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

Hiroshi Mukunoki† Hirofumi Okoshi‡

May 19, 2021

Abstract

We explore the new roles of rules of origin (ROO) when multinational enterprises (MNEs) manipulate their transfer prices to avoid a high corporate tax. The ROO under a free trade agreement (FTA) require exporters to identify the origin of exports to be eligible for a preferential tariff rate. We find that a value-added criterion of ROO restricts abusive transfer pricing by MNEs. Interestingly, an FTA with ROO can induce MNEs to shift profits from a low- to high-tax country. Because the ROO augment tax revenues inside FTA countries, they can transform a welfare-reducing FTA into a welfare-improving one.

Keywords: Rules of origin; Free trade agreement; Transfer pricing; Profit shifting

JEL classification codes: F13; F15; F23; H26

*This study was conducted as a part of the Project "Analyses of Offshoring" undertaken at the Research Institute of Economy, Trade, and Industry (RIETI). We wish to thank Jay Pil Choi, Ruud Aloysius de Mooij, Carsten Eckel, Clemens Fuest, Taiji Furusawa, Andreas Haufler, Jung Hur, Jota Ishikawa, Hiro Kasahara, Yoshimasa Komoriya, Ngo Van Long, Kiyoshi Matsubara, Kaz Miyagiwa, Monika Mrazova, Martin Richardson, Kensuke Teshima, and the participants of the Canadian Economic Association meeting, RIETI, 58th congress of ERSA, Microeconomics Workshop at the University of Tokyo, Summer Workshop on Economic Theory at Otaru University of Commerce, and 21st annual conference of ETSG, Workshop on International Economics at Osaka University, 76th Annual Congress of IIPF. Hiroshi Mukunoki acknowledges financial support from JSPS KAKENHI (Grant Numbers JP19H00594 and JP20K01659). Hirofumi Okoshi acknowledges the financial support from Deutsche Forschungsgemeinschaft. (German Science Foundation GRK1928). The usual disclaimer applies.

†Gakushuin University, Faculty of Economics, Mejiro 1-5-1, Toshima-ku, Tokyo171-8588, Japan

‡Okayama University, Faculty of Economics, 3-1-1 Tsushima-naka, Kita-ku Okayama-shi, Okayama, 700-8530, Japan

1 Introduction

Tax avoidance by multinational enterprises (MNEs) has become controversial in the last two decades of rapid globalization. The Organisation for Economic Co-operation and Development (OECD) estimates that countries lose 4–10% of corporate income tax revenue annually because of profit shifting.[1] One way to shift profits across countries is to manipulate the price of intra-firm trade (transfer price), which is known as abusive transfer pricing. Because MNEs determine the prices of transactions among related companies, they manipulate these prices to decrease profits in high-tax countries and conversely increase profits in low-tax countries. Some empirical research has provided evidence of transfer pricing being used to save tax payments.[2] Because the taxes paid by firms are one of the main sources of government revenues, tax avoidance by MNEs has become a serious issue, as trade liberalization and the creation of global value chains increase intra-firm trade and provide MNEs with greater opportunities to redistribute their tax base to low-tax countries.

Our primarily focus is on how such losses of tax revenues are linked to trade liberalization driven by trade agreements. Trade agreements among countries facilitate firms' exports and imports. They also influence firm behaviors in other respects including transfer pricing and generate more complicated welfare effects. In particular, the specific rules needed to implement trade agreements complicate the effects of trade liberalization.

We focus on the rules of origin (ROO) of a free trade agreement (FTA), which require exporters in member countries making tariff-free exports to other member countries to prove that the exported products originated within the FTA.[3] To meet the ROO, firms may change their strategies such as their input procurement. Conconi et al. (2018) concludes that the ROO of the North American Free Trade Agreement (NAFTA) reduce imports of inputs from non-member

[1] See `http://www.oecd.org/tax/beps/`, accessed on March 11, 2020.

[2] For instance, Swenson (2001), Clausing (2003), Cristea and Nguyen (2016) and Davies et al. (2018) provided empirical evidence of the transfer price manipulation. Blouin et al. (2018) found that MNEs have conflicting motives for using transfer pricing to lower corporate tax and tariff payments.

[3] Regional trade agreements in goods are classified into FTAs and customs unions. Unlike customs unions, member countries of an FTA can set their own tariff schedule against non-member countries. This offers an opportunity for firms producing outside the FTA to save tariff payments by choosing as a transit country the member country whose tariff against non-member countries is low, and then re-exporting from that country to other FTA member countries whose tariffs against non-member countries are higher. Stoyanov (2012) presents evidence of firms' incentive to transship a good through FTA members. To forestall firms from tariff avoidance, FTA members stipulate ROO.

countries, suggesting that such rules cause inefficiency in input procurement. Considering tax avoidance, this also implies that the ROO can hinder MNEs from shifting profits within the firm because they may need to consider whether their intra-firm transactions satisfy the requirements of the ROO.

One way to prove the origin is to satisfy the value-added (VA) criterion, which is closely related to transfer price manipulation.[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 the input materials, which are used per unit of final good production and do not originate in the FTA. The VA criterion typically requires that the VA ratio $(p-r)/p$ is above the specified level. This method of calculating the VA content is called the "transaction value method." The value of the input materials depends on the transfer price if MNEs procure inputs from related companies outside FTA countries, and, hence, a VA criterion can constrain MNEs from engaging in tax avoidance through abusive transfer pricing. However, a VA criterion allows the MNE an option to meet the ROO without changing its input sources. This violates the principal purpose of the ROO to increase local procurement of inputs.

Although this possibility has been overlooked in the economic literature on transfer pricing and FTA, it has been noted by some policy researchers. LaNasa III (1996) stated that "[v]alue-added rules of origin may be circumvented by the use of transfer pricing, ... to increase the amount of local value added to ensure that the good qualifies as originating in the country of assembly, related parties could reduce the price of the imported materials used in the finished product." Eden (1998) examined the 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) indicated 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) similarly 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

[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 study focuses only on the VA criterion.

country in the free trade area."[5] The World Customs Organization notes that a disadvantage of the VA criterion of ROO is the possible exposure to transfer pricing.[6] A recent report by Deloitte Touche Tohmatsu Limited reports that exporters should consider an adjustment of transfer prices when making use of FTAs.[7]

As Estevadeordal and Suominen (2003) reported that 68 of the 87 FTAs they analyzed employ a VA criterion at least in a particular product category, the aforementioned statements suggest that the effect of FTAs on tax avoidance and welfare can be understood clearly if we investigate the role of the ROO in restricting the abusive use of transfer pricing. Given the experience of tax avoidance countries face, analyzing the anti-tax avoidance aspect of the ROO holds crucial policy implications. In reality, different groups of policymakers, namely, customs and tax authorities, are responsible for designing trade policies and regulating transfer pricing. The interaction between these two authorities has been rare. According to WCO (2018), ". . . the WCO is working with the OECD and World Bank Group to encourage Customs and tax administrations 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 an MNE's business, its compliance record and can make informed decisions on the collect revenue liability." Thus, the increasing number of FTAs and volume of intra-firm trade necessitates us to explore the relationship between transfer prices and ROO.

Against this backdrop, this study builds an international monopoly model to investigate an MNE's response to an FTA formation with two new elements: transfer pricing and ROO.[8] We consider a situation wherein an MNE produces final goods within an FTA member country and exports the goods to other FTA member countries. The MNE procures inputs from either an

[5]Some practitioners see the link as one factor to be considered, noting 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 3, 2018).

[6]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.

[7]The report states that ". . . in cases where the preferential calculation is based on the Value Added Rule and the required threshold is barely reached, an adjustment of transfer prices might lead to the loss of the preferential status of an article." See https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/tax/deloitte-ch-en-making-use-of-free-trade-agreements.pdf, accessed on May 1, 2021.

[8]If we consider local firms among FTA members and oligopoly in the final goods market, the fundamental properties of our results remain unchanged, although the analysis becomes more complicated. See Mukunoki and Okoshi (2019) for the oligopoly version of the model.

FTA member country or a low-tax country outside the FTA. In the absence of ROO, the MNE always prefers to produce inputs outside the FTA by itself and avoids tax by setting a high transfer price. However, the presence of ROO restricts the manipulation of the transfer price because a high transfer price reduces the VA ratio of the final product inside the FTA. Thus, the MNE chooses one of three options: (i) fully manipulating its transfer price to avoid tax payments at the expense of the preferential tariff of the FTA, (ii) procuring inputs inside an FTA to comply with the ROO and eliminate tariffs, or (iii) adjusting its transfer price to comply with the ROO to eliminate tariffs and pursue partial tax avoidance.[9] This model exhibits the MNE's choice of the "tariff elimination versus tax avoidance" via its location choice of producing inputs and/or transfer price manipulation.[10]

When the MNE chooses the second option, it no longer avoids high tax. When it chooses the third option, its transfer price deviates from the optimal abusive transfer price, which retains part of the MNE's tax base in the high-tax country. As the ROO restrict the abusive use of transfer pricing through either a change in input procurement or an adjustment of the transfer price, tax revenues in a high-tax country increase. Thus, the VA criterion works as an anti-tax avoidance policy. Interestingly, the direction of shifted profits can be from a low-tax country to a high-tax country when the MNE adjusts the transfer price to meet the VA criterion. Empirical analyses on transfer pricing should consider the possibility that the VA criterion of ROO affects transfer pricing.

The ROO can also increase the welfare gains of FTA countries because members can collect tax revenues from the MNE. Although consumers' gains from the FTA are smaller than those without ROO, tax revenues from the MNE can cover the smaller consumers' gains and the loss of tariff revenues. Our results present a new role of the ROO in preventing abusive transfer pricing and increasing the welfare gains of FTA formation for member countries.

[9] Although the MNE uses its transfer price for complying with the ROO, it can still shift profits from one country to another to save tax payments when the VA requirement is less stringent and the tax gap is large. Nevertheless, the overall tax payments become larger because the transfer price is suboptimal from the viewpoint of tax savings.

[10] We use the terms "tax rate" and "tax revenue" to represent the corporate tax rate and corporate tax revenue, respectively, which we distinguish from the tariff rate and tariff revenue.

1.1 Literature review

Our model contributes to the literature on transfer pricing policies since MNEs have been accused of tax avoidance activities. How to regulate transfer prices has been a central issue in policy debates. Several studies examine the effects of policies on transfer price manipulation. Elitzur and Mintz (1996) investigated the determinants of transfer prices when tax authorities use the cost-plus method to infer the appropriate transfer price. Nielsen et al. (2003) compared the use of transfer prices under two international tax systems, namely, separating account and formula apportionment.[11]

Bond and Gresik (2020) examined a high-tax country's unilateral adoption of border adjusted taxes and cash flow taxes when heterogeneous firms choose either arm's length transactions or to establish their own subsidiaries in their input sourcing. Choi et al. (2020) examined the effect of the arm's length principle on a monopolistic MNE's transfer pricing and tax competition.[12] As their focus was on direct regulation on transfer pricing, the role of ROO in preventing abusive transfer pricing is overlooked in the literature.

Our second contribution is to the literature on FTAs with ROO. Krishna and Krueger (1995) showed that the ROO may work as hidden protection against input suppliers outside the FTA. Ju and Krishna (2005) showed that ROO can either increase or decrease the price of FTA-made inputs, depending on the number of firms complying with the rules. However, their focus was on intermediate goods markets, and they did not consider how ROO affect consumers.

Demidova and Krishna (2008) extended the work of Ju and Krishna (2005) by including the heterogeneity of productivity of final good producers. They showed that productivity sorting ensures the negative relationship between the stringency of the rules and the demand for FTA-made inputs (i.e., wages). Ishikawa et al. (2007) focused on final good markets, showing that the ROO have a role to segment markets within the FTA, and that both inside and outside

[11]The traditional international corporate tax system is the separating account system that computes MNEs' national tax base by regarding intra-firm transactions as inter-firm transactions. Conversely, under the formula apportionment system, the tax payments of MNEs to one country depend on their consolidated tax base and the proportion of activity operated in the country. See more details in Chapter XVI and Article 86 of European Commission (2011).

[12]Bauer and Langenmayr (2013), Choe and Matsushima (2013), and Kato and Okoshi (2019) also investigated the effect of the arm's length principle on the input procurement decision, tacit collusion, and input production location, respectively.

firms producing final goods may benefit from the ROO at the cost of consumers. Mukunoki (2017) showed that an FTA with ROO may harm consumers if it changes outside firms' location decisions. Mukunoki and Okoshi (2021) investigated a firm's export price manipulation to comply with the ROO, particularly how an MNE's transfer price manipulation affects the inputs imported from outside the FTA. None of these studies, however, consider transfer price manipulation to meet ROO. Felbermayr et al. (2019) suggested that there is little rationale for ROO because tariff circumvention is not profitable for 86% of bilateral trade owing to the small differences in external tariffs and non-negligible transport costs. This study thus provides a new rationale for the ROO from the viewpoint of tax avoidance by an MNE.

This study also examines the connection between transfer pricing and trade policy. In this regard, Horst (1971) showed that the optimal transfer price is influenced by not only tax differentials but also 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. Subsequently, the optimal transfer price is influenced by both corporate tax avoidance and tariff avoidance.[13] Kant (1988) regarded the transfer price as a tool to repatriate profits when a foreign subsidiary is not fully owned by the parent firm. The study found that even when the tax rate in the home country is higher than that in the host country, an MNE has an incentive to remit all the profits earned in the low-tax host country. These studies, however, did not explicitly consider trade liberalization by forming an FTA, let alone the effects of ROO on transfer prices.

The rest of the paper is organized as follows. Section 2 presents the model and derive an equilibrium. Section 3 investigates the effects of FTA formation on profit shifting, consumers, and the MNE's profit. Section 4 discusses total welfare of member countries and the effect of input tariff. The last section concludes. Online Appendix explains the robustness of the main results by relaxing some key assumption.

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

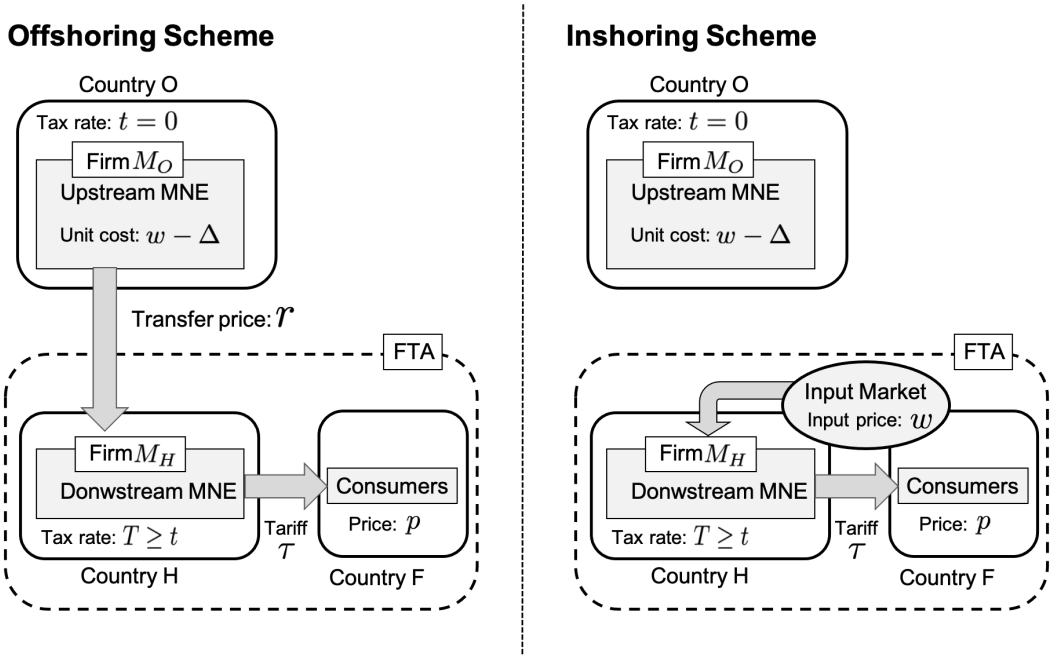

Figure 1: Model

2 Model

There are three countries, H, F, and O; countries H and F are potential FTA members. Fig. 1 illustrates the model. A single firm, an MNE, produces a final good using inputs and sells it in country F. For simplicity, the benchmark model ignores the output market in country H and focuses only on the consumers in country F. This assumption does not qualitatively change our main results as long as the two markets are segmented in the sense that the MNE can make a separate decision in each market. The representative consumer's utility in country F is given by $U = ax - \frac{x^2}{2}$, where x is the consumption of the final good. By utility maximization, the demand function becomes $x = a - p$.

One of the two member countries, country H, has a location advantage for final good production because of low factor prices, a large pool of skilled labor and so on. Therefore, country H hosts a downstream affiliate of the MNE (firm M_H). The MNE's upstream affiliate (firm M_O) is located in country O. Firm M_O may also produce an input for final good production, as explained below.

Firm M_O has already operated in country O and generates positive profits, $\overline{\pi}$, which are exogenously given. To produce the final product, firm M_H needs to procure one unit of inputs

for the production of one unit of final products.[14] Firm M_H can procure the input from a perfectly competitive input market inside FTA countries, which supply the input at the price of w. Alternatively, firm M_O located in country O can produce the input at the cost of $w - \Delta$. We assume $\Delta > 0$ and $\Delta \in (0, w]$. Therefore, input production in country O is more efficient than that in country H. This implies that the self-production of the input in country O gives the MNE not only a lower input cost but also a tax-saving opportunity via the manipulation of the transfer price, which is denoted by r. We assume away transfer pricing that realizes negative reported profits because tax authorities can audit tax avoidance.

Without the FTA, country F imposes a specific tariff, τ, on imports of the final good. We consider the case in which $\tau < a - w + \Delta$ holds to rule out zero output in the equilibrium. The governments in countries O and H, respectively, levy t and T as a corporate tax on reported profits as well.[15] To focus on the effect of FTA formation on the final good market, tariffs on inputs are assumed away.[16] Hereafter, we focus on the case in which $T \geq t$ holds. Without loss of generality, we set $t = 0$.

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 foreign direct investment (FDI), where the FDI firm exports from the host country to other countries. For example, see Tekin-Koru and Waldkirch (2010) for Mexican evidence of its increasing role as a host of export platform FDI. Tintelnot (2017) showed the share of output exported to countries outside the host country by the United States' MNEs. For instance, the share of exports in Belgium was 63% in 2004.

2.1 The equilibrium without ROO

Let us first derive the market equilibrium without the ROO in each scheme of the MNE choice. In the inshoring scheme, denoted as scheme I, the MNE purchases the input from local producers.

[14] We can consider a more general situation wherein the MNE uses a continuum of inputs, and determines the extent to which it uses intra-firm inputs for final good production. As explained in Online Appendix B.1, this modification does not change the qualitative results of the benchmark model.

[15] In this model, we postulate that the governments in countries O and H adopt a territorial tax system instead of a worldwide one. Most OECD countries, except Chile, Israel, Mexico, and South Korea, have adopted a territorial tax system. The United States moved from a worldwide tax system to a territorial tax system in December 2017.

[16] This assumption is relaxed in section 4.2.

Firm M_H earns profits under the cost of inputs w and tax rate T. In the offshoring scheme, denoted as scheme O, the MNE's upstream affiliate in country O, firm M_O, produces the input at the production cost of $w - \Delta$. Firm M_O sells the input to firm M_H at the input price, denoted by r. Thus, r is the transfer price of the MNE.

In the inshoring scheme, the MNE maximizes the post-tax profit, $\Pi = (1-T)(p - w - \lambda\tau)x + \overline{\pi}$, subject to p. The equilibrium price and sales are respectively given by $p^I = \frac{a+w+\lambda\tau}{2}$ and $x^I = \frac{a-w-\lambda\tau}{2}$ where λ is a state variable that takes zero if the MNE qualifies for an FTA tariff rate, and unity otherwise. By substituting them, the equilibrium post-tax profits under the inshoring scheme become

$$\Pi^I = (1-T)\underbrace{(x^I)^2}_{\pi_H^I} + \underbrace{\overline{\pi}}_{\pi_O^I}. \tag{1}$$

π_i^s represents the reported profits of firm M_i under scheme $s \in \{I, O\}$.

In the offshoring scheme, the MNE maximizes

$$\Pi^O = (1-T)\underbrace{(p - r - \lambda\tau)x}_{\pi_H^O} + \underbrace{[\{r - (w-\Delta)\}x + \overline{\pi}]}_{\pi_O^O} \tag{2}$$

with respect to r and p, subject to $\pi_H^O \geq 0$ and $\pi_O^O \geq 0$. Since $\frac{\partial \Pi_O}{\partial r} = Tx > 0$ always holds, the MNE is willing to set the optimal transfer price as high as possible. Therefore, the optimal abusive transfer price is set at the level that transfers all the profits earned in a high-tax country to a low-tax country, $r = p - \lambda\tau$.[17] Next, the post-tax profits are rewritten as $\Pi^O = [\{p - \lambda\tau - (w-\Delta)\}x + \overline{\pi}]$. By maximizing them with respect to p, the price, sales, and transfer price in equilibrium are given by $p^O = \frac{a+w-\Delta+\lambda\tau}{2}$, $x^O = \frac{a-w+\Delta-\lambda\tau}{2}$, and $r^O = \frac{a+w-\Delta-\lambda\tau}{2}$, respectively. Thus, the post-tax profits under the offshoring scheme are given by

$$\Pi^O = (x^O)^2 + \overline{\pi}. \tag{3}$$

Irrespective of the formation of an FTA, the MNE always prefers the offshoring scheme to

[17] We assume there is no cost of shifting profits across countries. This is a conventional way of determining the optimal transfer price in the literature, when the cost of profit shifting is absent. We relax this assumption by introducing a standard convex concealment cost in Online Appendix B.2.

the inshoring scheme, as

$$\Pi^O - \Pi^I = (x^O)^2 - (1-T)(x^I)^2 \geq 0 \tag{4}$$

holds given λ, because $x^O > x^I$. Intuitively, the offshoring generates more profits because procurement from its upstream affiliate provides the MNE with both efficient input production and the opportunity to shift profits.

For notational convenience, we use the superscript "$*$" for the variables in the pre-FTA case and "$\widehat{\ }$" for the post-FTA variables without ROO hereafter.

2.2 The equilibrium with ROO

Let us next consider an FTA formation with the ROO. As stated in the Introduction, our focus is on the VA criterion of ROO. Specifically, a VA criterion is applied to exports of the final good in the FTA. For notational convenience, we use "$\widetilde{\ }$" as a circumflex for the variables in the presence of the ROO. After an FTA is formed, firm M_H needs to meet the VA criterion to be eligible for the elimination of τ. Specifically, the ROO require firm M_H to add a proportion of at least $\underline{\alpha}$ ($\in (0,1]$) of the values of exported goods within FTA. There are three cases that we explain sequentially below.

First, if firm M_H chooses the offshoring of input production and sets an abusive transfer price, $r = p$, the VA ratio is always zero, which fails to meet the requirements of the ROO. Hence, the final goods exports of the MNE incur tariff τ, even after the formation of the FTA. We call this case scheme N (non-compliance).[18] The equilibrium outcomes of this scheme are obtained by setting $\lambda = 1$ in p^O, x^O, and r^O, as well as in the other corresponding welfare components. Second, if firm M_H chooses the inshoring of input production (scheme I), the VA ratio is 1, and it satisfies the requirement of the ROO. The equilibrium outcomes are obtained by setting $\lambda = 0$ in p^I and x^I.

Third, if firm M_H chooses the offshoring of the input production and sets p and r such that

[18]Some empirical evidence shows that not all firms use FTA tariffs, because of the existence of the ROO, that is, the effects of FTA formation are heterogeneous across firms. See, for example, Takahashi and Urata (2010) and Hayakawa et al. (2013).

they satisfy

$$\alpha \equiv \frac{p-r}{p} \geq \underline{\alpha}, \tag{5}$$

it complies with the ROO, and the tariff is thus eliminated. A combination of p and r that realizes (5) with strict inequality cannot be the equilibrium price. In section 2.1, we show that the MNE's post-tax profit after FTA formation is maximized by setting $p = r$ without the ROO. This implies that, as long as $\frac{p-r}{p} > \underline{\alpha}$ holds, the MNE always has an incentive to reduce $p - r$ by adjusting p and r.[19] Therefore, the MNE optimally sets p and r such that (5) is satisfied with equality, which yields:

$$r = (1 - \underline{\alpha})p. \tag{6}$$

We call this case scheme B (binding ROO).

By substituting (6), and $\lambda = 0$ into (2), the post-tax profits under scheme B are given by

$$\Pi^B = (1-T)\underbrace{\underline{\alpha}px}_{\pi_H^B} + \underbrace{[\{(1-\underline{\alpha})p - (w-\Delta)\}x + \overline{\pi}]}_{\pi_O^B} = \{1 - \underline{\alpha}T\}(p - c_M)x + \overline{\pi}, \tag{7}$$

where $c_M = \frac{w-\Delta}{1-\underline{\alpha}T} (> w - \Delta)$ represents the "perceived marginal cost" of producing the final good.[20] The perceived marginal cost is higher than the physical marginal cost, $w - \Delta$, as long as $\underline{\alpha}$ is positive and $T > 0$. Both an increase in the stringency of the ROO (i.e., $\underline{\alpha}$) and the tax in country H (i.e., T) increases the perceived marginal cost. We can interpret the perceived marginal cost as follows. Without any ROO, the MNE shifts all the profits to a low-tax country by setting $r = p$. From $r = p$, the introduction of the ROO decreases the transfer price by as much as $\underline{\alpha}p$ and increases the per-unit tax payments of the MNE by as much as $\underline{\alpha}pT > 0$. This means that the ROO decrease the MNE's marginal gains from selling the final good in terms of the post-tax profits. Therefore, the MNE becomes less aggressive in the product market under scheme B. The lower incentive to sell the final good is reflected in the perceived marginal cost. The price setting of the MNE is made with c_M instead of $w - \Delta$, that is, the binding ROO

[19] For instance, the MNE can always reduce its tax payments and increase the post-tax profits by increasing r and setting p such that it does not affect its sales in the product market.

[20] The terminology "perceived marginal cost" is often used in the analysis of vertically related industries in the context of industrial organization. See Choi et al. (2020) for an application of this terminology in the tax avoidance literature.

increase the equilibrium price, since $c_M > w - \Delta$ holds.

By maximizing (7) with respect to p, we obtain equilibrium price, sales, and transfer price as $\widetilde{p}^B = \frac{a+c_M}{2}$, $\widetilde{x}^B = \frac{a-c_M}{2}$, and $\widetilde{r}^B = (1-\underline{\alpha})\left(\frac{a+c_M}{2}\right)$, respectively. We thus have the following proposition:

Proposition 1. Suppose that an FTA is formed and the MNE chooses offshoring. The ROO induce the MNE to set a lower transfer price and a higher output price if it complies with the ROO.

Proof. See Appendix A.1.

The MNE adjusts both transfer price and output price to satisfy (5). If the MNE keeps $p = \widehat{p}^O$ and only lowers r to satisfy the VA ratio, the increase in the tax burden hurts the MNE more. If the MNE only raises p keeping $r = \widehat{r}^O$, it loses its profit in the product market even more. The rise in output price implies that the tariff pass-through is smaller with the ROO than without them because a part of tariff reduction is countered by an adjustment of the output price to meet the VA criterion. A smaller tariff pass-through reduces the consumers' gains from an FTA formation, as shown in section 3.2.

By substituting the equilibrium price and sales, the equilibrium post-tax profit of the MNE under scheme B becomes

$$\widetilde{\Pi}^B = \frac{\{(1-\underline{\alpha}T)a - w + \Delta\}^2}{4(1-\underline{\alpha}T)} + \overline{\pi}. \tag{8}$$

$\widetilde{\Pi}^B$ is a decreasing function of $\underline{\alpha}$because an increase in $\underline{\alpha}$ forces the MNE to set a transfer price and an output price that deviate more from the levels at which it avoids tax payments in the high-tax country.

2.3 The MNE's choice of scheme

In section 2.1, we have argued that the MNE always chooses scheme O before an FTA formation, and also after an FTA formation without the ROO. In an FTA formation with the ROO, schemes I, N, and B are possible equilibrium outcomes. Among the three possible schemes (I, N, and B), the MNE chooses the one that maximizes its profits.

Let us first compare $\widetilde{\Pi}^I$ with $\widetilde{\Pi}^N$. Since both profits are independent of the VA threshold, $\underline{\alpha}$, the tariff level and tax differential determine which profit is larger. The MNE faces a trade-off between tax avoidance and tariff avoidance. If the tax differential is large, the MNE prefers scheme N to scheme I because of the stronger incentive to avoid tax payments in country H. If the tax differential is small, scheme I is preferable for the MNE. Thus, there exists a unique threshold of T, $\widetilde{T}$, such that $\widetilde{\Pi}^I = \widetilde{\Pi}^N$ holds. As a higher tariff discourages the MNE from choosing scheme N, $\frac{\partial \widetilde{T}}{\partial \tau} > 0$ holds.[21]

Next, we compare the profits in scheme B with those in schemes N and I. $\widetilde{\Pi}^B$ is decreasing in $\underline{\alpha}$, and $\widetilde{\Pi}^B = \widehat{\Pi}^O$ holds at $\underline{\alpha} = 0$, which is larger than $\widetilde{\Pi}^N$ and $\widetilde{\Pi}^I$. Therefore, we can derive a unique threshold, $\underline{\alpha}^N$ (resp. $\underline{\alpha}^I$), above which the MNE prefers scheme N (resp. scheme I) to scheme B. Intuitively, under a less-strict ROO, the MNE prefers scheme B to schemes N and I because adjusting the transfer price to comply with those ROO becomes less costly as the VA criterion becomes less stringent. In other words, the profits of MNE from tariff elimination, by adjusting the transfer price, become smaller as the FTA is attached to more stringent ROO.

Putting the above comparisons together, we characterize the equilibrium outcomes as follows:

Proposition 2. The MNE chooses offshoring (scheme O) before an FTA formation or after an FTA formation without the ROO. After an FTA formation with the ROO, the MNE chooses (i) inshoring (scheme I) if $T \leq \widetilde{T}$ and $\alpha > \underline{\alpha}^I$ hold, (ii) offshoring and its exports incur a tariff (scheme N) if $\widetilde{T} < T$ and $\alpha > \underline{\alpha}^N$ hold, and (iii) offshoring, and it then adjusts its transfer price to meet the ROO (scheme B) if $\underline{\alpha} \leq \min\{\underline{\alpha}^I, \underline{\alpha}^N\}$ holds.

Proof. See Appendix A.2.

The equilibrium outcomes with the ROO is illustrated in Figure 2. The MNE always chooses the self-production of inputs before an FTA is formed. After an FTA is formed, this proposition suggests that the MNE may change its input procurement from self-production to the purchase of local inputs, despite higher production cost. As Conconi et al. (2018) showed,

[21] Formally, the threshold is calculated as $\widetilde{T} = 1 - \left(\frac{a-w+\Delta-\tau}{a-w}\right)^2 < 1$. We can confirm that $\widetilde{T}$ is positive if and only if $\Delta < \tau$ holds. To secure the existence of the equilibrium with scheme I, we additionally assume $\Delta < \tau$ hereafter.

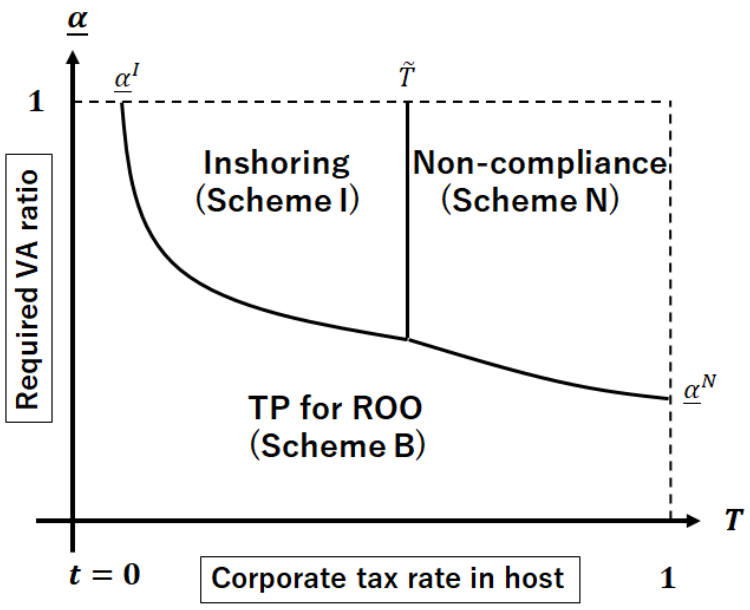

Figure 2: The equilibrium MNE's choice

the ROO lower the likelihood of input procurement from non-FTA countries. This "input trade diversion" corresponds to the area of scheme I in Figure 2.

Further, as Takahashi and Urata (2010) and Hayakawa et al. (2013) noted, some firms may not use FTA tariffs because of the burden of the ROO. This possibility corresponds to the area of scheme N in Figure 2. A standard explanation for the non-use of an FTA is that export firms must incur additional costs to meet the ROO. Our model suggests another burden of meeting the ROO: It increases tax payments by restricting the MNE's freedom to adjust its transfer price.

3 Effects of FTA formation

We have explored the equilibrium outcomes of an FTA with the ROO. This section analyzes how an FTA formation prevents the MNE's profit shifting and how it affects the consumer surplus and the MNE's profit.

3.1 FTA as an anti-tax avoidance policy

Let us explore how an FTA formation affects the MNE's tax avoidance. When the MNE engages in transfer pricing, an FTA with the ROO enables member countries to recover some of the

MNE's tax bases. When the MNE procures the input from the local input market (scheme I), there are no opportunities to shift profits, and all the tax bases are retained in country H. When it adjusts its transfer price to meet the VA criterion of the ROO, a part of the tax base is retained in country H because of the limited use of abusive transfer pricing.

Notably, we can confirm that the ROO reverse the direction of profit shifting across countries. To see this point more clearly, it is useful to decompose the optimal transfer prices into the "tax avoidance motive" and "tariff elimination motive." In the pre-FTA equilibrium, the optimal transfer price is always above the marginal cost of input production:

$$r = w - \Delta + \underbrace{\frac{a - w + \Delta - \lambda\tau}{2}}_{\text{Tax avoidance motive}}. \tag{9}$$

The second term of (9) represents the tax avoidance motive, which makes the transfer price as high as making the profit of the downstream affiliate of the MNE zero.

In scheme B of the post-FTA equilibrium, the tariff elimination motive counters the tax avoidance motive. The optimal transfer price is expressed as

$$\widetilde{r}^{B} = w - \Delta + \underbrace{\frac{a - w + \Delta}{2}}_{\text{Tax avoidance motive}} - \underbrace{\frac{\underline{\alpha}\{(1 - \underline{\alpha}T)a + (1 - T)(w - \Delta)\}}{2(1 - \underline{\alpha}T)}}_{\text{Tariff eliminative motive}}. \tag{10}$$

The third term of (10) captures the tariff elimination motive, which is zero at $\underline{\alpha} = 0$ and increasing in $\underline{\alpha}$. If the tariff elimination motive is sufficiently large, such that $\widetilde{r}^{B}$ is lower than $w - \Delta$, then the profits of the MNE shift from a low-tax country to a high-tax country, which is in sharp contrast to the conventional effect of transfer pricing.

Therefore, the direction of profit shifting relies on the size of the two motives. Indeed, we can derive a unique threshold of $\underline{\alpha}$, $\underline{\alpha}^{r}$, such that $\widetilde{r}^{B} < w - \Delta$ holds and profits shift from a high-tax country to a low-tax country if $\underline{\alpha} > \underline{\alpha}^{r}$ holds. Figure 3 illustrates the reversal of profit shifting.[22] The dotted curve represents $\underline{\alpha}^{r}$, and the dotted area in the figure represents the case in which profits flow from a low-tax country to a high-tax country. The following proposition summarizes the effect on tax revenue.

[22] We use the following parameters for the figure: $a = 1$, $w = \frac{1}{2}$, $\Delta = \frac{1}{32}$, and $\tau = \frac{1}{4}$.

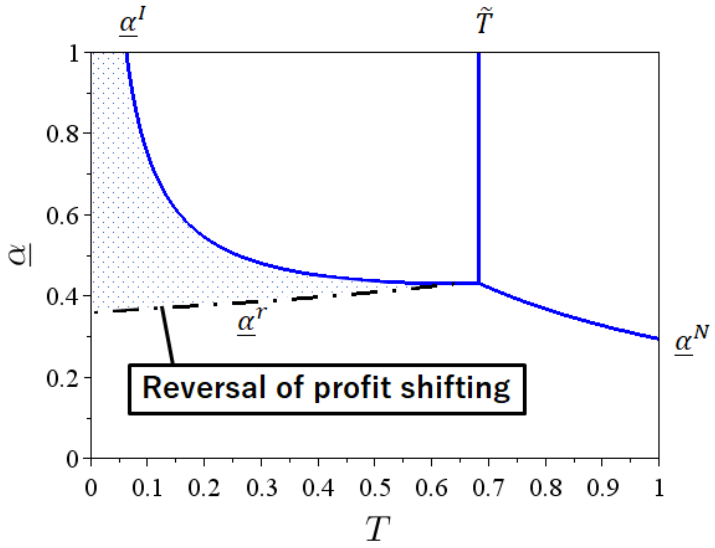

Figure 3: The direction of the MNE's shifted profits

Proposition 3. An FTA formation with the ROO reduces the profits of the MNE that are shifted from a high-tax country to a low-tax country if the MNE uses an FTA tariff. The MNE shifts its profits from a low-tax country to a high-tax country if the tariff elimination motive of transfer pricing is sufficiently large.

Proof. See Appendix A.3.

This proposition sheds new light on the role of the ROO that has been overlooked in policy debates. As Proposition 2 shows, an FTA formation with the ROO can induce the MNE to abandon the self-production of inputs as an opportunity to avoid tax.[23] This result suggests that a VA criterion provides another channel to keep MNEs away from tax avoidance by restricting the extent of abusive transfer pricing. Although the main purpose of imposing the ROO is to prevent trade circumvention, the ROO also play a role in preventing tax avoidance.

Furthermore, Proposition 3 provides a new empirical implication for estimating transfer pricing. As our model shows, the optimal transfer price depends on the stringency of the VA criterion, suggesting that the observed transfer prices can reflect not only the tax avoidance motive but also the tariff elimination motive.[24]

[23]This effect is observed in the other two criteria of ROO, that is, the change in tariff classification criterion and specific process criterion.

[24]Although there are other ways of shifting profits, such as using internal debt and making royalty payments, the tariff elimination motive behind the transfer pricing of tangible assets remains.

3.2 Effects on consumer surplus and the MNE's overall profit

Let us next explore the effect of an FTA formation on consumers and the MNE's overall profit. Under scheme I, FTA formation increases the marginal cost of production from $w - \Delta$ to w because the MNE changes the location of its input procurement. However, the FTA formation also eliminates the tariff, τ, faced by the MNE. We can easily confirm the following: The MNE chooses scheme I only if $\Delta < \tau$ holds (see footnote 21). Therefore, an FTA always decreases the MNE's marginal cost of exports whenever scheme I becomes the equilibrium outcome, and it always increases the exports of the MNE.[25] Tariff elimination increases the free-on-board (f.o.b) price, $p - \tau$, but its degree is always less than the tariff level, and it decreases the consumer price. Thus, an FTA formation benefits consumers and the MNE under scheme I.

Under scheme B, the MNE also faces a higher marginal cost because the perceived marginal cost is higher than $w - \Delta$. As in scheme I, however, the MNE chooses scheme B only if the cost reduction from tariff elimination dominates the increase in the marginal cost of production (see Appendix A.2 for details). Therefore, the FTA always increases the exports of the MNE whenever scheme B becomes the equilibrium outcome. As before, the tariff elimination increases the f.o.b price and decreases the consumer price. Thus, an FTA formation benefits consumers and the MNE under scheme B.

Under scheme N, the situation is the same as that in the pre-FTA equilibrium, and the FTA has no effects on the equilibrium outcomes. Putting the two cases together, we have the following proposition:

Proposition 4. An FTA formation with the ROO always benefits consumers and the MNE if the MNE uses an FTA tariff and has no effect on consumers and the MNE otherwise. The presence of the ROO decreases both consumers' and the MNE's gains.

Proof. See Appendix A.4.

Although FTA formation is beneficial for consumers and the MNE, the ROO decrease their gains because of the increase in the production cost owing to the inefficient procurement of inputs (scheme I) or increase in the perceived marginal costs (Scheme B). In scheme B, the

[25] Specifically, the change in exports becomes $\Delta x^{I*} = \widetilde{x}^I - x^{O*} = \frac{\tau - \Delta}{2}$, which is positive if $\Delta < \tau$ holds.

MNE gives up full tax avoidance, and the adjustment of transfer price to meet the ROO increases the MNE's perceived marginal cost. We should recognize this export-decreasing effect of the ROO driven by the change in the MNE's transfer pricing.

4 Discussion

We have shown that an FTA with ROO can prevent the MNE's tax avoidance. In this section, we explore how FTA formation affects the total welfare of member countries. This study employs a partial-equilibrium model focusing on one industry, and we should be careful about evaluating the welfare impacts of FTA because there should also be gains/losses in other industries. Nevertheless, our analysis provides new element to consider the desirability of FTA in the real world.

Besides that, the benchmark model also assumes that country H does not impose a tariff on inputs imported from country O. The member countries may have an incentive to prevent tax avoidance by the MNE by setting a high tariff on inputs, which reduces the gains from producing inputs outside the FTA. This diminishes the role of the ROO in preventing abusive transfer pricing. We show that the main results of the benchmark model remain unchanged even if country O optimally sets the level of input tariff.

4.1 Total welfare of member countries

To explore the welfare effects of FTAs, we focus on the total welfare of FTA countries. The total welfare of FTA countries under scheme $s \in \{I, O, B, N\}$ is the sum of the consumer surplus in country F (CS_F^s), tax revenues of country H paid by the MNE (TR_H^s), and tariff revenues in country F (TR_F^s):

$$W^s = CS_F^s + TR_H^s + TR_F^s = \frac{(x^s)^2}{2} + T\pi_H^s + \lambda\tau x^s. \tag{11}$$

Total welfare does not include the post-tax profits of the MNE because it is owned by residents outside the FTA. We have the following proposition. See Online Appendix B.3 for the proof.

Proposition 5. An FTA formation without the ROO benefits member countries when the initial

tariff rate is high ($\tau > \tau^W$) and hurts them when it is low ($\tau < \tau^W$). An FTA formation with the ROO benefits member countries if (i) the post-FTA scheme is scheme I, and $T > \widetilde{T}^W$ holds or (ii) the post-FTA scheme is scheme B and $\underline{\alpha} > \underline{\alpha}^W$ holds. It has no effect on member countries if the post-FTA scheme is N. Otherwise, an FTA with the ROO hurts member countries.

Let us first explain the welfare effect of an FTA formation without ROO. The post-FTA equilibrium scheme is always scheme O. The member countries cannot collect tax revenues both before and after the FTA formation. The FTA generates a trade-off between an increase in the consumer surplus and disappearance of tariff revenues. When the initial tariff rate is high ($\tau > \tau^W$), the consumers' gains exceed tariff revenues and the FTA formation increases the total welfare of member countries.

The presence of the ROO changes the welfare effect of FTA formation. The FTA has no effect if the post-FTA equilibrium scheme is scheme N, but it affects the total welfare in scheme I or scheme B. As discussed in section 3.2, the ROO reduce consumer gains from FTA formation in country F in these schemes. However, the ROO also help generate tax revenues in country H if the MNE changes its input procurement from country O to country I, or it adjusts its transfer price to comply with the ROO. Thus, the ROO can either increase or decrease the welfare gains from FTA formation.

If the post-FTA scheme is scheme I, an FTA formation improves the total welfare of the FTA members when the tax gap is relatively high, and the positive effects from generating a tax revenue is large enough. Similarly, if the post-FTA scheme is scheme B and $\underline{\alpha}$ is high, then the MNE needs to adjust its transfer price to comply with the ROO. In this case, the positive effect from gaining the tax revenue is large enough to improve the total welfare.

We found there exists a case wherein an FTA without the ROO worsens the total welfare, but an FTA with the ROO improves it. There is also a case where an FTA with the ROO improves the total welfare, but an FTA without the ROO worsens it (see Online Appendix B.3 for details). However, note that this study employs a partial equilibrium model that focuses on one specific sector. In reality, an FTA should affect many sectors. The welfare effects discussed here explain only a part of the overall effects of the FTA. Nevertheless, the analysis of this study is distinct in that it suggests a new mechanism through which the ROO change the welfare effects of an

FTA formation.

4.2 A tariff on inputs

In the benchmark model, we have assumed that country H imposes no tariff on imported inputs. Given the zero external tariff, only the ROO can hinder the MNE's transfer pricing. However, country H may have an incentive to set an input tariff to prevent the loss of the tax revenue. Even if we consider an input tariff, and its degree is determined endogenously by country H, the main results of the baseline model holds, as long as the corporate tax in country H is not very high. See Online Appendix B.4 for the detailed discussion.

This is because, with a low corporate tax in country H, the main source of country H's welfare is not the tax revenue from the MNE, but the tariff revenue from the input tariff. Let ξ be a tariff on inputs in country H. Then, country H's welfare is either the tax revenue from the MNE ($TR_H^I = T\pi_H^I$) if inputs are produced in country H, or the tariff revenue ($TR_H^O = \xi x^O$) if inputs are produced in country O. If ξ exceeds the threshold level, ξ_M, then the MNE chooses inshoring over offshoring. If $T < \frac{\xi x^O}{\pi_H^I}$ is satisfied for some $\xi < \xi_M$ such that $TR_H^O > TR_H^I$ holds, the optimal external tariff to maximize tariff revenue becomes $\xi_T = \frac{a-w+\Delta-\tau}{2}$. In this case, country H does not impose a tariff that induces inshoring and prevents profit shifting. If T is large enough, $TR_H^O \leq TR_H^I$ always hold and country H sets the tariff that induces inshoring, $\xi \geq \xi_M$. As our main results are obtained when T is small, they are not qualitatively changed by introducing an input tariff of country H.

Although introducing an input tariff does not change the main results, it is worth discussing which members of the FTA, country H or country F, benefit more from imposing an import tariff that hinders transfer pricing (i.e., $\xi \geq \xi_M$). If $a - w - \tau < \Delta$ holds, then $\xi_T < \Delta$, and the MNE's marginal cost is lower under offshoring. In this case, country F prefers offshoring because the MNE's exports to country F is larger. If $a - w - \tau > \Delta$ holds, we have $\xi_T > \Delta$ and the MNE's marginal cost is lower under inshoring conditions. In this case, country F prefers inshoring. Therefore, when $TR_H^O \leq TR_H^I$ and $a - w - \tau < \Delta$ holds, an input tariff that prevents transfer pricing hurts country F, but it benefits country H. When $TR_H^O > TR_H^I$ and $a - w - \tau > \Delta$ holds, the prohibitive input tariff benefits country F but hurts country H. In other

cases, the prohibitive input tariff either benefits or hurts the two countries at the same time.

5 Conclusion

The recent proliferation of FTAs has been advancing trade liberalization among countries, and the cross-border economic activities of MNEs prevail globally. This study investigated a vertically integrated MNE's input production and pricing strategies to analyze the welfare effects of FTA formation when the MNE can manipulate its transfer price of intra-firm trade. As in previous studies, the MNE uses its transfer price to avoid a high corporate tax. After the formation of an FTA, however, there emerges another reason for transfer price manipulation in the presence of the ROO. Specifically, if the ROO of the FTA employ a VA criterion, the FTA induces the MNE to manipulate the transfer price to comply with the ROO and be eligible for tariff elimination.

When the VA criterion of the ROO is low, the MNE prefers transfer price manipulation, since adjusting the transfer price is straightforward. However, once the required VA level is high, the transfer price adjustment decreases the efficiency of tax avoidance, such that the manipulation of the transfer price for the ROO is suboptimal. If the tax gap between a country outside an FTA and a member country is large, the MNE produces a necessary input in the outside country at the expense of the FTA tariff rate because the gain from tax avoidance is large. If it is small, the MNE procures the input into the inside country to qualify for the FTA tariff. This result is in line with empirical and anecdotal evidence that (i) FTAs sometimes induce input relocation to inside FTA countries, (ii) not all firms export using the preferential tariffs of FTAs, and (iii) transfer price manipulation is a factor in the difference in corporate tax rates and the required VA criterion of ROO.

Our model also showed the possibility that ROO can prevent profit shifting by an MNE by either a change in procurement strategy or another use of transfer prices. The formation of FTAs with the ROO is expected to work as an effective policy to not only induce trade liberalization, but also keep MNEs away from tax avoidance.

A remarkable result is that the direction of the MNE's shifted profits is contrary to that under

common knowledge when the MNE manipulates the transfer price for the ROO. Although the ROO reduce consumers' gains of the FTA formation, member countries can benefit more with the ROO owing to the emergence of the MNE's tax base. There is a case where the ROO can transform a welfare-reducing FTA into a welfare-improving one.

There remains room for further research. We assumed that tax rates and tariff rates are exogenously given. It would be intriguing to investigate how the formation of an FTA affects the outcomes of tax competition among countries as well as the optimal tariffs set by FTA members. Another direction in which to extend the model is to examine the effects of regulations on transfer pricing, such as the arm's length principle, 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

By comparing the equilibrium output prices, we have $\widetilde{p}^B - \widehat{p}^O = \frac{c_M - (w-\Delta)}{2} > 0$. By comparing the equilibrium transfer prices, we obtain $\widetilde{r}^B - \widehat{r}^O = -\frac{\alpha[\{1-\alpha T\}a-(1-T)(w-\Delta)]}{2(1-\underline{\alpha}T)} < -\frac{(1-\alpha)Ta}{2(1-\underline{\alpha}T)} < 0$, where the first inequality is due to $a > (w - \Delta)$. ■

A.2 Proof of Proposition 2

Without the ROO, (4) indicates that the MNE always chooses scheme O before an FTA formation. With the ROO, the post-tax profits of the MNE under schemes N and I are given by

$$\widetilde{\Pi}^N = \frac{(a - w + \Delta - \tau)^2}{4} + \overline{\pi}, \tag{A-1}$$

$$\widetilde{\Pi}^I = \frac{(1 - T)(a - w)^2}{4} + \overline{\pi}. \tag{A-2}$$

The condition under which the MNE prefers scheme I to scheme N is given by

$$\widetilde{\Pi}^I - \widetilde{\Pi}^N > 0 \iff T < 1 - \left(\frac{a - w + \Delta - \tau}{a - w}\right)^2 \equiv \widetilde{T}. \tag{A-3}$$

From (8), we can easily confirm that the following inequality holds:

$$\widetilde{\Pi}^B|_{\underline{\alpha}=0} = \frac{(a-w+\Delta)^2}{4} + \overline{\pi} > \max\{\widetilde{\Pi}^N, \widetilde{\Pi}^I\}. \tag{A-4}$$

Further, the first derivative of $\widetilde{\Pi}^B$ with respect to $\underline{\alpha}$ is

$$\frac{\partial \widetilde{\Pi}^B}{\partial \underline{\alpha}} = -\frac{T\{(1-\underline{\alpha}T)a - w + \Delta\}(1-\underline{\alpha}T + w - \Delta)}{4(1-\underline{\alpha}T)} < 0. \tag{A-5}$$

Let $\underline{\alpha}^x$ denote the cutoff level of $\underline{\alpha}^x$ such that $\widetilde{x}^B = x^{O*}(= \widetilde{x}^N)$ holds. Specifically, we have

$$\widetilde{x}^B \gtreqless x^{O*} \iff \underline{\alpha} \lesseqgtr \frac{\tau}{(w-\Delta+\tau)T} \equiv \underline{\alpha}^x. \tag{A-6}$$

If evaluated at $\underline{\alpha} = \underline{\alpha}^x$, (8) becomes

$$\widetilde{\Pi}^B|_{\underline{\alpha}=\underline{\alpha}^x} = \frac{(w-\Delta)(a-w+\Delta-\tau)^2}{4(w-\Delta+\tau)} + \overline{\pi}\left(< \widetilde{\Pi}^N\right). \tag{A-7}$$

This implies that there exists a unique cutoff level of $\underline{\alpha}$, $\underline{\alpha}^N \in (0, \underline{\alpha}^x)$, such that $\widetilde{\Pi}^B \geq \widetilde{\Pi}^N$ holds with $\underline{\alpha} \leq \underline{\alpha}^N$ and $T \geq \widetilde{T}$. Moreover, remember that $\frac{\partial \widetilde{\Pi}^I}{\partial T} < 0$ and $\widetilde{\Pi}^I = \widetilde{\Pi}^N$ holds at $T = \widetilde{T}$. Then,

$$\widetilde{\Pi}^I > \widetilde{\Pi}^I|_{T=\widetilde{T}} = \widetilde{\Pi}^N > \widetilde{\Pi}^B|_{\underline{\alpha}=\underline{\alpha}^x} \tag{A-8}$$

holds for any $T \in [0, \widetilde{T}]$. Note that $\widetilde{\Pi}^B > \widetilde{\Pi}^I$ holds if the following condition is satisfied:

$$\widetilde{\Pi}^B|_{\underline{\alpha}=1} > \widetilde{\Pi}^I \iff T < 1 - \left(\frac{w-\Delta}{w}\right). \tag{A-9}$$

This implies that there exists a unique cutoff level of $\underline{\alpha}$, $\underline{\alpha}^I \in (0, \underline{\alpha}^x)$, such that $\widetilde{\Pi}^B \geq \widetilde{\Pi}^I$ holds with $\underline{\alpha} \leq \underline{\alpha}^I$ and $1 - \left(\frac{w-\Delta}{w}\right) \leq T < \widetilde{T}$. ■

A.3 Proof of Proposition 3

From (10), we obtain

$$\frac{\partial \widetilde{r}^B}{\partial \underline{\alpha}} = -\frac{(1-\underline{\alpha}T)^2 + (1-T)(w-\Delta)}{2(1-\underline{\alpha}T)^2} < 0. \tag{A-10}$$

Therefore, $\widetilde{r}^B = w - \Delta + \frac{a-w+\Delta}{2} > w - \Delta$ holds at $\underline{\alpha} = 0$ and $\widetilde{r}^B$ takes the minimum value at $\underline{\alpha} = 1$, which is given by

$$\widetilde{r}^B|_{\underline{\alpha}=1} = 0 < w - \Delta. \tag{A-11}$$

Scheme B is the equilibrium at any $\underline{\alpha}$ if $T < \widetilde{T}$ holds. Therefore, there exists a unique $\underline{\alpha}^r$ such that $\widetilde{r}^B < w - \Delta$ holds when $\underline{\alpha} > \underline{\alpha}^r$ holds. ■

A.4 Proof of Proposition 4

Under scheme I, the changes in the amount of supplies from the pre-FTA equilibrium to the post-FTA equilibrium without the ROO are

$$\widetilde{x}^I - x^{O*} = \frac{\tau - \Delta}{2} > 0, \tag{A-12}$$

$$\widetilde{x}^I - \widetilde{x}^O = -\frac{\Delta}{2} < 0, \tag{A-13}$$

because $\tau > \Delta$ holds. Under scheme B, the FTA formation increases the amount of exports to country F when $\underline{\alpha} < \underline{\alpha}^x$ holds. From Proposition 2, we know that $\underline{\alpha} < \underline{\alpha}^x$ holds under scheme B and we always have $\widetilde{x}^B > x^{O*}$. In addition, we can easily confirm that

$$\widetilde{x}^B - \widetilde{x}^O = -\frac{(w-\Delta)\underline{\alpha}T}{2(1-\underline{\alpha}T)} < 0 \tag{A-14}$$

holds. ■

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.

Bond, E.W., Gresik, T.A., 2020. Unilateral tax reform: Border adjusted taxes, cash flow taxes, and transfer pricing. Journal of Public Economics 184, 104160.

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., 2020. Transfer pricing regulation and tax competition. Journal of International Economics 127, 103367.

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.

Davies, R.B., Martin, J., Parenti, M., Toubal, F., 2018. Knocking on tax haven's door: Multinational firms and transfer pricing. Review of Economics and Statistics 100, 120–134.

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.

Felbermayr, G., Teti, F., Yalcin, E., 2019. Rules of origin and the profitability of trade deflection. Journal of International Economics 121, 103248.

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.

LaNasa III, J.A., 1996. Rules of origin and the uruguay round's effectiveness in harmonizing and regulating them. American Journal of International Law , 625–640.

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. Tariff elimination versus tax avoidance: Free trade agreements and transfer pricing. RIETI Discussion Paper Series 19-E-099 .

Mukunoki, H., Okoshi, H., 2021. Rules of origin and consumer-hurting free trade agreements. The World Economy forthcoming.

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.

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.

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