Online Appendix for "tariff elimination versus Tax Avoidance: Free Trade Agreements and Transfer Pricing"

Hiroshi Mukunoki Hirofumi Okoshi

B.1 Partial procurement of inputs

We have assumed that the MNE makes a binary choice about input procurement, that is, a "make all or buy all" choice. It would be more realistic to suppose that the MNE purchases some proportion of parts from local suppliers and procures the rest from intra-firm transactions, which we refer to as scheme P.

Suppose that the MNE uses a continuum of inputs indexed in the $[0, 1]$ space. Let $\beta \in [0, 1]$ denote the proportion of inputs that firm M_H procures from firm M_O in country O. This means that the $1 - \beta$ proportion of the input is procured within the FTA countries. The amount of the intra-firm trade becomes βx, and the modified VA ratio is given by $\frac{p-\beta r}{p}$. The MNE maximizes

$$\Pi = \beta\{r - (w - \Delta)\}x + \overline{\pi} + (1 - T)[\{p - \beta r - (1 - \beta)w\}x] \tag{B-1}$$

with respect to x, r, and β, and subject to $\pi_H \geq 0$, $\pi_O \geq 0$, and $\frac{p-\beta r}{p} \geq \underline{\alpha}$.

Since $\frac{\partial \Pi}{\partial r} > 0$ holds, the optimal transfer price, r^P, is set at the level such that $p - \beta r^P - (1 - \beta)w = 0$ is satisfied. We can also confirm that $\frac{\partial \Pi}{\partial \beta} > 0$ always holds, which implies that the MNE sets β as high as possible. Therefore, the optimal β becomes $\beta = \max[0, \beta^P]$, where $\beta^P \equiv 1 - \frac{\underline{\alpha} p}{w}$. Let x^P denote the corresponding optimal level of x. Then, the equilibrium post-tax profits in scheme P are

$$\Pi^P = \left(1 - \frac{\underline{\alpha}\Delta}{w}\right)(p - c_M^P)x^P + \overline{\pi}, \tag{B-2}$$

where $c_M^P \equiv \frac{w(w-\Delta)}{w-\underline{\alpha}\Delta}$ is the modified perceived marginal cost, which falls between $w - \Delta$ and w.

Even if the MNE can choose β, the optimal level of β can be zero and the main results of the benchmark model remain unchanged. Given that the MNE sets the abusive transfer price as r^P, the VA ratio with $r = r^P$ becomes $\frac{(1-\beta)w}{p}$. If $\frac{w}{p} \leq \underline{\alpha}$ holds, the MNE can never comply with the ROO when $\beta > 0$.

Because $\frac{w}{p} < 1$, there exists a unique cut-off of $\underline{\alpha}$, $\underline{\alpha}^P_{\beta=0}$, above which the MNE sets $\beta^P = 0$. Thus, schemes N, I, and B of the benchmark model are still the equilibrium outcomes when $\underline{\alpha} \geq \underline{\alpha}^P_{\beta=0}$ holds. When $\underline{\alpha} < \underline{\alpha}^P_{\beta=0}$ holds, the MNE chooses a positive β, and scheme B is replaced with scheme P.

Furthermore, if some key inputs must be produced outside the FTA for technical reasons, there should be the upper bound of β, which is denoted by $\overline{\beta}$. If $\overline{\beta} < \beta^P$ holds, then the MNE still manipulates the transfer price to comply with ROO even with $\beta > 0$, although to a lesser degree than in the case with $\beta = 0$.

B.2 Concealment costs for transfer price manipulation

In the benchmark model, we have assumed that the MNE can costlessly manipulate the transfer price. Here, we show that the assumption is not critical to obtain the main results.

In practice, MNEs need to explain the plausibility of transfer pricing to shift profits across countries. As MNEs shift more profits among countries, it becomes more difficult to explain why their intra-firm prices deviate from the appropriate prices, such as the prices charged at arm's length transaction. Following the literature on transfer pricing, we introduce the "concealment cost" in the case of offshoring, which is increasing in the gap between the transfer price and the production cost of inputs:

$$C(r, x^O) = \frac{\delta\{r - (w - \Delta)\}^2 x^O}{2}, \tag{B-3}$$

where the parameter δ captures the difficulty of concealing tax avoidance. For example, a higher δ reflects a well-enforced tax authority. The post-tax profits under offshoring are modified as

$$\begin{aligned} \Pi^O &= \{r - (w - \Delta)\}x^O + \overline{\pi} + (1 - T)[(p - r - \lambda_M \tau)x^O] - C(r, x^O), \\ &= \overline{\pi} + (1 - T)(p - c^C_M)x^O, \end{aligned} \tag{B-4}$$

where $c^C_M = \frac{(w-\Delta)+(1-T)\lambda\tau - Tr}{1-T} + \frac{\delta\{r-(w-\delta)\}^2}{2(1-T)}$ is the effective marginal cost.

The concealment cost can prevent the MNE from transferring all the profits from country H to O. In other words, the MNE can choose the transfer price and final good price, such that $p > r + \lambda\tau$ holds, even in the absence of the ROO. Given that some of the MNE's tax bases remain in country H, our welfare analysis needs several modifications.

First, the welfare effect of an FTA formation in the absence of the ROO now depends on country H's tax, T. Substituting the optimal level of r that maximizes Π^O into c_M^C, the perceived marginal cost in the equilibrium is calculated as $c_M^C = w - \Delta + \lambda\tau - \frac{T^2}{2\delta}$. Thus, as the tax difference between countries widens, the perceived marginal cost lowers. Since the MNE becomes less willing to increase r because of the concealment cost, it has an incentive to increase x^O and lower p, which saves the MNE's tax payments in country H by narrowing the gap between p and $r + \lambda\tau$. This incentive is reflected in the perceived marginal cost. Because the elimination of tariffs increases the gap between p and $r + \lambda\tau$, it gives the MNE an extra incentive to increase x^O to avoid tax payments. The increase in x^O benefits consumers, and, thereby, improves the joint welfare of member countries. Therefore, in the presence of concealment costs, an FTA formation without the ROO is more likely to benefit member countries as the tax gap widens.

Second, $p > r + \lambda\tau$ implies that the VA ratio is positive, even for the absence of the ROO. In the benchmark model, the MNE always chooses zero VA ratio in FTA countries in the absence of the ROO. This implies that the VA requirement of the ROO affects neither the MNE's transfer pricing nor its location choice when $\underline{\alpha}$ is sufficiently small. This is because the MNE has already satisfied the required VA ratio in this case.

Even if we consider these two elements, the nature of our results does not change. Specifically, an FTA formation without the ROO may harm member countries, and ROO can transform a welfare reducing FTA into a welfare-improving one. The opposite case is also possible, where the ROO transform a welfare-improving FTA into a welfare-reducing one.

B.3 Total welfare of member countries

Let us start with an FTA formation without the 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 change in the total welfare by the FTA formation is given by

$$\begin{aligned}\widehat{W}^O - W^{O*} &= \widehat{CS_F}^O - CS_F^{O*} - TR_F^{O*} \\ &= \frac{-2(a-w+\Delta)+3\tau}{8} \gtreqless 0 \iff \tau \gtreqless \frac{2(a-w+\Delta)}{3} \equiv \tau^W. \end{aligned} \tag{B-5}$$

As (B-5) shows, an FTA without the ROO generates a trade-off between an increase in the consumer surplus and disappearance of tariff revenues. When the initial tariff rate is high, the consumers' gains exceed the tariff revenues, and the FTA formation increases the total welfare of member countries. Therefore, an FTA formation without ROO benefits member countries when the initial tariff rate is high ($\tau > \tau^W$) and hurts them when it is low ($\tau < \tau^W$).

Let us discuss how the presence of the ROO changes the welfare effect of FTA formation. As discussed in section 3.2, the ROO reduce consumers' gains from an FTA formation in country F. However, the ROO also help generate tax revenues in country H if the MNE changes its input procurement from country O to country H, or adjusts its transfer price to comply with the rules. Thus, the rules can either increase or decrease the welfare gains from an FTA formation.

Total welfare under scheme I is sum of the consumer surplus and tax revenue from the MNE:

$$\widetilde{W}^I = \frac{(a-w)^2(1+2T)}{8}. \tag{B-6}$$

We have

$$\widetilde{W}^I \gtreqless W^{O*} \iff T \gtreqless \widetilde{T}^W \equiv \frac{2(a-w)(\Delta+\tau)-(\tau-\Delta)(\Delta+3\tau)}{2(a-w)^2}. \tag{B-7}$$

Total welfare under scheme B also includes tax revenue from the MNE, given by

$$\widetilde{W}^B = \frac{1}{8}\left\{a - \frac{w-\Delta}{1-\underline{\alpha}T}\right\}^2 + T\underline{\alpha}\left[\frac{a^2}{4} - \left\{\frac{w-\Delta}{2(1-\underline{\alpha}T)}\right\}^2\right]. \tag{B-8}$$

At $\underline{\alpha} = 0$, regime B is identical to the post-FTA equilibrium without the ROO ($\widetilde{W}^B = \widehat{W}$). Starting from $\underline{\alpha} = 0$, an increase in the stringency of the ROO has two opposite effects on $\widetilde{W}^B$. On the one hand, a stricter VA requirement reduces the transfer price, and, thereby, increases the tax revenue that country H collects. On the other hand, it diminishes consumers' gains from an FTA formation by increasing the MNE's perceived marginal cost and reducing the amount of exports. There is thus an inverted U-shaped relationship between $\widetilde{W}^B$ and $\underline{\alpha}$. Specifically, the former effect dominates the latter and $\frac{\partial \widetilde{W}^B}{\partial \underline{\alpha}} > 0$ holds when $\underline{\alpha}$ is small, whereas the latter effect dominates the former and $\frac{\partial \widetilde{W}^B}{\partial \underline{\alpha}} < 0$ holds when $\underline{\alpha}$ is large.[1] We can specify a threshold of $\underline{\alpha}$, $\underline{\alpha}^W$, at which $\widetilde{W}^B = W^{O*}$ holds.[2] When $\underline{\alpha}^W < \min[\underline{\alpha}^I, \underline{\alpha}^N]$ holds, we have $\widetilde{W}^B \gtreqless W^{O*} \Longleftrightarrow \underline{\alpha} \gtreqless \underline{\alpha}^W$.[3] If the post-FTA scheme is scheme N, an FTA formation with the ROO does not affect the total welfare of member countries.

Suppose $\tau < \tau^W$, with which an FTA without the ROO reduces total welfare. Since stricter ROO improve post-FTA welfare in scheme B, the post-FTA welfare can be larger than the pre-FTA welfare, even with $\tau < \tau^W$. The left figure of Figure B.1 provides a numerical example to show that $\underline{\alpha}^W < \min[\underline{\alpha}^I, \underline{\alpha}^N]$ holds.[4] In the shaded area in scheme B, an FTA formation improves the total welfare. Further, an FTA formation improves the total welfare in scheme I if $T > \widetilde{T}^W$ holds. Therefore, the ROO can transform a welfare-reducing FTA into a welfare-improving one because the increased tax revenue from the MNE compensates for the tariff revenue loss.

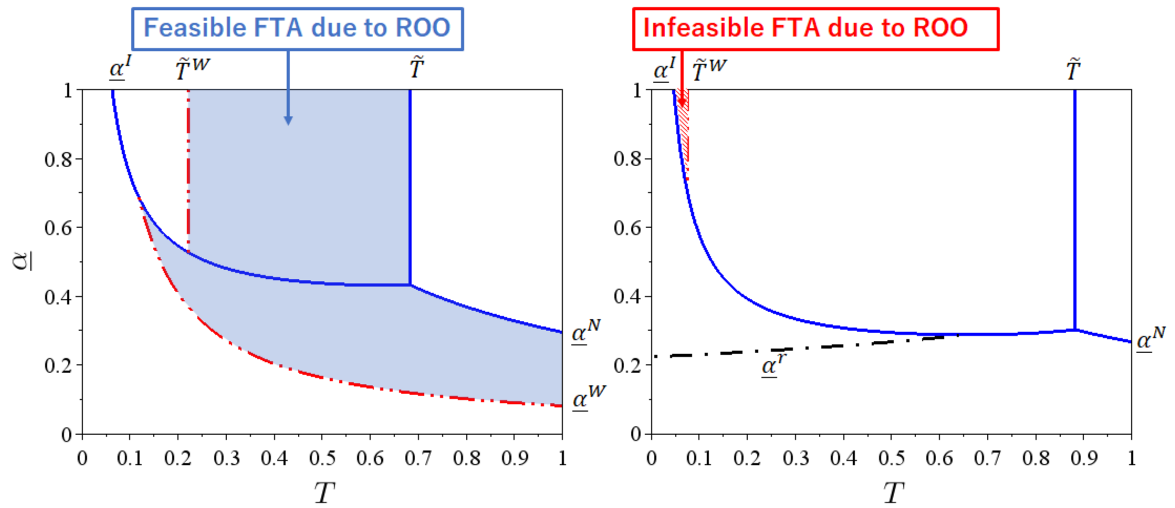
However, ROO may negatively affect total welfare and transform a welfare-improving FTA into a welfare-reducing one. The right figure of Figure B.1 corresponds to the case wherein $\tau > \tau^W$, and the formation of an FTA without the ROO is beneficial for member countries. The dotted curve in scheme B represents $\underline{\alpha}^r$, above which we see the reversal of profit shifting, as discussed in the previous section. The figure shows that an FTA stays feasible even if we take the ROO into account, such that the ROO increase the gains of forming an FTA under scheme B in the equilibrium because $\frac{\partial \widetilde{W}^B}{\partial \underline{\alpha}} > 0|_{\underline{\alpha}=0}$ holds. If both $\underline{\alpha}$ and T are high, such that scheme N is the equilibrium outcome, there is no welfare change from an FTA formation. Moreover, if

[1]Let $\underline{\alpha}^x$ be the threshold of $\underline{\alpha}$ such that the amount of exports is constant before and after the FTA is formed, $\widetilde{x}^B = x^{O*}$. Since we have $\frac{\partial \widetilde{W}^B}{\partial \underline{\alpha}}|_{\underline{\alpha}=0} > 0$, $\frac{\partial^2 \widetilde{W}^B}{\partial \underline{\alpha}^2} < 0$, and $\frac{\widetilde{W}^B}{\partial \underline{\alpha}}|_{\underline{\alpha}=\underline{\alpha}^x} < 0$, there exists a unique threshold, $\underline{\alpha}_0^W \in [0, \underline{\alpha}^x)$, such that $\frac{\widetilde{W}^B}{\partial \underline{\alpha}}|_{\underline{\alpha}=\underline{\alpha}_0^W} = 0$ holds.

[2]At $\underline{\alpha} = \underline{\alpha}^x$, an FTA formation does not change the MNE's exports of the final goods or consumer surplus. We can simplify the change in total welfare to the sum of tax revenue and the loss of tariff revenue, $\widetilde{W}^B - W^{O*}|_{\underline{\alpha}=\underline{\alpha}^x} = T\underline{\alpha}^x\widetilde{\pi}_H^B - \tau x^{O*} = \tau x^{O*}\left[\left\{\frac{a+w-\Delta+\tau}{2(w-\Delta+\tau)}\right\} - 1\right] > 0$. This means that a threshold $\underline{\alpha}^W \in [0, \underline{\alpha}^x)$ exists that satisfies $\widetilde{W}^B = W^{O*}$.

[3]Although $\widetilde{W}^B$ is an inverted U-shaped curve in $\underline{\alpha}$, we always have $\widetilde{W}^B > W^{O*}$ at $\underline{\alpha} = \min[\underline{\alpha}^I, \underline{\alpha}^N]$ because $\min[\underline{\alpha}^I, \underline{\alpha}^N] < \underline{\alpha}^x$ holds. This means that an FTA formation always improves total welfare for $\underline{\alpha}^W < \underline{\alpha} \leq \min[\underline{\alpha}^I, \underline{\alpha}^N]$.

[4]The parameters are set as follows: $a = 1$, $\Delta = \frac{1}{32}$, and $\tau = \frac{1}{4}$. The left figure is drawn with $w = \frac{1}{2}$, whereas the right figure uses $w = \frac{2}{3}$.



(i) $\tau < \tau^W$ Infeasible FTA w/o ROO (ii) $\tau > \tau^W$ Feasible FTA w/o ROO

Figure B.1: rules of origin and the welfare effect of an FTA formation

the equilibrium outcome is scheme I and $T \leq \widetilde{T}^W$ holds, the gains from increased tax revenue are smaller than the loss from the lower consumer surplus. In this case, the ROO transform a welfare-improving FTA formation into a welfare-reducing one.

B.4 A tariff on inputs

Because there is no market in country H, only tariff revenues or tax revenues are the welfare components of country H. Let ξ be a tariff on inputs in country H. If the MNE locates its upstream affiliates in country O, the amount of the imported inputs from country O becomes $x^O = \frac{a-w+\Delta-\xi-\lambda\tau}{2}$. Before the FTA formation, $\lambda = 0$. The MNE chooses to produce in country H if

$$\Pi^I \geq \Pi^O \iff \xi \geq \Delta + (a-w-\tau)(1-\sqrt{1-T}) \equiv \xi_M \quad \text{(B-9)}$$

holds. ξ_M increases with T and $\xi_M|_{T=0} = \Delta$. In scheme I, country H earns no tariff revenue, but it earns corporate tax revenues, which are given by TR_H^I. Given the input production in country O, the input tariff that maximizes country H's tariff revenue, $TR_H^O = \xi x^O$, is given by:

$$\frac{\partial TR_H^O}{\partial \xi} = 0 \quad \rightarrow \quad \xi_T = \frac{a-w+\Delta-\tau}{2}. \quad \text{(B-10)}$$

Because $\xi_T - \xi_M|_{T=0} = \frac{a-w-\Delta-\tau}{2}$, we have $\xi_T > \xi_M|_{T=0}$ if $a - w - \tau > \Delta$ holds. There exists a threshold level of T, T_ξ, such that $\xi_T < \xi_M$ holds if $T_\xi < T$ is satisfied. If $a - w - \tau \leq \Delta$ holds, $\xi_T < \xi_M$ is always satisfied as long as $T > 0$. Moreover, TR_H^O is an inverse U-shaped curve in ξ. We can confirm that

$$TR_H^O\big|_{\xi=\xi_T} \gtreqless TR_H^I \iff T \lesseqgtr T_1 \equiv \frac{(a-w+\Delta-\tau)^2}{2(a-w-\tau)^2}. \tag{B-11}$$

If $a - w - \tau \leq \Delta$ and $T < T_1$ hold, or if $a - w - \tau > \Delta$ and $T_\xi < T < T_1$ hold, country H prefers the offshoring scheme and sets its input tariff at $\xi = \xi_T$. Otherwise, TR_H^I is always higher than TR_H^O, irrespective of the level of T and ξ. In this case, country H prefers the inshoring scheme and sets the input tariff at $\xi \geq \xi_M$ to induce the MNE's inshoring.

After an FTA is formed, by (B-9) and (B-10), the optimal input tariff becomes larger. However, Article XXIV of GATT prohibits an increase in the external tariff above the pre-FTA level. Therefore, by replacing Δ with $\Delta' \equiv \Delta - \xi$, we qualitatively obtain the same results as the benchmark analysis.