RIETI Discussion Paper Series 19-E-099

Tariff Elimination versus Tax Avoidance: Free Trade Agreements and Transfer Pricing

MUKUNOKI, Hiroshi
Gakushuin University

OKOSHI, Hirofumi
University of Munich

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 19-E-099
December 2019

Tariff Elimination versus Tax Avoidance: Free Trade Agreements and Transfer Pricing[1]

MUKUNOKI, Hiroshi
Gakushuin University

OKOSHI, Hirofumi
University of Munich

Abstract

Free trade agreements (FTAs) entail rules of origin (ROO), which require exporters to identify the origin of exports to prove eligibility for preferential tariff rates. This paper investigated how a multinational enterprise (MNE) in an international oligopoly model reacts to an FTA with ROO when it can manipulate its transfer price for intra-firm trade. Before the formation of an FTA, the MNE uses the transfer price to avoid a high corporate tax or to shift profits from the rival firm in the final-goods market. After the FTA formation, ROO can force the MNE to set the transfer price such that it meets the value-added requirement of ROO, or to change the location of its input production. We show that an FTA with ROO may decrease the profits of both the MNE and the local firm, even when they comply with ROO and take advantage of the tariff elimination provided for in the FTA. Furthermore, there is a case where ROO increase the consumer's gains from an FTA and transform a welfare-reducing FTA into a welfare-improving FTA.

Keywords: rules of origin, free trade agreement, transfer pricing
JEL classification: F13, F15, F23, H26

[1]This study is conducted as a part of the Project "Analyses of Offshoring" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). We thank Jay Pil Choi, Taiji Furusawa, Andreas Haufler, Jung Hur, Jota Ishikawa, Yoshimasa Komoriya, Ngo Van Long, Kiyoshi Matsubara, Eiichi Tomiura, and Makoto Yano for their helpful comments and suggestions. We are especially grateful to Martin Richardson for detailed discussions and suggestions at RIETI 2018 International Workshop "Frontiers in Research on Offshoring". Hiroshi Mukunoki acknowledges financial support from the JSPS KAKENHI Grant Numbers JP16H03620, JP17K03706. Hirofumi Okoshi acknowledges financial support from Deutsche Forschungsgemeinschaft (German Science Foundation, GRK1928). The usual disclaimer applies.

1 Introduction

For over half a century, trade liberalization has progressed all over the world. Although multilateral liberalization under non-discrimination principle played a major role in early years, the trend of trade liberalization has shifted to a formation of regional trade agreements (RTAs) and they have been a major driving force to eliminate or reduce trade barriers among countries in recent years. As of November 2019, 302 RTAs are in force.[1] Understanding the effects of RTAs has been one of the most important policy issues, and many papers have investigated them both theoretically and empirically.[2]

Trade liberalization usually lowers consumer price and raises export price, benefiting consumers and exporters. We expect that RTAs have the same effects for member countries, but the preferential nature of RTAs may make their effects more complicated that they seem because they have specific rules to implement trade agreements. Among others, setting rules of origin (ROO) is indispensable to form a free trade agreement (FTA), and it affects exporting firms' strategies such as their input procurement or locations choices. Conconi et al. (2018) concluded that ROO stipulated in NAFTA reduced imports from the non-member countries to be qualified for tariff-free trade, which indicates ROO cause inefficiency in input procurement. When firms in member countries make tariff-free exports to other member countries, ROO require the firms to prove that the exported products are originated within the FTA.[3] One way to prove the origin is to satisfy a value added (VA) criterion, which is closely related to the market outcome.[4] The VA criterion requires firms to add a sufficient value inside FTA member countries. Specifically, let p denote the export price of the product and r denote the value of input materials, which are used per unit of final-good production and not originating in the FTA. Then, a VA criterion typically requires that the value-added ratio, $(p-r)/p$, is larger than the specified level. This method of calculating the value-added content is called the "transaction value method." Estevadeordal and Suominen (2003) reported that among 87 FTAs they analyzed, 68 FTAs employed this method, at least in a particular product category.

When an intra-firm trade between related companies arises, multinational enterprses (MNEs) that operate these companies freely determine the price on the intra-firm trade. The MNEs' pricing

[1] See `http://rtais.wto.org/UI/PublicAllRTAList.aspx`.

[2] See Freund and Ornelas (2010) for the review of the literature on RTAs.

[3] Unlike custom union, member countries of an FTA are able to set their own tariff schedule against non-member countries, which provides an opportunity for firms producing outside the FTA to save tariff payment by choosing the member country whose tariff against the non-member countries is low as a transit country and re-exporting from the country to other FTA member countries whose tariffs against the non-member countries are higher. See or example, Stoyanov (2012) for the evidence on firms' incentive to transship the good going through the FTA members. To forestall firms from tariff avoidance, WTO stipulates ROO.

[4] Other ways to prove the origins of products include change in tariff classification criterion and specific process criterion. Although the effects of these criteria are also important, this paper focus only on VA criterion.

on the intra-firm trade has been usually argued in the context of tax avoidance. Namely, the related entity is able to save overall corporate tax payments by shifting profits via price manipulation on the traded products. This price manipulation is called transfer pricing and the manipulated price is called transfer price. Some empirical researches have provided the evidences on transfer pricing to save tax payments.[5]

In the presence of FTAs, MNEs need to consider a different effect of transfer pricing: MNEs cannot set a high transfer price to be eligible for tariff elimination in FTAs. Specifically, if MNEs import inputs from their related companies to produce the good inside an FTA, they must set their transfer prices such that they can meet the VA criterion of ROO. In other words, the VA criterion of ROO prevents MNEs from using their transfer prices solely for tax avoidance motive. Although this possibility has been overlooked in the economic literature of transfer pricing and that of FTA, it has been pointed out by some policy papers. For instance, Eden (1998) examined ROO of NAFTA and suggested that "... underinvoicing parts coming outside North America and overinvoicing locally made parts would increase the North American content." Falvey and Reed (1998) pointed out that the VA criterion "... allows room for manipulation of prices as well as quantities, and may generate additional incentives for transfer pricing by multinationals." Reuter (2012) 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."[6] Furthermore, the World Customs Organization suggests that one of the demerits of the VA criterion of ROO is possible exposure to transfer pricing.[7] These statements tell us that investigating a role of transfer price in complying with ROO is important.

Transfer pricing to comply with ROO also suggests the importance of policy coordination. Both designing trade policies and regulations on transfer pricing are important issues that policy makers pay attentions to, but the link between them are rarely discussed. This is because customs are responsible for tariff issues while tax authorities are responsible for tax avoidance by MNEs, and there are no interaction between them. Recently, collaboration between the two institutes begins by sharing transfer price documents handed in tax authorities. According to WCO (2018), "... the WCO is working with the OECD and World Bank Group to encourage Customs and tax administrations

[5]For instance, Swenson (2001), Clausing (2003) and Cristea and Nguyen (2016) provide empirical evidences of transfer price manipulation. Blouin et al. (2018) found that the conflicting motives of transfer price when MNEs use it for corporate tax saving and also for tariff saving.

[6]Some practitioners see the link as one factor to be considered and state that "If transfer pricing changes the value of local content, then the ROO as applied may remove any FTA benefit that was previously available" (see `https://www.expertguides.com/articles/oecd-beps-project-and-trade-new-perspectives/AREXIEUO`, accessed on May 03, 2018).

[7]See `http://www.wcoomd.org/-/media/wco/public/global/pdf/topics/origin/overview/origin-handbook/rules-of-origin-handbook.pdf`, accessed on May 3, 2018

to establish bilateral lines of communication in order to exchange knowledge, skills and data, where possible, which will help ensure that each authority has the broadest picture of a MNE's business, its compliance record and can make informed decisions on the correct revenue liability." Thus, as the number of FTAs and volume of intra-firm trade have increased, exploring the relationship between transfer price and ROO is an urgent issue.

1.1 Preview of the model and the results

Against this backdrop, this paper builds an international duopoly model to investigate an MNE's response to an FTA formation with two new elements: transfer pricing and ROO. The MNE produces a final good within an FTA member country and exports the good to the other FTA member country while the MNE's location of the input production is either in the final good production country or in a low tax country outside the FTA. The MNE can shift profits across countries by manipulating transfer price if it decides input production outside the FTA countries. The MNE competes with the local firm, which also produces the final good within the FTA and exports its product to the same country.

In the absence of ROO, the MNE prefers to locate its upstream and downstream affiliates in different countries when the tax gap is large. This is because larger tax differential increases the MNE's gains from tax savings. The decision makings of the two affiliates are centralized in this case. When the tax gap is small, the MNE prefers to locate both affiliates in the same country and the decision making of the downstream affiliate is decentralized. Then, the upstream affiliate sets a low input price to make the downstream more competitive in the product market. This is in line with Schjelderup and Sorgard (1997), which showed that, when the decisions of the headquarters and foreign affiliates are decentralized and those affiliates compete with rival firms in the product market, MNEs may use its transfer price as a strategic tool to shift rents from the rival firms. We show that the formation of an FTA can induce the MNE's input production relocation from a country outside the FTA to a country inside the FTA. A notable result is that an FTA formation with the MNE's input relocation may hurt the local firm even though the local firm's export is always subject to tariff elimination of the FTA. This is because the loss from the strategic effect of transfer pricing outweighs the gain from tariff elimination for the local firm.

In the presence of ROO, the MNE chooses one of the three strategies: (i) producing inputs inside an FTA to comply with ROO, (ii) manipulating transfer price to reduce tax payments (i.e., for tax avoidance), and (iii) manipulating transfer price to comply with ROO (i.e., for tariff elimination). If

the MNE produces the input in the same country where its downstream affiliate locates, the MNE always complies with ROO and export its product without tariff, though it cannot save tax payments by transfer pricing. If the MNE locates its upstream affiliate in a country outside the FTA, the MNE either saves corporate tax payments by setting a high transfer price or complies with ROO to avoid tariff burden by setting a low transfer price. Therefore, this model exhibits the MNE's choice of "tariff elimination versus tax avoidance" via its input procurement strategies and/or transfer price manipulation.

Besides the above two options, the MNE may manipulate transfer price to meet the VA criterion of ROO, rather than to save tax payments. In this case, the MNE enjoys tariff elimination but faces larger tax payments.[8] Remarkably, when the MNE optimally chooses the strategy of manipulating transfer price for complying with ROO, its post-tax profits can decrease with a formation of FTA. Furthermore, unlike the case without ROO, an FTA formation may reduce the profits of both the MNE and the local firm, even though both of them comply with ROO and they make tariff-free exports within the FTA. This happens when the MNE relocates its upstream production to an inside FTA country. The MNE's (post-tax) profit decreases because it no longer save tax payments, while the local firm's profit decreases because the strategic effect of low intra-firm input price of the MNE intensifies the market competition.

We also demonstrates that imposition of ROO can be a devise to improve total welfare inside FTA countries even though it is at the expense of the local firm. As the MNE shifts tax base to an outside FTA country, the welfare effect of an FTA formation without ROO can be negative. However, ROO mitigate the outflows of tax base from the inside country. Therefore, a strict ROO can turn a welfare-reducing FTA into a welfare-improving FTA.

1.2 Relationship to the literature

The welfare effects of FTA with ROO have been previously analyzed by some papers, but their focuses are mainly on intermediate goods markets. Krishna and Krueger (1995) showed that ROO may work as a hidden protection against the input suppliers outside the FTA. Ju and Krishna (2005) showed that ROO increase the price of FTA-made inputs and reduce the total output if ROO are not so stringent such that all firms comply with ROO, but they have the opposite effects if ROO are sufficiently stringent such that some firms choose not to comply with ROO. In Ju and Krishna (2005), however, the price of the output is fixed and they did not consider how ROO affect

[8]Even though the MNE uses its transfer price for ROO compliance, it can still shift profits from one country to another to save tax payments when the VA requirement is not so stringent and the tax gap is large. Nevertheless, the overall tax payments become larger because the level of transfer price is sub-optimal from the viewpoint of tax savings.

consumers. Demidova and Krishna (2008) extended Ju and Krishna (2005) to include productivity heterogeneity of final-good producers and showed that productivity sorting ensures the negative relationship between the stringency of ROO and the demand for FTA-made inputs (i.e., wages). Ishikawa et al. (2007) focused on final good markets and showed that ROO have a role to segmented the markets within FTA and both inside and outside firms producing final goods may benefit from ROO at the cost of consumers. Mukunoki (2017) showed that FTA with ROO may be consumer-hurting if it changes the outside firms' location decisions. None of these papers, however, have considered transfer price manipulation to meet ROO. A companion paper of ours, Mukunoki and Okoshi (2019), investigates a firm's export price manipulation for complying with ROO. The focus of this paper is closer but apparently different because this paper investigates an MNE's transfer price manipulation on inputs imported from outside FTA.

There exist some papers that investigated the relationship between transfer pricing and trade barriers including tariffs. Horst (1971) showed that optimal transfer price is influenced by not only tax differentials but also by tariffs. Schjelderup and Sorgard (1997) showed that if the importing country imposes an ad valorem tariff on inputs, an MNE can save tariff payments by reducing its export price.[9] Then, the optimal transfer price is influenced by both corporate tax avoidance and tariff avoidance. Kant (1988) regards transfer price as a tool to repatriate profits when a foreign subsidiary is not fully owned by the parent firm. With a partial ownership of the foreign affiliate, a profit shifted from home to foreign country is partly distributed to the other owners. The paper found that even when the tax rate in home country is higher than the tax rate in the host country, an MNE has an incentive to remit all the profit earned in the low-tax host country, if both tariff and the proportion of the MNE's ownership shares in the foreign affiliate are low. These papers, however, did not explicitly consider trade liberalization by forming an FTA, let alone the effects of ROO on transfer prices.

This model also contributes to the literature on policies on transfer pricing since MNEs have been accused of tax avoidance activities and how to regulate transfer prices has been one of the central issues in policy debates. Several papers have examined the impacts of policies on transfer price manipulation. Elitzur and Mintz (1996) investigated the determinant of transfer price when tax authorities use cost-plus method to infer appropriate transfer price. Nielsen et al. (2003) compared the use of transfer price under two international tax systems (i.e., separating account (SA) versus formula apportionment (FA)[10]. Choi et al. (2018) examined the impact of arm's length principle

[9]Given multiple roles of transfer prices, recent papers have examined the optimal MNEs' strategies (Hyde and Choe, 2005; Nielsen et al., 2008; Dürr and Göx, 2011). None of them, however, link transfer pricing and ROO.

[10]Traditionally prevailing international corporate tax system is SA which computes MNEs' national tax base by re-

(ALP), by which MNEs should set the same price for intra-firm transactions as the price of the same transaction conducted between independent firms.[11] As their focus was on such a direct regulation on transfer pricing, transfer pricing for meeting ROO has been overlooked in the literature.

The rest of paper is organized as follow. In Section 2, we set up a model. Section 3 derives the equilibrium and analyze the effect of FTA without ROO. Section 4 investigates the effect of FTA with ROO. Section 5 discusses the robustness of the main results by relaxing some key assumptions. Section 6 summarizes and concludes the paper.

2 Model

We consider a three-country model with two firms, an MNE (firm M) and a local firm (firm L).[12] The MNE has a downstream affiliate (firm M_D) that produces a final good and an upstream affiliate (firm M_U) that produces inputs used for the production of firm M_D. The two of the three countries are potential FTA member countries, while the rest is the outside, non-member country (country O). The MNE is owned by a country other than these three countries. The model is illustrated in Figure 1.

We assume the two downstream firms, firm M_D and firm L, locate in one of the member countries, which is referred to as the host country (country H), because the country has location advantages to attract firms, such as low factor prices, a large pool of skilled labors, and so on. The two downstream firms produce homogeneous goods and serve them to consumers in a foreign member country (country F), which is a potential member of FTA with country H.[13] Without FTA, country F imposes a specific tariff, τ, on imports of final goods.[14] The governments in country O and H respectively impose t and T as a corporate tax on reported profits.[15,16] In order to focus on the

garding intra-firm transaction as inter-firm transaction. On the other hand, alternative system is FA under which MNEs' tax payments to one country depend on its consolidated tax base and the proportion of activity operated in the country. See more detail in European Commission (2011) chapter XVI article 86.

[11] Bauer and Langenmayr (2013), Choe and Matsushima (2013) and Kato and Okoshi (2019) have also investigated the effect of ALP on input procurement decision, tacit collusion, and input production location, respectively.

[12] This type of foreign direct investment (FDI) is known as export-platform FDI whose feature is to export from the host country to neighbour countries to avoid high trade cost from the origin country. For example, see Tekin-Koru and Waldkirch (2010) for the Mexican evidence of an increasing role as an export platform. Tintelnot (2017) also shows the share of output exported to countries outside the host country by U.S. MNEs. The share of Belgium was 63% in 2004, which was the third highest share.

[13] For simplicity, we ignore the output market in country H in the main analysis. This assumption does not qualitatively change our main results if the two markets are segmented. We relax this assumption in section 5.1.

[14] We focus on the case where both firms always supply their products in country F. This requires both a tariff rate and corporate tax rates are low. Specifically, exports by both the MNE and firm L become positive when τ satisfies $\tau < \min\left\{\frac{a-w+2\Delta}{2}, a-w-\left(\frac{1-t}{1-T}\right)\Delta\right\}$. Besides that, firm L's exports are positive when $T \leq T^{max} \equiv 1-\frac{(1-t)\tau}{a-w-\tau} < 1$ holds.

[15] Note that we use the term "tax rate" and "tax revenue" to represent corporate tax rate and corporate tax revenue, respectively. These tax rate and tax revenue are distinguished from tariff rate and tariff revenue.

[16] In this model, we postulate that both government in country O and H adopt territorial tax system instead of world-

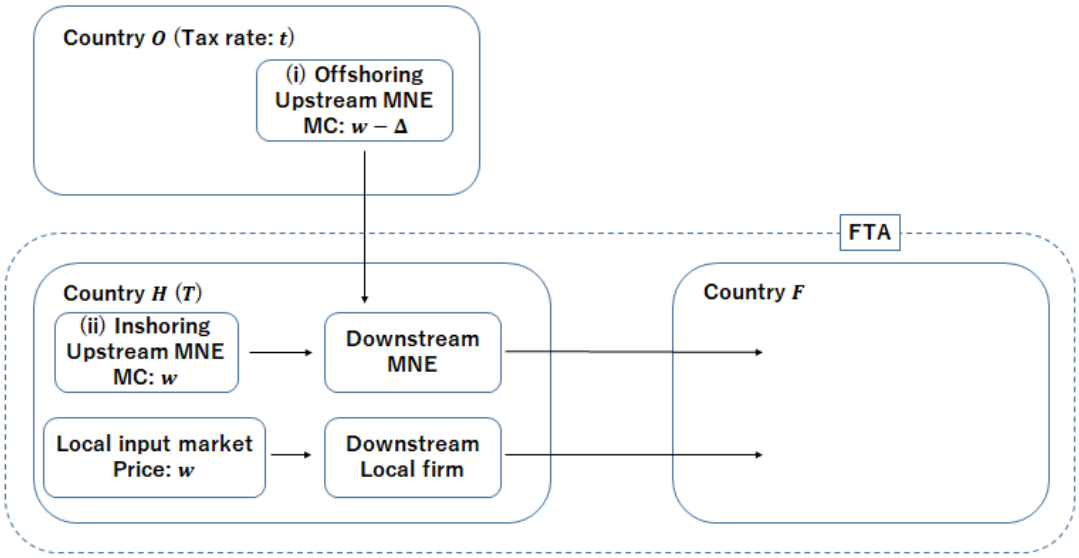

Figure 1: Model

impact of FTA formation on the competition in the final-good market, tariffs on inputs are assumed away. Hereafter, we focus on the case with $T \geq t$, with which main findings are obtained.[17]

For the consumer side, the representative consumer's utility in country F is given by $U = a(x_L + x_M) - (x_L + x_M)^2/2$, where x_i is the consumption of the final good produced by firm i ($i \in \{L, M\}$). By the utility maximization, the inverse demand function becomes $p = a - (x_L + x_M)$, where p is the price of the final good.

To produce the final goods, the two firms need to procure inputs made in either country H or O. Firm L always procures inputs with the input price w from perfectly competitive input market in country H. However, firm M is able to locate its related upstream affiliate (firm M_U) either in country O or country H. We assume input production in country O is more efficient than that in H. Specifically, if firm M_U produces inputs in country H, its marginal cost is given by w, while its marginal cost is given by $w - \Delta$ if inputs are produced in country O. This implies that locating firm M_U in country O gives firm M_U not only a cost advantage over local input suppliers but also a tax saving opportunity. Without loss of generality, we assume that both downstream firms use the same production technology, where one unit of inputs is transformed into one unit of final products and

wide one. After the U.S. moved from worldwide tax system to territorial tax system, most of OECD countries adopt territorial tax system.

[17]This situation is consistent with the real-world observation. For instance, Mexico and Belgium have higher corporate taxes than other countries, and these countries are major host countries of export-platform FDIs. See also footnote 12.

the additional cost of final good production is constant and normalized to zero.[18,19] If the MNE locates firm M_U in country O, firm M_U exports the produced inputs to its downstream affiliate (firm M_D) by charging an intra-firm, transfer price denoted by r^O. Alternatively, if the MNE locates the upstream firm in country H, we assume the decision makings of the MNE are decentralized. Namely, the downstream firm (firm M_D) maximizes its own profits, while the objective of the upstream affiliate is to maximize the total profits of upstream and downstream affiliates.[20] Thus, even though the two affiliates locates in the same country, firm M manipulates intra-firm price r^I for its strategic purpose.

This organization structure of the MNE, namely a centralized versus a decentralized decision, is a reasonable assumption under $T \geq t$. When the MNE produces both inputs and final goods in the same country, the MNE has an incentive to decentralize the quantity decision to make its downstream affiliate (firm M_D) more aggressive in the product market and shifts rents from the local firm. On the other side, when the MNE separately locates their production, decentralized decision is not profitable since setting a high transfer price to avoid high tax of the host country the increases procurement cost for firm M_D, which worsens its performance in the product market. Nielsen et al. (2008) demonstrates that centralization is more profitable than decentralization when input production country sets the lower corporate tax than the final production country and the tax gap is large. When the tax gap is small, decentralization brings the higher profits for the MNE.

We solve the following three-stage game. In the first stage, the headquarters of the MNE decides the location of firm M_D. In the second stage, the headquarters determines the optimal transfer price. In the third stage, the MNE and firm L compete à la Cournot in country F. As briefly mentioned above, we assume that the MNE's decision on quantity is based on centralization when the headquarters prefers country O as the input production country while on decentralization when it chooses country H as the input production country. Thus, the MNE faces a trade-off in the location of its input production: the MNE is able to use the transfer price to save tax payments in producing inputs in country O, while it is able to use the transfer price to take advantage of a strategic effect of decentralization in producing them in country H.

[18]We assume only the MNE has an option to procure inputs from country O, even though both firms produce with the same technology. This assumption is supported by empirical evidences that some firms engage in global production such as outsourcing and foreign direct investment while the others do not even though they have similar productivity. For example, see Tomiura (2007).

[19]We can consider a more general situation where the MNE uses a continuum of inputs and decides the extent to which it uses the intra-firm inputs for final good production. As is explained in Section 5.2, this modification does not change the qualitative results of the benchmark model.

[20]The decentralization gives the MNE a strategic effect, which is known as "managerial advantage" in the literature of Industrial organization. For example, Fershtman and Judd (1987) and Sklivas (1987) showed that the firm owners have incentives to give their managers a right of decision making on the product market. In the literature of transfer pricing, Schjelderup and Sorgard (1997) first pointed out the strategic motive of transfer pricing manipulation.

3 The equilibrium without ROO

In this section, we derive the pre-FTA equilibrium and the post-FTA equilibrium in the absence of ROO. In the next section, we introduce ROO and derive the post-FTA equilibrium with ROO.

3.1 Market equilibrium

Let us first derive the market equilibrium in the last stage. The unit cost of firm L in producing a final good and exporting it to country F is given by $c_L = w + \lambda_L \tau$ where λ_i ($i \in \{L, M\}$) is a state variable which takes zero if firm i is qualified for FTA tariff after the FTA formation and takes unity otherwise. Therefore, firm L maximizes the following (pre-tax) profit,

$$\pi_L = (p - c_L)x_L. \tag{1}$$

There are two schemes as the MNE's input sourcing: (1) Offshoring scheme in which the MNE produces inputs in country O (Regime O) and (2) inshoring scheme in which the MNE produces inputs in country H (Regime I).

Offshoring scheme When the MNE locates its input production in country O, the MNE centralizes its decision makings and determines the amount of supply to maximize the following global post-tax profit:

$$\begin{aligned}\Pi_M^O &= (1-t)(r^O - (w-\Delta))x_M^O + (1-T)(p - r^O - \lambda_M\tau)x_M^O \\ &= (1-T)\left(p - \underbrace{\frac{(1-t)(w-\Delta) + (1-T)\lambda_M\tau - (T-t)r^O}{1-T}}_{c_M^O=\text{Percieved marginal cost}}\right)x_M^O.\end{aligned} \tag{2}$$

As seen in the above expression, the centralized MNE's behaviour is based on the *perceived marginal cost*, which is different from the sum of the input production cost and trade cost $w - \Delta + \lambda_M\tau$.[21] In this cross-border production, the global unit cost is adjusted by the tax differential. As a marginal increase in the transfer price, r^O, saves the per-unit tax payments as much as $(T - t) > 0$, it reduces effective marginal cost of firm M_D in the production of the final good. Therefore, the "perceived marginal cost" becomes lower and the MNE supplies more as r^O becomes higher. Note that the perceived marginal cost under offshoring scheme c_M^O is equivalent to $w - \Delta + \lambda_M\tau$ only if $T = t$

[21]The terminology "perceived marginal cost" is often used in the analysis of vertically related industry in the context of industrial organization. See Choi et al. (2018) for an application of this terminology into tax avoidance literature.

holds and is decreasing in T and increasing in t. This means the perceived marginal cost is less than true marginal cost $c_M^O \leq w - \Delta + \lambda_M \tau$ when $T > t$ holds. In other words, transfer pricing makes the MNE more aggressive in the product market under input offshoring.

Inshoring scheme When the MNE locates its input production in country H, the MNE's decision makings are decentralizated. Thus, how much to produce the final good is delegated to the manager of the downstream affiliate (i.e., firm M_D), who only takes into account the profit of firm D, which is given by:

$$\pi_D = (p - \underbrace{(r^I + \lambda_M \tau)}_{c_M^I}) x_M. \tag{3}$$

For expositional convenience, we denote c_M^I as the perceived marginal cost under inshoring scheme since the output decision by firm M_D is based upon c_M^I, whereas the true marginal cost of the MNE is $w + \lambda_M \tau$.

Given Eqs. (1) to (3), we can derive the optimal supply of each firm by solving the first-order conditions of profit maximization:

$$x_M^s = \frac{a - 2c_M^s + c_L}{3}, \quad \text{and} \quad x_L^s = \frac{a - 2c_L + c_M^s}{3}, \qquad s \in \{O, I\}. \tag{4}$$

The equilibrium profit of each firm is given by:

$$\pi_L^s = (x_L^s)^2 \quad \text{and} \quad \Pi_M = \begin{cases} (1-T)(x_M^O)^2 \\ (1-T)\left(r - w + x_M^I\right) x_M^I \end{cases}.$$

The equilibrium consumer surplus in country F is given by

$$CS_F^s \equiv \frac{\left(x_L^s + x_M^s\right)^2}{2}. \tag{5}$$

3.2 Manipulation of the transfer price

Next, we consider how the MNE sets transfer price in the second stage. As described above, there are two cases in which the MNE manipulates transfer price. We derive the optimal level of transfer price separately in these cases.

Offshoring scheme Given Eqs.(2) and (4), the overall profit of the MNE is

$$\Pi_M^O = (1-T)\left(\frac{a+w-2r^O-(2\lambda_M-\lambda_L)\tau}{3}+\frac{2(1-t)(r^O-(w-\Delta))}{3(1-T)}\right)^2. \tag{6}$$

Since the first derivative of Π_M^O with respect to r is always positive, the optimal transfer price is as high as possible, that is, such that $p-\hat{r}^O-\lambda_M\tau=0$ holds;[22]

$$r^{O*} = w-\Delta+\underbrace{\frac{(1-T)\{a-w+2\Delta-(2\lambda_M-\lambda_L)\tau\}}{(1-t)+2(1-T)}}_{\text{Tax avoidance motive}}. \tag{7}$$

The corresponding equilibrium output and profits are;

$$x_M^{O*} = \frac{(1-t)(a-w+2\Delta-(2\lambda_M-\lambda_L)\tau)}{3-2T-t}, \text{ and } \Pi_M^{O*} = (1-T)\left(x_M^{O*}\right)^2. \tag{8}$$

Inshoring scheme Given Eq. (4), the overall profit of the MNE is

$$\Pi_M^I = (1-T)\left[\{r^I-w+p-(r^I+\lambda_M\tau)\}\left(\frac{a+w-2r^I-(2\lambda_M-\lambda_L)\tau}{3}\right)\right]. \tag{9}$$

By differentiating (9) with respect to r^I, the optimal transfer price becomes;

$$r^{I*} = w\underbrace{-\frac{a-w-(2\lambda_M-\lambda_L)\tau}{4}}_{\text{Strategic motive}}. \tag{10}$$

The corresponding equilibrium output and profits are;

$$x_M^{I*} = \frac{a-w-(2\lambda_M-\lambda_L)\tau}{2}, \text{ and } \Pi_M^{I*} = \frac{(1-T)}{2}\left(x_M^{I*}\right)^2. \tag{11}$$

Depending on the input location, setting the optimal transfer price is led by different motives. In the case of offshoring, the optimal transfer price is greater than the marginal cost of producing inputs. The transfer price is set to shift profits from a high-taxed country H to a low-taxed country O. The second term of (7) represents a *tax avoidance motive*, whose sign is always positive.[23] On

[22]This is a conventional way of determining the optimal transfer price in the literature, when the cost for profit shifting is absent. We relax this assumption by introducing standard convex concealment cost in section 5.3.

[23]Note that the incentive gets stronger as the tax differential gets wider $\partial\tilde{r}^O/\partial T<0$ because the quantity decision, or the perceived marginal cost is a function of T. On the one hand, wider tax gap increases the MNE's incentive to shift profits for tax purpose, which is captured by the quantity decision. On the other hand, transfer pricing also captures the aspect of "first mover advantage". If the MNE sets a high transfer price when the tax differential is wide, oversupply from the MNE's viewpoint arises because higher transfer price induces greater MNE's supply, which makes marginal operating profit much lower. Thus, the MNE's transfer pricing aim to report zero profit in high tax country and maximize the gains from first mover advantage.

top of that, Eq.(8) indicates that the MNE's output expands as corporate tax in country H, T, is higher. This is because the perceived marginal cost is lower as T is higher. As the output expansion decreases the equilibrium price of the good, p, the transifer price that realizes zero profits of firm M_D becomes lower.

In the case of inshoring scheme, on the other hand, the production and export decisions in the final-good market is delegated to firm M_D. Then, the MNE uses the transfer price to make firm M_D behave more aggressive in the product market by setting low transfer price. In other words, lowering the transfer price works as a "strategic intra-firm subsidy" and shifts rents from firm L to firm M. This is captured by the second term in Eq.(10), which is always negative. Unlike offshoring scheme, the optimal transfer price under in-sourcing is independent of T. Note that the perceived marginal cost is always lower than the marginal costs under both schemes.

3.3 Location choice of the input production

In the first stage, the MNE chooses between country O or H as the location of firm M_U. To distinguish between pre-FTA variables and post-FTA ones, we use asterisk "$*$" for the former while use hat "$\widehat{\ }$" for the latter. In the pre-FTA situation, the MNE prefers offshoring scheme to inshoring scheme if and only if:

$$\Pi_M^{O*} - \Pi_M^{I*} \geq 0 \iff T \geq T^* \equiv \frac{3-t}{2} - \frac{(1-t)(a-w+2\Delta-\tau)\sqrt{2}}{a-w-\tau}. \tag{12}$$

In the post-FTA situation without ROO, the MNE chooses offshoring scheme if and only if

$$\widehat{\Pi}_M^{O} - \widehat{\Pi}_M^{I} \geq 0 \iff T \geq \widehat{T} \equiv \frac{3-t}{2} - \frac{(1-t)(a-w+2\Delta)\sqrt{2}}{a-w}. \tag{13}$$

Intuitively, the MNE has a strong incentive to shift profits if corporate tax in country H is sufficiently high whereas it enjoys the strategic effect if country H's corporate tax rate is not so high.

It is easily verified that $T^* < \widehat{T}$ holds. This implies that an FTA formation without ROO changes the MNE's input sourcing strategy from offshoring to inshoring if $T^* < T < \widehat{T}$ holds. This is because the elimination of tariff magnifies strategic/tax avoidance motives of transfer pricing at different levels. Figure 2 illustrates the determination of the cut-off level of T and the MNE's location decision. The solid horizontal line and the solid curve represent the MNE's pre-tax profits under inshoring and offshoring, respectively after the FTA formation, while the dashed ones are those before the FTA formation. The solid line/curve is always above the dashed line because tariff

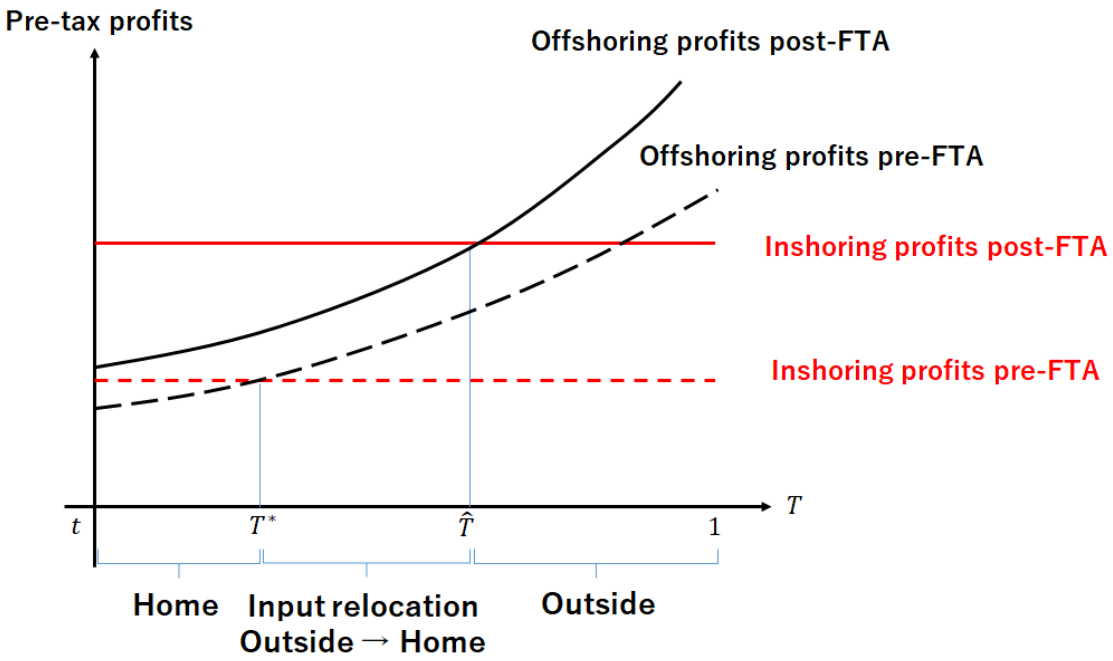

Figure 2: The MNE's Production

elimination increases the MNE's profits. An increase in T reduces the perceived marginal cost and raises the MNE's pre-tax profits under offshoring. The pre-tax profits under inshoring, on the other hand, are independent of T and remains constant. We can confirm that at $T = T^*$, the profit gap under inshoring is always larger than the profit gap under offshoring. Therefore, $T^* < \widehat{T}$ holds. We have the following proposition.

Proposition 1. In the absence of ROO, the MNE locates its upstream affiliate (i) in host country before and after an FTA formation if $T < T^*$ holds, (ii) in outside FTA country before and after the FTA formation if $\widehat{T} < T$ holds, and (iii) in outside FTA country before the FTA formation but in host country after the FTA formation if $T^* \leq T \leq \widehat{T}$ holds.

This proposition gives two new implications for empirical result. First, some empirical papers show an increase in FDI inside FTA countries after an FTA formation. For instance, Hanson et al. (2005) empirically shows a substantial increase in investment in Mexico and Canada after North America Free Trade Agreement (NAFTA). Our result suggests that the increase in FDI might be caused by input relocation of MNEs to exploit the magnified strategic effect of transfer pricing.

Second, the proposition indicates a new possibility of trade diversion effect. Traditionally, trade diversion effect captures a substitution effect between imports from member countries and those from non-member countries caused by preferential elimination of tariffs on imports from member countries. In our model, however, even in the absence of ROO, reduction in tariff on final product

may affect the MNE's input sourcing strategy. One notable thing is that this diversification effect is inefficient in the sense that input production is relocated from more efficient country whose production cost of the input is $w - \Delta$ to less inefficient country whose production cost is w.

3.4 The welfare effects of FTA formation without ROO

Despite the inefficiency of input production, we have confirmed that an FTA formation always favors consumers and the MNE. For consumers, the positive effect of tariff elimination dominates the negative effect of inefficient input production. For the MNE, the gains from the magnified strategic effect exceeds the loss of tax saving opportunity. However, firm L can be negatively influenced by the FTA if $T^* < T < \widehat{T}$ holds since the input relocation makes the MNE more aggressive in the product market and x_L^O becomes larger. Indeed, there exists a threshold T_L such that $\widehat{x}_L^I = x_L^{O*}$ holds. For $T < T_L$, $\widehat{x}_L^I < x_L^{O*}$ and for $T > T_L$, $\widehat{x}_L^I > x_L^{O*}$. We have $T_L \in (T^*, \widehat{T})$ if $\tau > \tau_L$ holds while $T_L > \widehat{T}$ holds if $\tau < \tau_L$. Thus, the FTA without ROO that accompanies the input relocation benefits firm L if and only if $\tau > \tau_L$ and $T_L < T < \widehat{T}$ hold. Otherwise, the FTA that comes with input relocation of the MNE hurts firm L, even though firm L is subject to tariff elimination of the FTA. Since x_L^{O*} gets smaller as tariff gets larger due to a direct impact on firm L, when the initial tariff is large enough, the gains from tariff elimination can be positive even if the MNE's input location changes. The following proposition summarizes the result.

Proposition 2. In the absence of ROO, an FTA formation always benefits consumers in a foreign country and the MNE. However, it hurts the local firm if the MNE changes the location of input production and $T^* < T < \min[T_L, \widehat{T}]$ hold. We have $T_L < \widehat{T}$ if $\tau > \tau_L$ holds and $T_L \geq \widehat{T}$ if $\tau \leq \tau_L$ holds. Otherwise, an FTA formation benefits the local firm.

We have shown that an FTA formation without ROO may hurt the local firm. FTA formation also changes both tariff revenues and tax revenues of member countries. Our next question is whether these effects are consistent with countries' incentives to form an FTA. We suppose that countries are able to arrange transfers of welfare between member countries upon FTA formation by making mutual concessions in other sectors, for instance. This means that a formation of an FTA is feasible if it improves the total welfare of member countries.

We exclude the MNE's profits from welfare in country H since the MNE is owned by residents of another country outside the model. We denote the MNE's taxable profits in country in

scheme/regime s ($s \in \{I, O\}$) as,

$$\pi_M^s = \begin{cases} (p - r - \lambda_M \tau) x_M^s, & \text{if} \quad s = O \\ \{r - w + p - (r + \lambda_M \tau)\} \left(\frac{a + w - 2r - (2\lambda_M - \lambda_L)\tau}{3} \right), & \text{if} \quad s = I. \end{cases} \tag{14}$$

Then, the equilibrium welfare of country H and F in scheme/regime s is respectively given by,

$$W_H^s \equiv \pi_L^s + T\pi_M^s, \tag{15}$$

$$W_F^s \equiv CS_F^s + TR_F^s, \tag{16}$$

where $TR_F^s \equiv \tau(\lambda_L x_L^s + \lambda_M x_M^s)$ is the equilibrium tariff revenue in country F. Then, the total welfare of the FTA countries is $W_{FTA}^s \equiv W_H^s + W_F^s$.

When the MNE produces inputs in country H regardless of FTA formation ($T < T^*$), FTA increases consumers' gains, the profit of the local firm, and tax revenue from the MNE while tariff revenue no longer exists. When the initial tariff is large, then the positive effects dominate the negative one. However, the negative effect can excel the other positive effect when the initial tariff is sufficiently low. When $\tau < \frac{2(a-w)}{13}$, the total net effect of firm L, consumers, and tariff revenue is negative. However, as the MNE pays tax in country H, the net negative effect can be covered by the increase in the MNE's tax payment. Since the increase in the tax revenue is greater as the corporate tax rate is higher, the overall net effect is positive when T is close to T^*.[24]

When $T^* < T < \widehat{T}$ holds, the MNE's input relocation takes place. As proposition 2 shows, firm L may lose from an FTA. However, the input relocation inflows tax revenue in country H after the FTA. Therefore, although there exist two negative impacts, positive impacts from an increase in consumers' gains and the collection of tax revenue cover the negative ones. Since an FTA formation without ROO hurts the local firm only in this range of T, countries have incentives to form an FTA that huts the local firm.

When the MNE does not change its input production country, namely $\widehat{T} < T$, the MNE always shift all the taxable profits in country H. From proposition 2, both firm L and consumers gains from FTA but no tariff revenue comes in. In this case, the loss of tariff revenue is larger as the tax gap is wider since the export by the MNE is greater due to the lower perceived marginal cost as T gets larger. Thus, when T is closer to $\widehat{T}$, the tariff revenue losses is minimized and can be covered by the positive impacts. As an increase in T creates three effects: larger losses of tariff revenue,

[24]Specifically, let $T_W^* \equiv \frac{2(a-w)-13\tau}{4\{2(a-w)-\tau\}}$ be a thresholds such that $W^{I*} = \widehat{W}^I$ holds. Then, there exist a unique threshold τ^* such that $t = T_W^*$ holds if $t < \frac{1}{4}$ holds. Thus, $W^{I*} > \widehat{W}^I$ holds when T lies in $[t, T_W^*)$ which arises when $\tau < \tau^*$ holds.

less profit of firm L, and more gains for consumers due to the MNE's aggressive behaviour, which effect dominates varies over T. An increase in T magnifies the negative effect the most among the three effects so that it is possible that an FTA formation worsens the total welfare when T is large. A further increase in T, however, reverses the dominant effect from the tariff losses to a growth in exports by the MNE. When T approaches to T^{max}, most of the exports are done by the MNE that is more efficient in production than firm L. Thus, a growth in consumer surplus dominates the tariff losses. In sum, the FTA benefits the member countries when T is either close to $\widehat{T}$ or T^{max}.[25] The following proposition summarizes the welfare effect for inside countries.

Proposition 3. In the absence of ROO, if the MNE always chooses the offshoring scheme ($T > \widehat{T}$), there exist cutoff levels of T, $\underline{T}_W^O(> \widehat{T})$ and $\overline{T}_W^O(< T^{max})$, such that an FTA formation decreases the joint welfare of inside countries for $\underline{T}_W^O < T < \overline{T}_W^O$. If the MNE always chooses the inshoring scheme ($T < T^*$), there exists cutoff level of T,$T_W^*(< T^*)$ such that an FTA formation also decreases the joint welfare when $\tau < \frac{2(a-w)}{13}$ and $T < T_W^*$ hold. Otherwise, it improves the joint welfare.

The MNE's manipulation of transfer price is the reason why an FTA can worsen the welfare of inside countries. When the MNE chooses inshoring and transfer price is used for a strategic motive, the MNE exports more than the local firm in the market equilibrium and the rent-shifting role of tariff is more significant than the market equilibrium without transfer pricing. The same is true when the MNE chooses offshoring and transfer price is used for a tax-saving motive, because the manipulation of transfer price lowers the perceived marginal cost of the MNE and its exports are larger than the market equilibrium without transfer pricing. Therefore, an FTA formation that eliminates a rent-shifting effect of tariff may worsen the joint welfare of inside countries if we consider the transfer pricing.

4 Equilibrium with ROO

In this section, we introduce ROO into the benchmark setup. We use " ˜ " as a circumflex for the case with ROO. After an FTA is formed, exporting firms need to meet the VA criterion of ROO to be eligible for the non-application of τ. Specifically, ROO require firms to add at least $\underline{\alpha}(\in [0,1])$ fraction of the values of exported goods within the FTA. This implies that firm L always meets ROO because it always uses FTA-made inputs so that its VA in country H is always 1. However, if firm

[25] Formally, the two thresholds, $\underline{T}_W^O$ and $\overline{T}_W^O$, are the tax level that attains$W^{O*} = \widehat{W}^O$, where $\underline{T}_W^O = \frac{(1+t)(a-w)-3(1-t)\Delta+(3-2t)\tau-(1-t)\sqrt{\theta}}{2(a-w)+\tau}$, $\overline{T}_W^O = \frac{(1+t)(a-w)-3(1-t)\Delta+(3-2t)\tau+(1-t)\sqrt{\theta}}{2(a-w)+\tau}$, and $\theta \equiv (a-w)^2 - 6(\tau-\Delta)(a-w) + 3(\tau-\Delta)(\tau-3\Delta)$.

M chooses offshoring and sets $\widehat{r}^O$, the VA is always zero, which results in violation of ROO so that the MNE incurs tariff even after an FTA is formed. We call this case Regime N (Non-compliance).[26]

In order for firm M to utilize FTA tariff, it has to comply with ROO by either (i) procuring inputs in the host country as we saw in the benchmark case (Regime I), or (ii) setting r^O such that

$$\alpha \equiv \frac{p^O - r^O}{p^O} \geq \underline{\alpha}. \tag{17}$$

is satisfied. It is apparent that α is decreasing in r and thereby $(p^O - r^O)/p^O$ is more likely to exceed $\underline{\alpha}$ as firm M sets lower r^O. Therefore, the other possible transfer price is such that the one is as high as possible and satisfies VA criterion, which we call Regime B (Binding ROO).

Given that $\lambda_L = \lambda_M = 0$ and Eq.(17), the following optimal transfer price is obtained under Regime B,

$$\widetilde{r}^B = \underbrace{w - \Delta + \frac{(1-T)(a-w+2\Delta)}{(1-t)+2(1-T)}}_{=\widehat{r}^O} \underbrace{- \frac{3(1-T)\{(1-T)(a+w)+(1-t)(w-\Delta)\}\underline{\alpha}}{\{(1-t)+2(1-T)\}\{(1-t)+2(1-T)-(T-t)\underline{\alpha}\}}}_{\text{Tariff elimination motive}} \tag{18}$$

where the third term is a *tariff elimination motive* which captures the adjustment term for tariff elimination. We can easily see that $\widetilde{r}^B$ is decreasing in $\underline{\alpha}$ and is equivalent to $\widehat{r}^O$ at $\underline{\alpha} = 0$. Thus, the more stringent VA criterion lowers the transfer price to meet ROO. The equilibrium profit becomes

$$\widetilde{\Pi}_M^B = (1-T)\left(\frac{(1-t)(a-w+2\Delta) - (T-t)(a+w)\underline{\alpha}}{3-2T-t-(T-t)\underline{\alpha}}\right)^2. \tag{19}$$

Intuitively, the post-tax profits under regime B is a decreasing function of $\underline{\alpha}$ as an increase in $\underline{\alpha}$ induces the MNE to set transfer price that deviates more from the optimal level in the absence of ROO.

Firm M chooses among Regimes I, N, and B. Let us first compare $\widetilde{\Pi}_M^I$ and $\widetilde{\Pi}_M^N$. Since both profits are independent of the VA threshold, $\underline{\alpha}$, the tariff level and the tax differential determine which profit is higher. For the same reason as the benchmark case, the MNE prefers Regime I to Regime N if tax differential is small while regime N is more preferable under a large tax differential. There exists a unique threshold $\widetilde{T}$ such that $\widetilde{\Pi}_M^I = \widetilde{\Pi}_M^N$ holds. Since the MNE needs to incur tariff under regime N, this threshold is greater than that of benchmark setup, $\widehat{T} < \widetilde{T}$. Moreover, as a larger tariff discourages firm M to choose Regime N, $\frac{\partial \widetilde{T}}{\partial \tau} > 0$ holds. Given the level of t, firm M

[26]Some empirical evidences show that not all firms use FTA tariffs because of the existence of ROO, which means the impacts of a FTA formation are heterogeneous across firms. See, for example, Takahashi and Urata (2010) and Hayakawa et al. (2013).

prefers regime I to regime N if the following inequality holds:

$$\widetilde{\Pi}_M^I - \widetilde{\Pi}_M^N \geq 0 \iff T \leq \widetilde{T} \equiv \frac{(3-t)}{2} - \frac{(1-t)\{a-w+2(\Delta-\tau)\}\sqrt{2}}{a-w},$$

Otherwise, it prefers regime N to regime I.

Let us next compare the profits in Regime B with those in Regime N and Regime I. Since $\widetilde{\Pi}_M^B = \widetilde{\Pi}_M^O$ at $\underline{\alpha} = 0$, which is larger than $\widetilde{\Pi}_M^N$ and $\widetilde{\Pi}_M^I$ when $T > \widehat{T}$, and $\widetilde{\Pi}_M^B$ is decreasing in $\underline{\alpha}$, there exist a unique threshold, $\underline{\alpha}^N$ (resp. $\underline{\alpha}^I$), above which firm M prefers Regime N (resp. Regime I) to Regime B. These thresholds are derived by

$$\underline{\alpha}^N \equiv \frac{(1-t)(3-2T-t)\tau}{(T-t)\{(1-T)a+(2-T-t)w-(1-t)(\Delta-\tau)\}}, \tag{20}$$

$$\underline{\alpha}^I \equiv \frac{2(1-t)(a-w+2\Delta)\sqrt{2}-(a-w)(3-2T-t)}{(T-t)\{2(a+w)\sqrt{2}-(a-w)\}}, \tag{21}$$

respectively. Intuitively, under less strict ROO, the MNE prefers Regime B to Regimes N and I, because using transfer price for complying with ROO and realizing tariff elimination become less costly as regulation on VA criterion becomes less stricter. In other words, the MNE's gains from tariff elimination become smaller as the FTA is attached with a stringent ROO. Putting the above comparisons all together, we characterize the equilibrium outcomes as follows.

Proposition 4. After an FTA with ROO is formed, the MNE chooses (i) inshoring if $T \leq \widetilde{T}$ and $\alpha > \underline{\alpha}^I$, (ii) offshoring and its exports incur tariff if $\widetilde{T} < T$ and $\alpha > \underline{\alpha}^N$, and (iii) offshoring and it uses its transfer price to meet ROO if $\underline{\alpha} \leq \min\{\underline{\alpha}^I, \underline{\alpha}^N\}$.

The equilibrium MNE's choice is depicted in Figure 3. We should learn the following three things from the figure. First, as Takahashi and Urata (2010) and Hayakawa et al. (2013) pointed out, some firms may not utilize FTA tariffs because of the burden of ROO, which corresponds to the area $\widetilde{T} \leq T$ and $\underline{\alpha}^N \leq \underline{\alpha}$ in the figure. Under Regime B or Regime I, the MNE no longer use its transfer price to avoid a high corporate tax of country H, and the increased tax payments is interpreted as a cost of meeting ROO.

Second, as Conconi et al. (2018) shows, ROO lower the likelihood of input procurement from outside FTA countries. In the figure, this "input trade diversion" corresponds to the area of $\widehat{T} \leq T \leq \widetilde{T}$ and $\underline{\alpha}^I \leq \underline{\alpha}$. As the MNE's input production takes place in country O in the area under no ROO, ROO work as an effective tool not only to promote trade liberalization but also to prevent the MNE from engaging in tax avoidance. Finally, the MNE manipulates transfer price to comply with

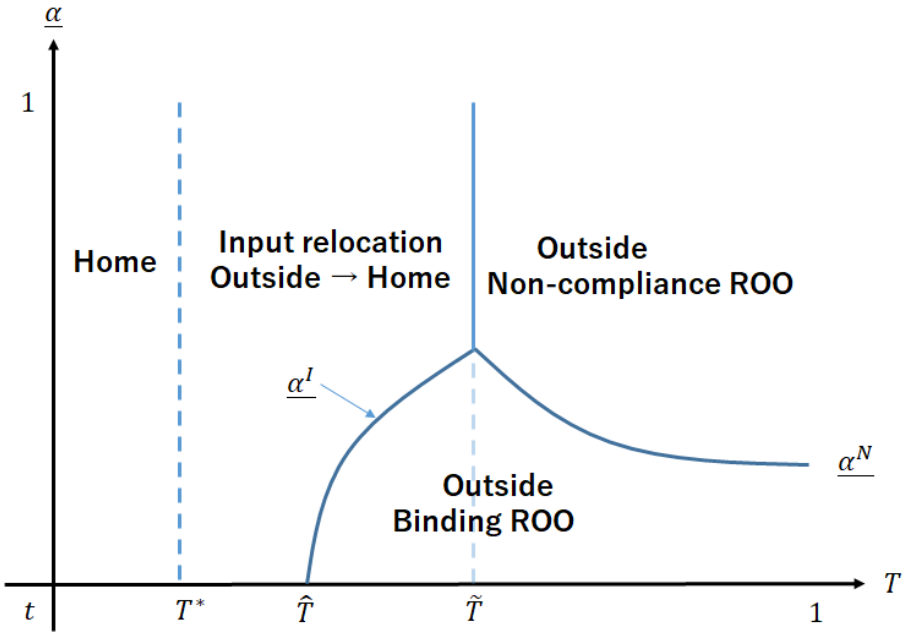

Figure 3: Equilibrium MNE's strategies

ROO only if the VA criterion is not so stringent.[27] ROO again play a role to prevent the MNE from tax avoidance.

4.1 The welfare effects of FTA formation with ROO

We have shown that a formation of FTA with ROO may change the MNE's strategies in transfer pricing or input location. With taking this effect of ROO into account, this section investigates the welfare impact of FTA formation.

4.1.1 The effect on the MNE

We investigate how an FTA formation affects the MNE's profits. Recall that, in the absence of ROO, the formation of an FTA increases the post-tax profits of the MNE. With ROO, however, the FTA can reduce the MNE's post-tax profits in each regime.

Under the situation where the MNE chooses Regime N, an FTA formation definitely reduces the profits of the MNE because only the MNE burdens tariff while the rival firm, firm L, takes advantage of tariff elimination. When the FTA formation leads to the input relocation of the MNE (i.e., Regime I), the MNE's profits decline if the tax differential is large. By producing inputs inside the FTA, the MNE gives up a tax-saving opportunity by manipulating its transfer price. The loss

[27] This result is the same as Mukunoki and Okoshi (2019) which concludes that an offshoring firm adjusts or increases export price to meet VA criterion in the sense that firms are likely to manipulate their price for ROO at any phase of price.

from paying a high corporate tax is significant if the tax differential is large. Without ROO, this reduction never happens because the possible tax range that makes Regime I being the post-FTA equilibrium outcome is sufficiently small. With ROO, however, the MNE chooses input relocation under a wider tax range. In other words, $\widehat{T} < \widetilde{T}$ holds and there exists a unique threshold T_M such that $\widetilde{\Pi}_M^I = \Pi_M^{O*}$ holds. The threshold is given by

$$T_M \equiv \frac{3-t}{2} - \frac{(1-t)(a-w+2\Delta-\tau)\sqrt{2}}{(a-w)}. \tag{22}$$

Only below this threshold, the FTA benefits the MNE. Note that $T_M < \widetilde{T}$ always holds, meaning that the MNE loses from the FTA formationa when the ROO input location and $T_M < T < \widetilde{T}$ holds.

When Regime B is the post-FTA equilibrium outcome, the MNE manipulates transfer price for complying with ROO. In this case, the MNE exploits tariff elimination but it can avoid tax payments to a high-tax country only partially, because the transfer price is smaller than the level that reallizes zero profits of firm M_D. As the stringency of the VA criterion of ROO increases, the extent to which the MNE save tax payments becomes smaller. On top of that, the lower transfer price due to the stricter ROO increases the perceived marginal cost of the MNE and reduces the MNE's exports. Thus, these negative impacts become larger as $\underline{\alpha}$ increases and they dominate the gains from tariff elimination when it is sufficiently high. Similar to T_M, we can obtain a unique threshold $\underline{\alpha}_M$ such that $\widetilde{\Pi}_M^B = \Pi_M^{O*}$ holds, which is specifically given by

$$\underline{\alpha}_M \equiv \frac{(3-2T-t)(1-t)\tau}{(T-t)\{2a(1-T)+w(2-T-t)-(1-t)(2\Delta-\tau)\}}, \tag{23}$$

where $0 < \underline{\alpha}_M < \underline{\alpha}^N$ holds because $\widetilde{\Pi}_M^B$ is decreasing in $\underline{\alpha}$ and $\widetilde{\Pi}_M^N < \Pi_M^{O*} < \widehat{\Pi}_M^O = \Pi_M^B|_{\underline{\alpha}=0}$ holds. The following proposition summarizes the effects of an FTA formation on the MNE's profits.

Proposition 5. A formation of an FTA with ROO hurts the MNE if the post-FTA MNE's choice is (i) not complying with ROO (Regime N), (ii) locating its upstream affiliate in host country (Regime I) and $T > T_M$ holds, or (iii) using transfer price to comply with ROO (Regime B) and $\underline{\alpha}_M < \underline{\alpha}$ holds. Otherwise, it benefits the MNE.

This proposition suggests that the MNE may be negatively influenced by an FTA formation even if it utilizes the FTA tariff in exporting. Even though the FTA formation hurts the MNE, such choices are still optimal for the MNE. This is because the MNE uses its transfer price to avoid a high corporate tax and to be stronger in the product market competition in the pre-FTA equilibrium, but the elimination of tariff on the rival firm forced the MNE to give up transfer pricing by input

relocation or to use the transfer price to comply with the ROO. Therefore, the profit loss from paying more corporate taxes and that from losing an advantageous position in the market competition can outweigh the positive profit effect from tariff elimination.

4.1.2 The effect on the local firm

As is stated in section 3.4, firm L gains from the FTA when the MNE does not change the location of its input production, or when it relocates the location from country O to country H and $T_L < T$ holds. This statement still holds with ROO and the gains become greater when the MNE gives up complying with ROO (regime N) or manipulates transfer price to meet VA criterion (regime B). This is because the MNE still needs to incur tariff or its perceived marginal cost gets higher due to a lower transfer price. In the case where the MNE changes the location of input production, however, the FTA formation may hurt firm L. The following proposition summarizes the effect of an FTA formation on the local firm.

Proposition 6. If $T^* < T < \min[\widetilde{T}, T_L]$ holds, a formation of FTA with ROO hurts the local firm. Otherwise, it benefits the local firm.

As in the benchmark model, a shift from offshoring to inshoring makes the MNE more aggressive in the product market and it has a negative effect on the local firm. If the negative effect exceeds the positive effect from tariff elimination, a FTA formation hurts the local firm.

Importantly, this shift is more likely to happen with ROO, because ROO increases the cost of choosing offshoring. Specifically, without ROO, we have changes in the location of input production if $T^* < T < \widehat{T}$ and $\tau < \tau_L$ hold or if $T^* < T < T_L$ and $\tau_L < \tau$ hold. With ROO, however, a range of T that induce changes in the location of input production is widened, because $\widehat{T} < \widetilde{T}$ holds. The ranking between $\widetilde{T}$ and T_L is ambiguous and depends on the level of tariff, τ.[28] Thus, firm L's profit decreases if $T^* < T < \min[\widetilde{T}, T_L]$.

It is counter-intuitive that ROO may hurt the local firm, although they only restrict the MNE's actions. It contrasts previous arguments, such as Krishna and Krueger (1995), that ROO work as a "hidden protection" policy for both the domestic upstream and downstream industries, in the sense that it promotes a local input production and raises the cost of the foreign firm. In our model, although ROO induce the MNE to procure inputs inside an FTA and they raise the production cost of the MNE, they also cause the MNE's decentralization decision. Then, the MNE uses its transfer price as a commitment devise to make the downstream MNE more aggressive in the product market.

[28] As is shown in Appendix A.5, $\widetilde{T} > T_L$ holds if $\tau > \tau_L^R$ and $\widetilde{T} < T_L$ holds if $\tau < \tau_L^R$, where τ_L^R is the cut-off level of tariff.

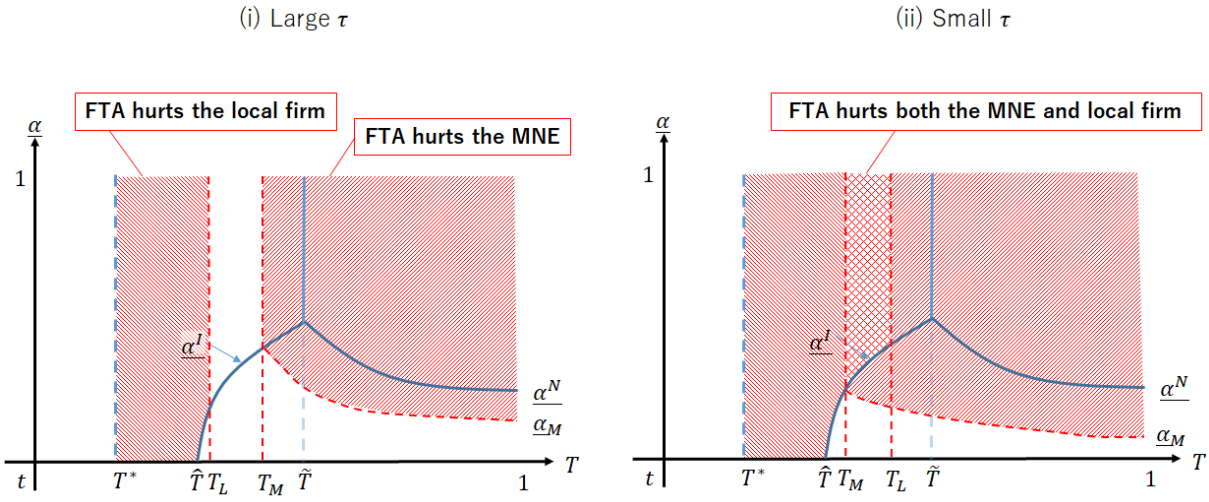

Figure 4: Effect of ROO on firms

In other words, our model indicates that ROO may have a pro-competitive effect in the domestic downstream industry.

Another intriguing result is that an FTA formation with ROO may hurt the MNE and firm L at the same time, even though both of them comply with ROO and are qualified for zero-tariff exports. Figure 4 illustrates the FTA impacts on the post-tax firms' profits with different τ.

On the left figure, τ is large such that $T_L < T_M$ holds. In this cae, the FTA can hurt either firm L or the MNE, but it never hurts both firms at the same time. When τ is high, if the tax is small enough such that an FTA formation hurts the local firm, the FTA formation always benefits the MNE because the positive effects from tariff elimination and strategic effect of decentralization outweigh the negative effect from input relocation. Conversely, the gains from tariff elimination is very large and the local firm always benefits from an FTA formation whenever the tax of the host country, T, is large enough to hurt the MNE.

On the right figure, however, τ is small such that $T_M < T_L$ holds. In this case, the magnitude of the positive effect from tariff elimination, which benefit both the MNE and the local firm, is not so large and there exists a range of $T \in (T_M, T_L)$ in which an FTA formation hurts both firms at the same time.[29]

This result is novel in the FTA literature because the existing studies suggest that at least some exporting firms producing within FTA benefits from tariff elimination. This paper suggest that, if we take input account the MNE's location choice, its manipulation of transfer pricing, and ROO, then an FTA formation decreases the post-tax profits of all exporting firms, even though they comply

[29]Specifically, this range appears when $\tau < \frac{(a-w)(2-\sqrt{2})}{4}$ holds.

with ROO and all tariffs are eliminated.

4.1.3 The effect on consumers

As Proposition 2 states, an FTA formation without ROO always benefits consumers. The presence of ROO does not change the MNE's actions for $T < \hat{T}$, where the MNE produces inputs in country H both before and after an FTA formation, and also for $T^* \leq T < \hat{T}$, where the MNE changes the location of input production with an FTA formation. In these cases, an FTA with ROO has the same effect as an FTA without ROO does, because the requirement of ROO is not binding for the MNE.

For $\hat{T} \leq T$, however, ROO change the MNE actions. In this range of T, the MNE always produces inputs in country O without ROO. With an FTA with ROO, the MNE either (i) changes its input location to comply with ROO (Regime I), (ii) manipulates the transfer price to comply with ROO (Regime B), or (iii) does not comply with ROO (Regime N), depending on the level of $\bar{\alpha}$ and T.

In case (i), it is ambiguous whether ROO increase or decrease the exports of the MNE. On the one hand, the input relocation induced by ROO increases the MNE marginal cost and has a negative impact on the volume of exports of the MNE. On the other hand, the input relocation is accompanied by the MNE's decentralization decision, which makes the MNE more aggresive in the product market and has a positive impact on the volume of its exports. If the tariff is sufficiently low ($\tau \leq \tau_{CS}^R$) or tax gap is sufficiently low ($\hat{T} < T < T_{CS}^R$), the latter effect dominates the former and ROO enhance the consumers' gains from FTA formation. Otherwise, ROO reduce the consumers' gains form FTA formation. In case (ii) and (iii), ROO diminish consumers' gains from FTA formation because lowering transfer price to comply with the VA criterion (Regime B) or the imposed tariff by not complying with ROO (Regime N) increases the MNE's (perceived) marginal cost of exporting and reduces exports of the MNE.

Even if ROO reduces the consumers' gains from FTA formation, the direct effect of tariff elimination of FTA always dominates the ROO effect and an FTA formation always benefits consumers. The following proposition summarizes the results.

Proposition 7. A formation of FTA with ROO always benefits consumers. The VA requirement of ROO increases the consumers' gains from FTA formation if it induces the input relocation of the MNE and either $\tau < \tau_{CS}^R$ or $\hat{T} < T < T_{CS}^R$ holds. Otherwise, the VA requirement of ROO diminishes or has no effects on the consumers' gains.

As long as the MNE continues to produce inputs outside an FTA, ROO diminish the consumers' gains from the FTA formation due to the manipulation of transfer price for complying with ROO or

the non-compliance with ROO. If ROO induce the MNE's input relocation, however, the consumers' gains may become larger because the MNE uses its transfer price for a strategic motive after the input relocation, and the resulting intensified competition gives consumers additional gains. If that is the case, ROO improve the consumers' gains at the expense of the profit of the local firm, even though the requirement of ROO is binding only for the MNE and it raises the production cost of the MNE. In other words, ROO increase the MNE's incentive to manipulate the transfer price for a strategic motive, rather than to use it for saving tax payments.

4.1.4 The effect on the welfare of inside countries

We have shown that an FTA without ROO can worsen the joint welfare of member countries. Here, we discuss how ROO change the welfare effect.

In the presence of ROO, regimes B and N are possible in addition to regimes I and O. When the VA requirement of ROO is not so stringent, the MNE prefers to manipulate transfer price for complying with ROO (Regime B). As $\underline{\alpha}$ gets higher, the transfer price gets lower, which leaves some of the MNE's taxable income in country H, increases the profits of firm L, decreases total exports. If $\underline{\alpha}$ is small enough, the positive effects of ROO dominate the negative ones. Note that the case at $\underline{\alpha} = 0$ is equivalent to the case without ROO. Proposition 3 shows that an FTA formation without ROO may worsen the joint welfare of member countries. This implies that increase in $\underline{\alpha}$ can turn a welfare-reducing FTA into a welfare-improving one.

If the governments set $\underline{\alpha}$ higher than $\underline{\alpha}^N$ and $T \geq \widetilde{T}$, however, the MNE gives up complying with ROO (Regime N) and it uses its transfer pricing solely for tax avoidance purpose. In this case, country H cannot collect tax revenue from the MNE both before and after the FTA formation. Thus, an FTA formation benefits consumers and firm L, while members lose a part of tariff revenues. The increase in firm L's profits is larger because tariff is still incurred to the MNE after the FTA formation. In this case, the MNE is treated as if it produces the final good outside the FTA, and the rent-shifting effect makes a welfare-improving FTA more likely.

Because of analytical difficulty to derive a concrete result, we rely on a numerical example to examine the welfare impacts of an FTA with ROO, as Figure 5 illustrates.[30] The doted areas are the case where the FTA reduce the joint welfare of members. On the left figure (Large initial τ), the doted area appears in the area of regime B for a low $\underline{\alpha}$. Here, a stricter VA criterion of ROO works as an effective tool to enhance the joint welfare. On the right figure (small τ), an FTA formation can be

[30] In Figure 5, the parameters are set at $a = 3$, $w = 1$, $t = 0.1$, $\Delta = 1/32$, $\tau = (a - w)/8$ for the large τ and $\tau = (a - w)/16$ for the small τ.

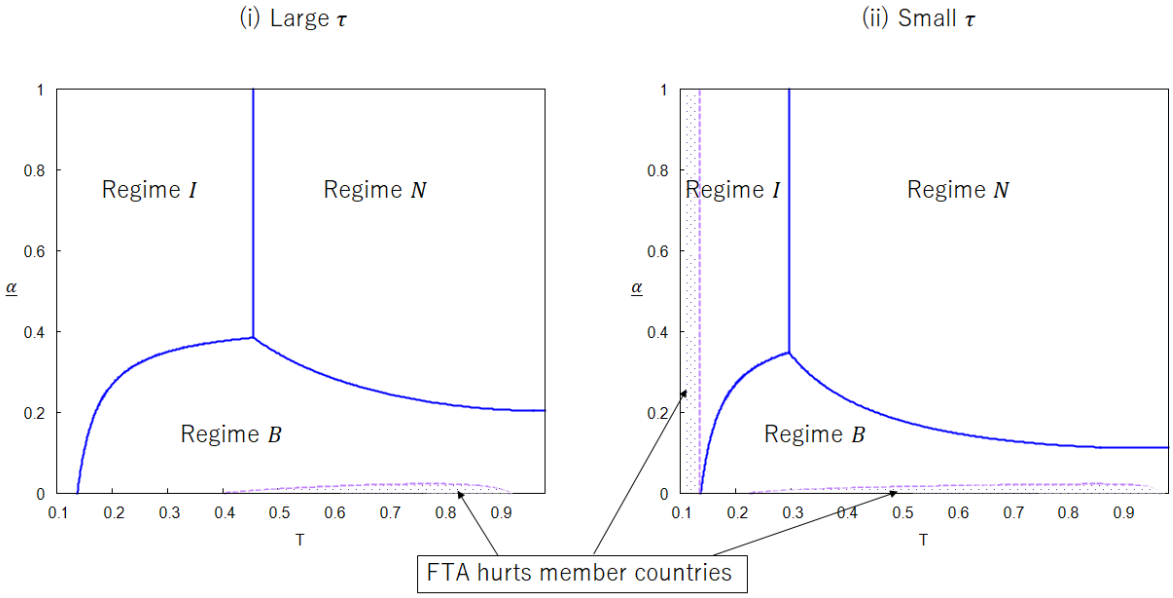

Figure 5: Welfare effect of ROO

in $T < T^*$. Unlike the case under regime B, the stringency of ROO does not affect the joint welfare because the MNE always chooses inshoring and *ROO* is unbinding in this area. This numerical example provides us with the following proposition.

Proposition 8. An FTA formation with ROO may decrease the joint welfare of FTA, but ROO reduce the likelihood of a welfare-reducing FTA compared to an FTA without ROO.

This proposition suggests that VA criterion of ROO contributes to increase the welfare gains from FTA. Without ROO, FTA may reduce the joint welfare of the FTA member countries when $\underline{T}_W^O < T < \overline{T}_W^O$ holds. With ROO, however, the MNE must lower its transfer price to comply with ROO in Regime B. In this case, the MNE cannot fully avoid a high tax in the host country and its perceived marginal cost becomes higher. In other words, a tariff-elimination motive of transfer pricing counteracts a tax saving motive. This change in MNE's strategy benefits firm L and a part of the increases in the downstream profit of the MNE is captured by country H as a tax revenue. We should note, however, that the welfare gains are materialized at the expense of consumers' gains because ROO weaken the product market competition.

If the VA criterion of ROO is stringent enough, then either Regime I or Regime N becomes the equilibrium outcome. Here, adjusting transfer price to comply with ROO is too costly for the MNE, and the MNE do a binary choice between tariff elimination by the input relocation (Regime I) or tax avoidance (Regime N). The MNE chooses tariff elimination if If the tax gap is small ($T < \widetilde{T}$) and tax avoidance otherwise. In either case, an FTA formation always improves the joint welfare of member

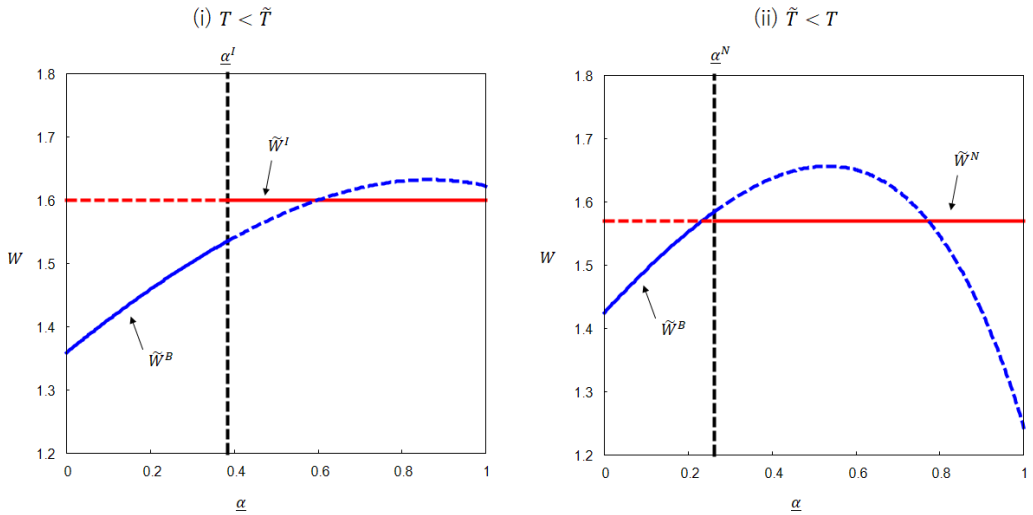

Figure 6: Optimal level of $\underline{\alpha}$

countries.

4.2 Optimal level of ROO

As the final argument, we examine the optimal level of VA criterion in a numerical example.[31] Figure 5 demonstrates that a stricter VA criterion of ROO can improve the total welfare inside FTA. This does not mean, however, the strictest criterion ($\underline{\alpha} = 1$) always lead to the highest welfare of members. Fig. 6 shows the change in the total welfare in response to a change in $\underline{\alpha}$ at $T = 0.45 < \widetilde{T}$ and $T = 0.65 > \widetilde{T}$, respectively.

In both figures, a blue curve shows the joint welfare of members under Regime B while the red horizontal line shows the joint welfare under either Regime I or Regime N. The solid parts are realized welfare in equilibrium. When $T < \widetilde{T}$ holds, which is shown in the left of Fig. 6, the stricter criterion increases the total welfare until $\underline{\alpha}$ reaches $\underline{\alpha}^I$ which is drawn with the vertical doted line. For $\underline{\alpha} \geq \underline{\alpha}^I$, the total welfare is constant since the MNE changes the location of its input production from country O to country H. As the total welfare under regime I is higher than that under regime B at $\underline{\alpha} = \underline{\alpha}^I$, the optimal VA criterion is $\underline{\alpha} \in (\underline{\alpha}^I, 1]$ in our model. Alternatively, when $T > \widetilde{T}$ holds, the MNE fully use its transfer price for a tax avoidance motive (Regime N). As the consumers' gains are smaller under Regime N due to the MNE's tariff burden, the total welfare under regime N, which is drawn in right from the vertical doted line, is lower than the highest welfare level under regime B, which is the left area from the vertical doted line. Thus, the optimal $\underline{\alpha}$ is marginally smaller than $\underline{\alpha}^N$.

[31] The same set of parameters are used with $\tau = (a - w)/8$.

These results suggest that policy makers should decide the optimal $\underline{\alpha}$ with taking into account a corporate tax rate and the MNE's transfer pricing. In reality, the VA thresholds are usually set between 30% and 60%.[32] When host countries of export-platform MNEs impose high tax rates, which is actually the case with NAFTA where Mexico has a higher corporate tax than U.S. and Canada, the VA criterion of ROO may have a positive role to prevent tax avoidance of MNEs and secure welfare gains for member countries. Moreover, as different VA thresholds are imposed on different products, the lower thresholds of VA criterion or other criterion of ROO should be set for the products/industries where MNEs are active in the industry, input relocation and other ways of profit shifting, such as royalty transfer or internal debt, are difficult to conduct,

5 Discussion

Our benchmark analysis has provided a set of new results which have not explored in the extant literature. In this section, we argue the robustness of these patterns of the MNE's strategies by relaxing some assumptions made in the basic model.

5.1 Home market

In the baseline model, the final product is consumed only in country F. Here, we discuss how the results are affected by considering the market in country H.

We have assumed that the MNE is able to have only a single input plant in either country O or H because of a high fixed cost of establishing plants. Besides that, we suppose that the production process in one plant cannot be separated depending on the destination of the products. Thus, it is not possible that the MNE procures local inputs to comply with ROO and at the same time it uses the input produced outside the FTA only for the domestic supply of the product. Therefore, we focus on the case where the MNE procures inputs only via self-production. For simplicity, we assume that the market size of country H is the same as that of country F.

As we assumed in the benchmark case, the MNE cannot report negative profits in each country. Then, the optimal transfer price under offshoring is the one such that $(p_H - r)x^O_{MH} + (p_F - r)x^O_{MF} = 0$ holds, while the first-order condition of the profit maximization provides the optimal internal price under inshoring. We put subscript HM for the case with home market competition, while no

[32] see http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=TAD/TC/WP(2015)28/FINAL&docLanguage=En.

subscript means the benchmark case. We have the following rankings of the input prices.

$$\widetilde{r}^I = \widetilde{r}^I_{HM} < r^{I*}_{HM} < r^{I*}, \tag{24}$$

$$\widetilde{r}^O < \widetilde{r}^O_{HM} < r^{O*} < r^{O*}_{HM} < \widehat{r}^O_{HM} = \widehat{r}^O, \tag{25}$$

The input price with the home market are the same as the benchmark case after an FTA formation without ROO ($\widehat{r}^O_{HM} = \widehat{r}^O$). Before the FTA, the MNE has a stronger incentive to exploit the strategic effect because the MNE also competes in country H in the case of inshoring. In the case of offshoring, the MNE also aggressively sets the transfer price so as to the operating profits from export market become negative but the negative profits are covered by the positive profits from the home country. Therefore, in both cases, the MNE's incentive to manipulate inputs price gets stronger if we introduce the home market competition.

ntuitively, we obtain similar input prices under both cases, and

As we face the analytical complication, we rely on a numerical analysis to confirm the existence of thresholds ($T^*, \widehat{T}, \widetilde{T}, \underline{\alpha}^I, \underline{\alpha}^N$). Figure 7 in the appendix shows us the same pattern of the MNE's strategy as the benchmark case.

5.2 Partial procurement of inputs

We have assumed that the MNE makes a binary choice on input procurement, that is, a "make all or buy all" choice. In the real world, however, the MNE may purchase some fraction of parts from the local suppliers and procure the rest of parts by the intra-firm transactions, which we refer to regime P.

Here, we maintain the supposition that the MNE cannot establish input plants in both countries O and H, because of the fixed cost of establishing plants. Thus, the MNE sets up an input plant in country O if it engages in partial procurement. We consider the situation where the MNE uses a continuum of inputs indexed in $[0,1]$ space. Let $\beta \in [0,1]$ denote the fraction of the inputs that the downstream affiliates procures from the upstream affiliate in country O, and r^P be the transfer price of that transaction. Alternatively, the inputs indexed by $\beta' \in (\beta, 1]$ are procured through the competitive input market as firm L does. Then, the amount of cross-border intra-firm trade becomes βx^P_M and the modified VA ratio becomes $(p - \beta r^P)/p$. We consider the modified four stage game. In the first stage, the MNE chooses the country where it produces inputs by itself. In the second stage, the headquarters decides the level of β to satisfy the VA criterion. In the third stage, the level of r^P is determined. In the last stage, the two firms engage in Cournot competition.

The MNE's post-tax profits are given by

$$\begin{aligned}\Pi_M^P &= (1-t)[\beta\{r^P-(w-\Delta)\}x_M^P]+(1-T)[\{p-\beta r^P-(1-\beta)w\}x_M^P] \\ &= (1-T)(p-c_M^P)x_M^P, \quad \text{where} \quad c_M^P=\left\{w(1-\beta)+\frac{\beta\{(1-t)(w-\Delta)-(T-t)r^P\}}{1-T}\right\}. \end{aligned} \tag{26}$$

Since $\frac{\partial \Pi_M^P}{\partial r^P}>0$ holds, the optimal transfer price is set as high as the level that realizes zero profit of the upstream affiliate. Namely, it is set such that $p-\beta r^P-(1-\beta)w=0$ holds:

$$\widetilde{r}^P=\underbrace{c-\Delta+\frac{(1-T)(a-w+2\Delta)}{3-2T-t}}_{=\widetilde{r}^O}+\underbrace{\frac{(1-T)(1-\beta)(a-w)}{\beta(3-2T-t)}}_{\text{Partial procurement adjustment}}. \tag{27}$$

In the above equation, the third term captures the effect of considering a partial procurement. As the MNE reduce the fraction of self-procurement, the MNE needs to increase its transfer price on inputs to shift profits for saving its tax payments.

The MNE's VA ratio is positive and calculated as

$$\alpha=\frac{p-\beta\widetilde{r}^P}{p}=\frac{(3-2T-t)(1-\beta)w}{(1-T)a+(2-T-t)w-\beta(1-t)\Delta}. \tag{28}$$

As we can confirm $\frac{\partial \widetilde{\Pi}_M^P}{\partial \beta}>0$, the optimal level of β is also set at the level that realizes zero profit of the upstream affiliate.[33] Therefore, the optimal input procurement fraction is,

$$\widetilde{\beta}^P=\frac{w(3-2T-t-\{(1-T)a+(2-T-t)w\}\underline{\alpha})}{w(3-2T-t)-(1-t)\Delta\underline{\alpha}}, \tag{29}$$

where $\widetilde{\beta}^P\Big|_{\underline{\alpha}=0}=1$ holds at $\underline{\alpha}=0$ because regime B and regime P are equivalent without the VA requirement. As we can easily confirm that $\frac{\partial \widetilde{\beta}^P}{\partial \underline{\alpha}}<0$ holds, the upper bound of $\underline{\alpha}$ with partial procurement becomes:

$$\widetilde{\beta}^P\geq 0 \iff \underline{\alpha}\leq\frac{(3-2T-t)w}{(1-T)a+(2-T-t)w}\equiv\underline{\alpha}^P(<1). \tag{30}$$

From this equation, we can conclude that regime I and N of the benchmark model are still possible equilibrium outcomes when $\underline{\alpha}>\underline{\alpha}^P$ holds, while the area of regime B is replaced with regime P when $\underline{\alpha}<\underline{\alpha}^P$.[34] In regime P, even though the equilibrium transfer price becomes higher than that

[33] Formally, $\widetilde{\Pi}_M^P=(1-T)\left(\frac{(1-t)(a-w+\Delta\beta)}{3-2T-t}\right)^2$ and $\frac{\partial \widetilde{\Pi}_M^P}{\partial \beta}=2(1-T)\left(\frac{(1-t)^2(a-w-\Delta\beta)\beta}{(3-2T-t)^2}\right)>0$ hold.

[34] In the appendix, we put Figure 8 to illustrate a numerical example of the modified model.

of regime B with a partial procurement, the qualitative nature of the model remain unchanged.

5.3 Concealment costs for transfer price manipulation

In the benchmark model, there is no cost of manipulating transfer price. In practice, MNEs need to explain the plausibility of transfer pricing in order to shift profits across country. This cost should increase as MNEs shift more profits because explaining the reasons of the greater deviation from the "appropriate price" (or arm's length price), is more difficult. Here, we show that introducing a cost of transfer pricing does not change the main results.

Following the literature on transfer pricing, we introduce the following quadratic concealment cost in the case of offshoring:

$$C(r) = \frac{\delta\{r - (w - \Delta)\}^2}{2}, \tag{31}$$

where δ is a parameter which captures the difficulty of concealing a tax avoidance. A higher δ corresponds to more difficult environment of profit shifting due to well enforced tax authority, for example. In the case of inshoring, the MNE does not incur such a cost of profit shifting.

The introduction of the concealment cost does not influence the MNE's actions under inshoring scheme. Therefore, we only discuss its effect under offshoring scheme. The modified post-tax profits are given by

$$\Pi_M^O = (1-t)[\{r - (w - \Delta)\}x_M^O] + (1-T)[(p - r - \lambda_M\tau)x_M^O] - C(r). \tag{32}$$

Unlike the benchmark case, the costs may prevent the MNE from shifting all the profits and the VA ratio of the downstream affiliate is positive even without the VA criterion. In this case, the VA criterion of ROO may not affect the MNE's actions when $\underline{\alpha}$ is sufficiently small. Let α_M be a VA ratio such that the MNE's optimal transfer price without the VA criterion of ROO just attains the required VA ratio. Thus, the VA criterion of ROO does not affect the MNE's actions when $\underline{\alpha} < \alpha_M$ holds. However, stricter ROO keep the MNE away from enjoying both profit shifting and FTA tariff, and the MNE has to choose one of the three regimes as our benchmark analysis. Due to the analytical difficulty, we rely on a numarical analysis and confirm that we have qualitatively the same thresholds $(T^*, \widehat{T}, \widetilde{T}, \underline{\alpha}^N, \underline{\alpha}^I)$ and results.[35]

[35]In the appendix, we show a numerical example of equilibrium regime. See Figure 9.

6 Conclusion

A recent proliferation of FTAs plays a key role to advance trade liberalization between countries, and cross-border economic activities of MNEs prevail all over the world. This paper has investigated a vertically integrated MNE's input production and pricing strategies to analyze the welfare effects of an FTA formation, when the MNE is able to manipulate its transfer price of intra-firm trade. As in the previous papers, the MNE uses its transfer price to avoid a high corporate tax and/or to shift profits from the rival firm in the final-good market. After the formation of an FTA, however, the MNE faces a constraint for transfer price manipulation in the presence of ROO. Specifically, if ROO of the FTA employ a VA criterion, the MNE's transfer price must be sufficiently low for complying with ROO and to be eligible for tariff elimination.

When the difference in corporate taxes between a country outside FTA and a country inside FTA is large, the MNE's produces a necessary input in the outside country because the gains from tax avoidance is large. If it is small, the MNE produces the input in the inside country to take advantage of strategic effect of transfer pricing. In the absence of ROO, an FTA formation induces input relocation from the outside country to the inside country because strategic effect becomes more important than tax avoidance. Such an input relocation can hurt the local firm while the MNE and consumers always benefit from the FTA.

The presence of ROO gives the MNE two additional options: (i) transfer price manipulation with an upper limit for partial tax avoidance and tariff elimination and (ii) transfer price manipulation for full tax avoidance without tariff elimination. Some empirical evidences and anecdotes imply that (i) FTAs sometimes induce input relocation to inside FTA countries, (ii) not all firms export with using preferential tariffs of FTAs, and (iii) transfer price manipulation is associated with difference in corporate tax rates and the required VA criterion of ROO. Our model can explain these facts. A remarkable result is that both multinational and national firms can lose from an FTA formation even though these firms comply with ROO and tariffs imposed on them are eliminated. Furthermore, ROO can increase the total gains from FTAs, though it is at the expense of the local firm's profit.

These results provide important policy implications. First and foremost, the policy makers should notice that even though firms comply ROO and make tariff-free exports, it does not always mean these firms gain from FTA formation when the MNEs manipulate their transfer prices and a VA criterion of ROO is employed. Next, although the practical cooperation between custom departments and tax authorities is rarely observed, policy makers should pay more attention to the link between tariffs and corporate taxes in evaluating the welfare effects of an FTA because intra-firm

trade and export-platform FDIs are prevalent in the real world. Furthermore, from the viewpoint of trade policy, ROO can diminish the positive effect of trade liberalization since MNEs may give up utilizing preferential tariffs or they relocate the input production from an efficient outside country to an inefficient inside country. Furthermore, from the viewpoint of tax avoidance, an FTA with a VA criterion of ROO can be an effective tool to prevent MNEs' tax avoidance and secure the tax revenues of high-tax countries via either MNEs' input relocation or MNEs' less efficient transfer pricing manipulation to avoid a high tax.

There remains room for further research. We have assumed that tax rates and tariff rates are exogenously given. It is intriguing to investigate how the formation of an FTA affects the outcomes of tax competitions among countries and how it affects optimal tariffs the FTA members set. Another direction of extending the model is to examine the effects of regulations on transfer pricing, such as ALP, in the presence of ROO. Finally, further empirical investigation on the relationship between ROO and transfer pricing is essential.

Appendix

A.1 Proof of Proposition 1

From Eqs.(8) and (11), we can obtain a unique thresholds such that the MNE is indifferent over input location before and after FTA formation,

$$\begin{aligned}\Pi_M^O - \Pi_M^I = (1-T)\left((x_M^O)^2 - \frac{(x_M^I)^2}{2}\right) \propto \left(\sqrt{2}x_M^O - x_M^I\right) \geq 0 \\ \iff T \geq \frac{3-t}{2} - (1-t)\left(\frac{a-w+2\Delta-(2\lambda_M-\lambda_L)\tau}{a-w-(2\lambda_M-\lambda_L)\tau}\right)\sqrt{2}.\end{aligned}$$

By substituting $(\lambda_M, \lambda_L) = (1,1)$ for the pre-FTA case and $(\lambda_M, \lambda_L) = (0,0)$ for the post-FTA case, we obtain, respectively,

$$T \geq \begin{cases} \frac{3-t}{2} - \frac{(1-t)(a-w+2\Delta-\tau)\sqrt{2}}{a-w-\tau} \equiv T^*, & \text{before FTA,} \\ \frac{3-t}{2} - \frac{(1-t)(a-w+2\Delta)\sqrt{2}}{a-w} \equiv \widehat{T}, & \text{after FTA.} \end{cases} \tag{33}$$

The above expressions also show $T^* \leq \widehat{T}$. ■

A.2 Proof of Proposition 2

By the proposition 1, the three possibilities exist. The comparison of supplies is equivalent to that of post-tax profits if the MNE's input location is unaffected by FTA formation. When $T \leq T^*$ ($T \geq T^*$) holds, the MNE produces input in country H (country O) irrespective of FTA formation. The supply by both firms are,

$$x_M^s = \begin{cases} \frac{a-w-(2\lambda_M-\lambda_L)\tau}{2} & \text{when} \quad T \leq T^* \\ \frac{(1-t)\{a-w+2\Delta-(2\lambda_M-\lambda_L)\tau\}}{3-2T-t} & \text{when} \quad T \geq \widehat{T} \end{cases} \tag{34}$$

$$x_L^s = \begin{cases} \frac{a-w-(3\lambda_L-2\lambda_M)\tau}{4}, & \text{when} \quad T \leq T^* \\ \frac{(1-T)(a-w)-(1-t)\Delta-\{(1-T)\lambda_L-(1-t)(\lambda_L-\lambda_M)\}\tau}{3-2T-t} & \text{when} \quad T \geq \widehat{T} \end{cases} \tag{35}$$

so that elimination of tariff clearly increases supply, and equivalently post-tax profits by both firms, which directly means that total supply is also increased so that consumers' benefit arises.

When $T^* \geq T \geq \widehat{T}$ holds, the MNE changes the country of input production from country O to country H. Similar to the proof for proposition 1, the change in the MNE's post-tax profits is computed by using Eqs.(34),

$$\begin{aligned} \widehat{\Pi}_M^I - \Pi_M^{O*} \propto \widehat{x}_M^I - \sqrt{2}x_M^{O*} = \frac{a-w}{2} - \frac{(1-t)(a-w+2\Delta-\tau)\sqrt{2}}{3-2T-t} \gtreqless 0 \\ \iff T \lesseqgtr \frac{3-t}{2} - \frac{(1-t)(a-w+2\Delta)\sqrt{2}}{a-w} \equiv T_M. \end{aligned} \tag{36}$$

Thus, although the MNE's post-tax profits would decrease if $T > T_M$ held, $\widehat{T} < T_M$ always holds, which means the MNE always benefits from the FTA.

On the effect on firm L, the supply comparison is again equivalent to the profit comparison. Thus, the FTA benefits firm L if and only if,

$$\widehat{x}_L^I = \frac{a-w}{4} > \frac{(1-T)(a-w-\tau)-(1-t)\Delta}{3-2T-t} = x_L^{O*} \iff T > 1 - \frac{(1-t)(a-w+4\Delta)}{2(a-w-2\tau)} \equiv T_L \tag{37}$$

holds. We check whether T_L lies in $[T^*, \widehat{T}]$ by computing,

$$\begin{aligned} \widehat{T} - T_L = (1-t)\left\{\frac{a-w-\tau+2\Delta}{a-w-2\tau} - \frac{(a-w+2\Delta)\sqrt{2}}{a-w}\right\} \gtrless 0 \\ \iff \tau \gtrless \frac{(\sqrt{2}-1)(a-w)(a-w+2\Delta)}{(2\sqrt{2}-1)(a-w)-4\Delta} \equiv \tau_L. \end{aligned} \tag{38}$$

Thus, the FTA hurts firm L when $T^* \leq T \leq T_L$ if $\tau \geq \tau_L$ holds or when $T^* \leq T \leq \widehat{T}$ if $\tau \leq \tau_L$ holds.

Finally, we analyze the change in total exports. From Eqs.(34) and (35), the change is positive if,

$$\left(\widehat{x}_M^I + \widehat{x}_L^I\right) - \left(x_M^{O*} + x_L^{O*}\right) = \frac{3(a-w)}{4} - \frac{(2-T-t)(a-w-\tau)+(1-t)\Delta}{3-2T-t} > 0$$
$$\iff T < \frac{(1+t)(a-w)+4(2-t)\tau-4(1-t)\Delta}{2(a-w+2\tau)} \equiv T_{CS} \tag{39}$$

holds. By subtracting T_{CS} from $\widehat{T}$,

$$T_{CS} - \widehat{T} = (1-t)\left(\frac{(a-w+2\Delta)\sqrt{2}}{a-w} - \frac{a-w-\tau+2\Delta}{a-w+2\tau}\right) \tag{40}$$

is obtained. Since the second term is decreasing in τ, $T_{CS} - \widehat{T}$ is minimized at $\tau = 0$. At $\tau = 0$, the difference is $(1-t)(\sqrt{2}-1)\left(\frac{a-w+2\Delta}{a-w}\right) > 0$ so that $\widehat{T} < T_{CS}$ always holds, which means consumers in country F always benefit from the FTA. ■

A.4 Proof of Proposition 4

Similar to the proof of proposition 1, a comparison of the (weighted) MNE's supply is equivalent to the comparison of post-tax profits. We get $\widetilde{x}_M^N$ by substituting $\lambda_M = 1$ and $\lambda_L = 0$ into x_M^O. Then, we obtain each threshold such that $\widetilde{\Pi}_M^I = \widetilde{\Pi}_M^N$, $\widetilde{\Pi}_M^I = \widetilde{\Pi}_M^B$, and $\widetilde{\Pi}_M^N = \widetilde{\Pi}_M^B$, respectively, hold is,

$$\widetilde{\Pi}_M^I - \widetilde{\Pi}_M^N \propto \left(\widetilde{x}_M^I\right) - \left(\widetilde{x}_M^N\right)\sqrt{2} = \frac{a-w}{2} - \frac{(1-t)(a-w+2\Delta-2\tau)\sqrt{2}}{3-2T-t} \geq 0$$
$$\iff T \leq \frac{3-t}{2} - \frac{(1-t)(a-w+2\Delta-2\tau)\sqrt{2}}{a-w} \equiv \widetilde{T}, \tag{41}$$
$$\widetilde{\Pi}_M^I - \widetilde{\Pi}_M^B \propto \left(\widetilde{x}_M^I\right) - \left(\widetilde{x}_M^B\right)\sqrt{2} = \frac{a-w}{2} - \frac{\{(1-t)(a-w+2\Delta)-(T-t)(a+w)\underline{\alpha}\}\sqrt{2}}{3-2T-t-(T-t)\underline{\alpha}} \geq 0$$
$$\iff \underline{\alpha} \geq \frac{2(1-t)(a-w+2\Delta)\sqrt{2}-(a-w)(3-2T-t)}{(T-t)\{2(a+w)\sqrt{2}-(a-w)\}} \equiv \underline{\alpha}^I, \tag{42}$$
$$\widetilde{\Pi}_M^N - \widetilde{\Pi}_M^B \propto \left(\widetilde{x}_M^N\right) - \left(\widetilde{x}_M^B\right) = \left(\frac{(1-t)(a-w+2(\Delta-\tau))}{3-2T-t} - \frac{(1-t)(a-w+2\Delta)-(T-t)(a+w)\underline{\alpha}}{3-2T-t-(T-t)\underline{\alpha}}\right) \geq 0$$
$$\iff \underline{\alpha} \leq \frac{(1-t)(3-2T-t)\tau}{(T-t)\{(1-T)a+(2-T-t)w-(1-t)(\Delta-\tau)\}} \equiv \underline{\alpha}^N. \tag{43}$$

These conclude the proposition. ■

A.5 Proof of Proposition 5

Eq.(11) shows inclement of global post-tax profits of the MNE when $T < T^*$ while Eq.(8) does reduction in those when the equilibrium regime is regime N.

From the proof of proposition 2, we know that, when $T^* < T < \widetilde{T}$ and $\underline{\alpha} > \underline{\alpha}^I$ hold, the MNE benefits from the FTA if and only if

$$\widetilde{\Pi}_M^I - \Pi_M^{O*} \geq 0 \iff T \leq T_M^F. \tag{44}$$

Since $\widehat{T} < T_M < \widetilde{T}$ holds, FTA formation hurts the MNE if $T_M < T < \widetilde{T}$ holds.

Similarly in the case of regime B, the MNE benefits from the FTA if and only if

$$\begin{aligned}\widetilde{\Pi}_M^B - \Pi_M^{O*} \geq 0 &\iff \widetilde{x}_M^B - x_M^{O*} \geq \\ &\iff \underline{\alpha} \leq \frac{(3-2T-t)(1-t)\tau}{(T-t)\{2(a(1-t)-w(2-T-t)-(1-t)(2\Delta-\tau))\}} \equiv \underline{\alpha}_M.\end{aligned}$$

As $\frac{\partial \widetilde{\Pi}_M^B}{\partial \underline{\alpha}} < 0$ and $\widetilde{\Pi}_M^N = \widetilde{\Pi}_M^B\Big|_{\underline{\alpha}=\underline{\alpha}^N} < \Pi_M^{O*} < \widetilde{\Pi}_M^B\Big|_{\underline{\alpha}=0}$ hold, $\underline{\alpha}_M$ always exists in $[0, \underline{\alpha}^N]$. ■

A.6 Proof of Proposition 6

Again since the profits made by firm L is always square of the quantity that firm L supplies, we compare the supply by firm L before and after an FTA is formed. By Eqs.(35), the equilibrium firm L's supply under regime N is greater than that under pre-FTA. Eqs.(35) also show the increase in supply by firm L when $T < T^*$ as in the proof of proposition 2.

From the proof of proposition 2, we know that the FTA reduces the profits of firm L if and only if $T < T_L$ holds. By subtracting T_L from $\widetilde{T}$,

$$\begin{aligned}\widetilde{T} - T_L \propto \frac{a-w-\tau+2\Delta}{a-w-2\tau} - \frac{(a-w-2\tau+2\Delta)\sqrt{2}}{a-w} \geq 0 \\ \iff \tau \geq \tau_L^R (> 0), \qquad (45) \\ \text{where} \quad \tau_L^R \equiv \frac{\left(2\sqrt{2}-1\right)(a-w)+2(1+2\Delta)\sqrt{2}+\sqrt{\kappa}}{8\sqrt{2}}, \\ \text{and} \quad \kappa \equiv \left(12\sqrt{2}-23\right)(a-w)^2 + 4\sqrt{2}\left\{\left(2\sqrt{2}-1\right)(1-4\Delta)+4\Delta\right\}(a-w) + 8(1+2\Delta)^2\end{aligned}$$

is obtained as a condition for $\widetilde{T} \geq T_L$. Thus, $\widetilde{T} \geq T_L$ holds if and only if $\tau \geq \tau_L^R$ holds.

By Eqs. (4) and (18), we have $\widetilde{x}_L^B = \frac{(1-T)(a-w)-(1-t)\Delta+(T-t)w\underline{\alpha}}{(3-2T-t-\overline{\alpha}(T-t)}$. Then, we can compute,

$$\begin{aligned}\widetilde{\Pi}_L^B - \Pi_L^{O*} &\propto \frac{(1-T)(a-w)-(1-t)\Delta+(T-t)w\underline{\alpha}}{(3-2T-t)-\overline{\alpha}(T-t)} - \frac{(1-T)(a-w)-(1-t)\tau-(1-t)\Delta}{3-2T-t} \\ &\propto (1-T)(3-2T-t)\tau + \underline{\alpha}(T-t)\{(1-T)(a+w-\tau)+(1-t)(w-\Delta)\} > 0, \qquad (46)\end{aligned}$$

which means that the supply by firm L under regime B is greater than that of pre-FTA. This concludes proposition 6. ■

A.7 Proof of Proposition 7

First, we show that ROO always benefit consumers and then move to show the impact of ROO compared to the case without ROO.

From the proof of proposition 2, we know that forming an FTA with ROO increases consumer surplus when $T \leq T^*$ holds. It also shows that the FTA increases exports under regime I if and only if $T < T_{CS}$ and $\underline{\alpha} \geq \underline{\alpha}^I$ hold. By taking the difference of $\widetilde{T}$ and T_{CS},

$$\begin{aligned}\widetilde{T} - T_{CS} &= \frac{a-w-\tau+2\Delta}{a-w+2\tau} - \frac{(a-w+2\Delta-2\tau)\sqrt{2}}{a-w} \\ &\propto -(\sqrt{2}-1)(a-w)(a-w+2\Delta) - \left(a-w+4\Delta\sqrt{2}\right)\tau + 4\tau^2\sqrt{2}, \end{aligned} \tag{47}$$

is obtained, which takes the minimum value at $\tau = \frac{a-w-4\Delta\sqrt{2}}{8\sqrt{2}}$. Note that the possible range of tariff is $\tau \leq \left(0, \min\left\{\frac{a-w+2\Delta}{2}, a-w-\left(\frac{1-t}{1-T}\right)\Delta\right\}\right]$ and we confirm,

$$\frac{a-w-4\Delta\sqrt{2}}{8\sqrt{2}} < \min\left\{\frac{a-w+2\Delta}{2}, a-w-\left(\frac{1-t}{1-T}\right)\Delta\right\} \equiv \tau^{max}. \tag{48}$$

When $T \leq \frac{a-w-2(2-t)\Delta}{a-w-\Delta}$ holds, $\tau^{max} = \frac{a-w+2\Delta}{2}$ at which,

$$\widetilde{T} - T_{CS}\Big|_{\tau=\frac{a-w+2\Delta}{2}} = -\frac{(a-w)(a-w+2\Delta)}{2} < 0, \tag{49}$$

holds. Alternatively, when $T > \frac{a-w-2(2-t)\Delta}{a-w-\Delta}$ holds, the difference is,

$$\begin{aligned}\widetilde{T} - T_{CS}\Big|_{\tau=a-w-\left(\frac{1-t}{1-T}\right)\Delta} &= -\left(\sqrt{2}-1\right)(a-w)(a-w+2\Delta) \\ &\quad + \left(a-w-\left(\frac{1-t}{1-T}\right)\Delta\right)\left\{\left(4\sqrt{2}-1\right)(a-w) - 4\sqrt{2}\left(\frac{2-T-t}{1-T}\right)\Delta\right\}, \end{aligned} \tag{50}$$

and the each element of the second term is decreasing in T, which means the difference gets the maximum value at lowest T. Note that the lowest T in this case is $\frac{a-w-2(2-t)\Delta}{a-w-\Delta}$, and by definition, $\tau^{max} = \frac{a-w+2\Delta}{2} = a-w-\left(\frac{1-t}{1-T}\right)\Delta$, which implies,

$$\widetilde{T} - T_{CS}\Big|_{\tau=a-w-\left(\frac{1-t}{1-T}\right)\Delta} = \widetilde{T} - T_{CS}\Big|_{\tau=\frac{a-w+2\Delta}{2}} < 0. \tag{51}$$

Thus, the total exports increased due to the FTA formation under regime I.

Under regime N, we can easily get by Eqs.(34) and (35),

$$\left(\widetilde{x}_M^N + \widetilde{x}_L^N\right) - \left(x_M^{O*} + x_L^{O*}\right) = \frac{(1-T)\tau}{3-2T-t} > 0, \tag{52}$$

so that the consumers benefit from the FTA when $T \geq \widetilde{T}$ and $\underline{\alpha}^N$ hold, or under regime N. In the case of regime B, the change in the total exports is computed as,

$$\begin{aligned}\left(\widetilde{x}_M^B + \widetilde{x}_L^B\right) - \left(\widetilde{x}_M^{O*} + \widetilde{x}_L^{O*}\right) &= \frac{(2-T-t)(a-w)+(1-t)\Delta-(T-t)a\underline{\alpha}}{3-2T-t-(T-t)\underline{\alpha}} - \frac{(2-T-t)(a-w-\tau)+(1-t)\Delta}{3-2T-t} \gtreqless 0 \\ &\iff \underline{\alpha} \lesseqgtr \frac{(3-2T-t)(2-T-t)\tau}{(T-t)\left\{(1-T)a-(1-t)\Delta+(2-T-t)(w+\tau)\right\}} \equiv \underline{\alpha}_{CS}\end{aligned} \tag{53}$$

and

$$\underline{\alpha}_{CS} - \underline{\alpha}^N \propto (1-T)\left\{(1-T)(a+w)+(1-t)(w-\Delta+\tau)\right\} + (1-t)\left\{2(2-T-t)w+(3-T-2t)\tau\right\} > 0 \tag{54}$$

holds so that the FTA increases the total exports to country F. These prove the first part of the proposition.

Next, we check the impact of ROO by comparing the total exports without ROO. Note that ROO does not have any impact on total exports when $T \leq \widehat{T}$ as the MNE chooses input production in country H irrespective of ROO. When $\widehat{T} < T < \widetilde{T}$ and $\underline{\alpha} \geq \underline{\alpha}^I$ hold, the change in the total exports is,

$$\begin{aligned}\left(\widetilde{x}_M^I + \widetilde{x}_L^I\right) - \left(\widehat{x}_M^O + \widehat{x}_L^O\right) &= \frac{3(a-w)}{4} - \frac{(2-T-t)(a-w)+(1-t)\Delta}{3-2T-t} \gtreqless 0 \\ &\iff T \lesseqgtr \frac{1+t}{2} - \frac{2(1-t)\Delta}{a-w} \equiv T_{CS}^R.\end{aligned} \tag{55}$$

By subtracting $\widehat{T}$ and $\widetilde{T}$ from T_{CS}^R,

$$T_{CS}^R - \widehat{T} \propto \left(\sqrt{2}-1\right)(a-w+2\Delta) > 0 \tag{56}$$

$$\begin{aligned}T_{CS}^R - \widetilde{T} &\propto \left(\sqrt{2}-1\right)(a-w+2\Delta) - 2\sqrt{2}\tau \gtreqless 0 \\ &\iff \tau \lesseqgtr \frac{(a-w+2\Delta)\left(\sqrt{2}-1\right)}{2\sqrt{2}} \equiv \tau_{CS}^R.\end{aligned} \tag{57}$$

Thus, when $T_{CS}^R < T < \widetilde{T}$ and $\underline{\alpha} \geq \underline{\alpha}^I$ hold, ROO decrease the exports compared to the case of the

Scheme/Regime	λ_M	λ_L	Export	Transfer price	MNE's post-tax profits	FTA Welfare
No FTA (w/ "$*$")						
Scheme I	1	1	x_i^{I*}	r^{I*}	Π_i^{I*}	W^{I*}
Scheme O	1	1	x_i^{O*}	r^{O*}	Π_i^{O*}	W^{O*}
FTA w/o ROO (w/ "$\widehat{\ }$")						
Scheme I	0	0	$\widehat{x}_i^I$	$\widehat{r}^I$	$\widehat{\Pi}_i^I$	$\widehat{W}^I$
Scheme O	0	0	$\widehat{x}_i^O$	$\widehat{r}^O$	$\widehat{\Pi}_i^O$	$\widehat{W}^O$
FTA w/ ROO (w/ "$\widetilde{\ }$")						
Regime I	0	0	$\widetilde{x}_i^I$	$\widetilde{r}^I$	$\widetilde{\Pi}_i^I$	$\widetilde{W}^I$
Regime N	1	0	$\widetilde{x}_i^N$	$\widetilde{r}^N$	$\widetilde{\Pi}_i^N$	$\widetilde{W}^N$
Regime B	0	0	$\widetilde{x}_i^B$	$\widetilde{r}^B$	$\widetilde{\Pi}_i^B$	$\widetilde{W}^B$

post-FTA without ROO, which can be possible only when $\tau > \tau_{CS}^R$ holds. Otherwise, they increase the amounts of exports.

Under regime N with ROO, we immediately obtain,

$$\left(\widetilde{x}_M^N + \widetilde{x}_L^N\right) - \left(x_M^{O*} + x_L^{O*}\right) = \frac{-(1-t)\tau}{3-2T-t} < 0, \tag{58}$$

while under regime B, the total exports is,

$$\left(\widetilde{x}_M^B + \widetilde{x}_L^B\right) = \frac{(2-T-t)(a-w) + (1-t)\Delta - (T-t)a\underline{\alpha}}{3-2T-t-(T-t)\underline{\alpha}}. \tag{59}$$

Note that $\left(\widetilde{x}_M^B + \widetilde{x}_L^B\right)\big|_{\underline{\alpha}=0} = \left(\widehat{x}_M^B + \widehat{x}_L^O\right)$ at $\underline{\alpha} = 0$ and

$$\frac{\partial\left(\widetilde{x}_M^B + \widetilde{x}_L^B\right)}{\partial\underline{\alpha}} \propto -(1-T)(a+w) - (1-t)(w-\Delta) < 0, \tag{60}$$

hold, which means the amount of the total exports with ROO is less than that without ROO. These conclude the proposition. ■

B. Additional Figures

The parameters used to draw the figures are $a = 3$, $w = 1$, $t = 0.1$, $\Delta = 1/32$, $\tau = (a-w)/8$, and $\delta = 2$ (for Fig. 9).

C. Key Symbols for notations

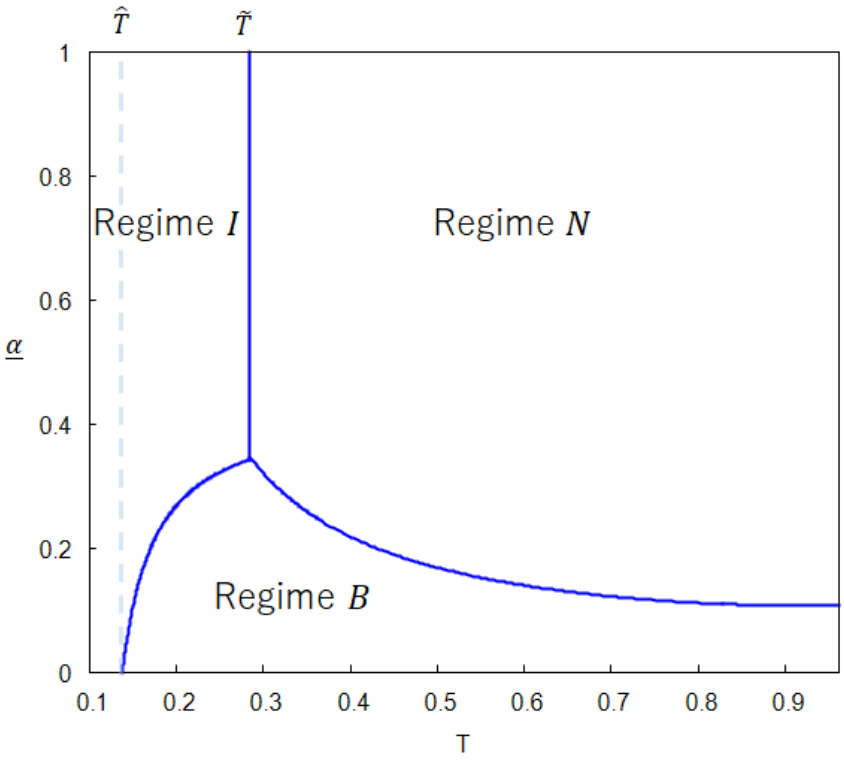

Figure 7: With home market competition

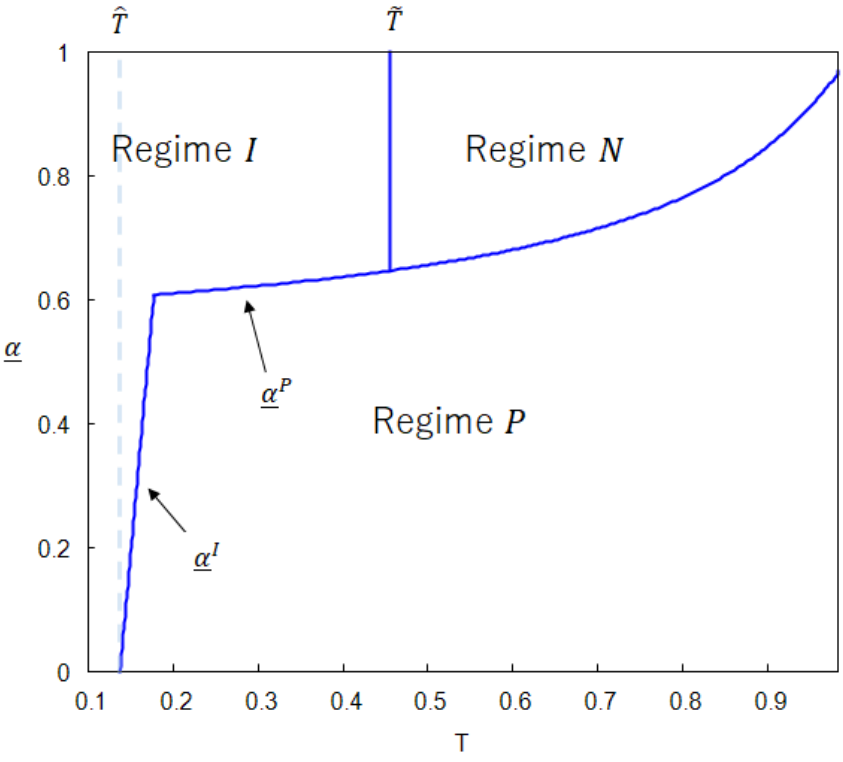

Figure 8: Partial procurement

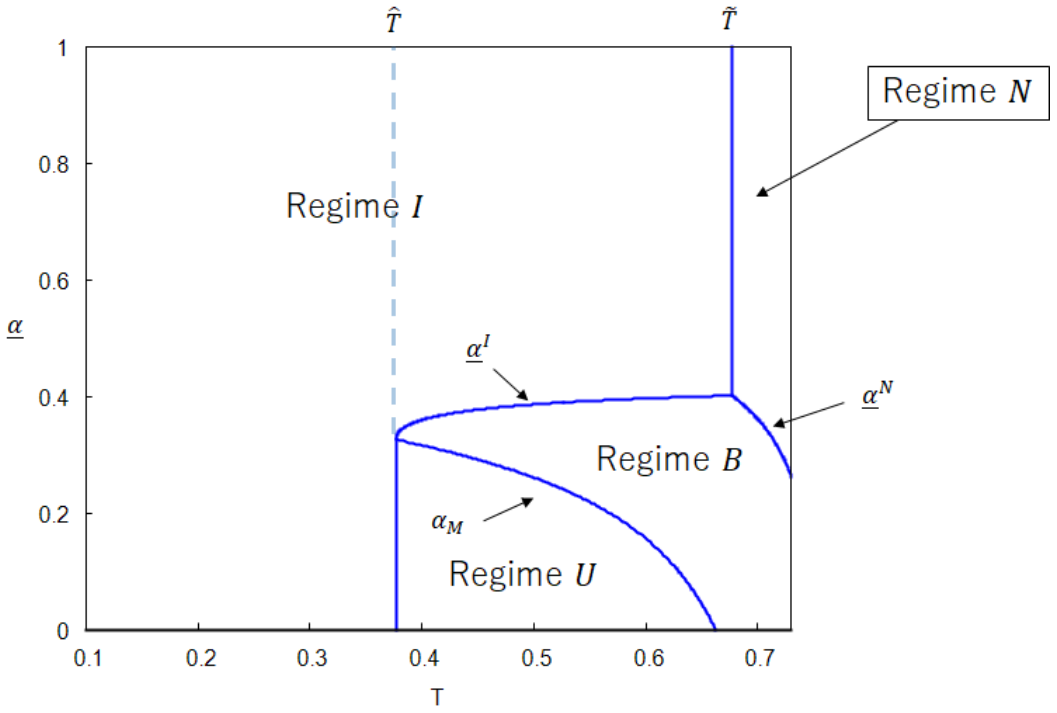

Figure 9: Concealment cost

References

Bauer, C.J., Langenmayr, D., 2013. Sorting into outsourcing: Are profits taxed at a gorilla's arm's length? Journal of International Economics 90, 326–336.

Blouin, J.L., Robinson, L.A., Seidman, J.K., 2018. Conflicting transfer pricing incentives and the role of coordination. Contemporary Accounting Research 35, 87–116.

Choe, C., Matsushima, N., 2013. The arm's length principle and tacit collusion. International Journal of Industrial Organization 31, 119–130.

Choi, J.P., Furusawa, T., Ishikawa, J., 2018. Transfer pricing and the arm's length principle under imperfect competition. CESifo Working Paper No. 7303.

Clausing, K.A., 2003. Tax-motivated transfer pricing and US intrafirm trade prices. Journal of Public Economics 87, 2207–2223.

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, 2335–65.

Cristea, A.D., Nguyen, D.X., 2016. Transfer pricing by multinational firms: New evidence from foreign firm ownerships. American Economic Journal: Economic Policy 8, 170–202.

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

Dürr, O.M., Göx, R.F., 2011. Strategic incentives for keeping one set of books in international transfer pricing. Journal of Economics & Management Strategy 20, 269–298.

Eden, L., 1998. Taxing multinationals: Transfer pricing and corporate income taxation in North America. University of Toronto Press.

Elitzur, R., Mintz, J., 1996. Transfer pricing rules and corporate tax competition. Journal of Public Economics 60, 401–422.

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).

European Commission, 2011. Proposal for a council directive on a common consolidated corporate tax base (CCCTB). COM (2011) 121 final .

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

Fershtman, C., Judd, K.L., 1987. Equilibrium incentives in oligopoly. The American Economic Review , 927–940.

Freund, C., Ornelas, E., 2010. Regional trade agreements. Annual Review of Economics 2, 139–166.

Hanson, G.H., Mataloni Jr, R.J., Slaughter, M.J., 2005. Vertical production networks in multinational firms. Review of Economics and statistics 87, 664–678.

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

Horst, T., 1971. The theory of the multinational firm: Optimal behavior under different tariff and tax rates. Journal of Political Economy 79, 1059–1072.

Hyde, C.E., Choe, C., 2005. Keeping two sets of books: The relationship between tax and incentive transfer prices. Journal of Economics & Management Strategy 14, 165–186.

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

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

Kant, C., 1988. Foreign subsidiary, transfer pricing and tariffs. Southern Economic Journal , 162–170.

Kato, H., Okoshi, H., 2019. Production location of multinational firms under transfer pricing: The impact of the arm's length principle. International tax and public finance 26, 835–871.

Krishna, K., Krueger, A., 1995. Implementing free trade areas: Rules of origin and hidden protection. Technical Report. National Bureau of Economic Research.

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., 2019. Rules of origin and consumer-hurting free trade agreements. Munich Discussion Papers .

Nielsen, S.B., Raimondos-Møller, P., Schjelderup, G., 2003. Formula apportionment and transfer pricing under oligopolistic competition. Journal of Public Economic Theory 5, 419–437.

Nielsen, S.B., Raimondos-Møller, P., Schjelderup, G., 2008. Taxes and decision rights in multinationals. Journal of Public Economic Theory 10, 245–258.

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

Schjelderup, G., Sorgard, L., 1997. Transfer pricing as a strategic device for decentralized multinationals. International Tax and Public Finance 4, 277–290.

Sklivas, S.D., 1987. The strategic choice of managerial incentives. The RAND Journal of Economics , 452–458.

Stoyanov, A., 2012. Tariff evasion and rules of origin violations under the Canada-US free trade agreement. Canadian Journal of Economics/Revue canadienne d'économique 45, 879–902.

Swenson, D.L., 2001. Tax reforms and evidence of transfer pricing. National Tax Journal , 7–25.

Takahashi, K., Urata, S., 2010. On the use of FTAs by japanese firms: Further evidence. Business and Politics 12, 1–15.

Tekin-Koru, A., Waldkirch, A., 2010. North–south integration and the location of foreign direct investment. Review of International Economics 18, 696–713.

Tintelnot, F., 2017. Global production with export platforms. The Quarterly Journal of Economics 132, 157–209.

Tomiura, E., 2007. Foreign outsourcing, exporting, and FDI: A productivity comparison at the firm level. Journal of International Economics 72, 113–127.

WCO, 2018. WCO guide to customs valuation and transfer pricing, World Customs Organization.