under Regime NC are computed as

$$W^{NC} - W^N = \frac{\tau[2\{(4-3b^2-b^3)a - 2(2-b^2)b\Delta\} - 2(1-b)^2(2+b)^2 c + (4-3b^2-2b^3)\tau]}{(2-b)^2}.$$

We have

$$(W^{NC} - W^N)|_{b=0} = \frac{[2(a-c)+\tau]\tau}{8} > 0,$$
$$(W^{NC} - W^N)|_{b=1} = -\frac{(4\Delta+\tau)\tau}{18} < 0, \text{and}$$
$$\frac{\partial^2(W^{NC} - W^N)}{(\partial b)^2} = -\frac{\Theta\tau}{(2-b)^4(2+b)^4} < 0,$$

where $\Theta \equiv (2b^5+9b^4+64b^3+56b^2+96b+16)\tau + (4b^3+30b^2+86b+80)b^2c + 32(a-c) + 2b\{(48+56b+32b^2+9b^3+b^4)a - 2(20+b^2)b^2\Delta\} > 0$. Thus, there exists a unique cut-off level of b, $b^W_{NC} \in (0,1)$, such that $W^{NC} > W^N$ holds for $b < b^W_{NC}$, $W^{NC} = W^N$ holds at $b = b^W_{NC}$, and $W^{NC} < W^N$ holds for $b > b^W_{NC}$. ■

References

Cadot, O., Carrère, C., de Melo, J., Portugal-Pérez, A., 2005. Market access and welfare under free trade agreements: textiles under NAFTA. The World Bank Economic Review 19, 379–405.

Conconi, P., García-Santana, M., Puccio, L., Venturini, R., 2018. From final goods to inputs: the protectionist effect of rules of origin. American Economic Review 108, 2335–2365.

Demidova, S., Krishna, K., 2008. Firm heterogeneity and firm behavior with conditional policies. Economics Letters 98, 122–128.

Estevadeordal, A., Suominen, K., 2003. Rules of origin in the world trading system, in: document presented at the Seminar on Regional Trade Agreements of WTO (Geneva, 14 November).

Falvey, R., Reed, G., 1998. Economic effects of rules of origin. Weltwirtschaftliches Archiv 134, 209–229.

Francois, J., Hoekman, B., Manchin, M., 2006. Preference erosion and multilateral trade liberalization. The World Bank Economic Review 20, 197–216.

Goldberg, P.K., Pavcnik, N., 2016. The effects of trade policy, in: Bagwell, K., Staiger, R.W. (Eds.), Handbook of commercial policy. Elsevier. volume 1. chapter 3, pp. 161–206.

Hayakawa, K., Hiratsuka, D., Shiino, K., Sukegawa, S., 2013. Who uses free trade agreements? Asian Economic Journal 27, 245–264.

Hayakawa, K., Laksanapanyakul, N., Mukunoki, H., Urata, S., 2019. Impact of free trade agreement use on import prices. The World Bank Economic Review 33, 643–660.

Ishikawa, J., Mukunoki, H., Mizoguchi, Y., 2007. Economic integration and rules of origin under international oligopoly. International Economic Review 48, 185–210.

Jinji, N., Mizoguchi, Y., 2016a. Optimal rules of origin with asymmetric compliance costs under international duopoly. Journal of Industry, Competition and Trade 16, 1–24.

Jinji, N., Mizoguchi, Y., 2016b. Rules of origin and technology spillovers from foreign direct investment under international duopoly. Japan and the World Economy 40, 47–60.

Ju, J., Krishna, K., 2005. Firm behaviour and market access in a free trade area with rules of origin. Canadian Journal of Economics 38, 290–308.

Krishna, K., Krueger, A., 1995. Implementing free trade areas: Rules of origin and hidden protection. NBER Working Paper No. W4983.

Mukunoki, H., 2017. The welfare effect of a free trade agreement in the presence of foreign direct investment and rules of origin. Review of International Economics 25, 733–759.

Mukunoki, H., Okoshi, H., 2020. Tariff elimination versus tax avoidance: Free trade agreements and transfer pricing. Munich Discussion Papers.

Reuter, P., 2012. Draining development? Controlling flows of illicit funds from developing countries. Washington DC: The World Bank.

Takahashi, K., Urata, S., 2010. On the use of free trade agreements by japanese firms, in: Free Trade Agreements in the Asia Pacific. World Scientific, pp. 241–257.