Preferential Trade Agreements, Technology Adoption and the Speed of Attaining Free Trade*

Hiroshi Mukunoki†

Gakushuin University

March 23, 2012

Abstract

This paper compares sequential trade liberalization through a preferential trade agreement (PTA) and one-shot multilateral trade liberalization with respect to the speed with which countries attain multilateral free trade. We build a three-country oligopoly model, including one developing country whose domestic firm initially uses old technology. Firm-level adoption of new technology and country-level conclusions of trade agreements are endogenously determined. When a PTA that includes the developing country is feasible, but a multilateral trade agreement is infeasible due to prior to technology adoption in the developing country, a free trade area (FTA) accelerates technology adoption and the realization of multilateral free trade; however a customs union (CU) delays these processes. The opposite case is obtained if PTAs are infeasible prior to technology adoption, or if they are formed between developing countries. Even if increased realization of free trade via an FTA improves world welfare, a developed country may prefer liberalization via a CU.

Key words: free trade area, customs union, technology adoption, international oligopoly

JEL classification numbers: F12, F13, F15, O14

---

*I appreciate Professors Jim Brander, Taiji Furusawa, Arghya Ghosh, Jota Ishikawa, Chia-Hui Lu, Pascalis Raimondos-Møller, Friska Parulian, Larry Qiu, Dan Richards, John Ries, Barbara Spencer, Frank Stähler, Johan Stennek, Andrey Stoyanov, Aleksandr Vashchilko, Alan Woodland, and the participants of the seminars at the University of British Columbia, the UNSW International Trade Workshop, the International Industrial Organization Conference, the annual conference of the Canadian Economic Association, the annual conference of the Western Economic Association International, and the 6th Australasian Trade Workshop for helpful comments and suggestions. I began the research while I was visiting the Sauder School of Business, University of British Columbia. I appreciate their great support and hospitality. The usual disclaimer applies.

†*Corresponding author*: Faculty of Economics, Gakushuin University, Mejiro 1-5-1, Toshima-ku, Tokyo 171-8588, Japan; E-mail: hiroshi.mukunoki@gakushuin.ac.jp

1 Introduction

Recent decades have witnessed growth in the participation of developing countries in the multilateral trading system. These countries currently compose about two-thirds of the membership of the World Trade Organization (WTO). In contrast to the growing interest in multilateral trade cooperation, it is becoming more difficult to build consensus in trade negotiations among countries with different economic backgrounds. In fact, the WTO negotiations in the Doha Development Round have repeatedly broken down since its inception in December 2001.

One problem that thwarts multilateral trade liberalization is a technological disparity between developed countries and developing countries. A technology difference between the countries can be a source of comparative advantage and lead to the traditional gains from reciprocal trade liberalization, which benefits all countries. However, the technology gap may also be a source of a lopsided distribution of the gains from trade liberalization if markets are imperfectly competitive and intra-industry trade prevails in some industries.[1] For instance, suppose firms in different countries engage in international oligopolistic competition and the firms in foreign countries have more efficient production technologies than the domestic firms. Under this situation, a reciprocal tariff reduction results in a relatively small increase in the profit generated in the foreign markets and a relatively large decrease in the profit earned in the domestic market, which may reduce the profits of the domestic firms. If the loss of firms' profits is large enough to exceed the efficiency gains from trade liberalization, then the country will oppose trade liberalization. Under this circumstance, technological advancements by the technologically lagging countries play an important role in advancing multilateral trade negotiations. In fact, the WTO has placed great emphasis on "capacity building" by developing countries in the Doha Development Round.

Another notable feature of the recent world economy has been a surge in the number of preferential trade agreements (PTAs). Faced with the sluggish WTO negotiations, many countries now resort to PTAs as an alternative method to pursue reciprocal trade liberalization. As of October 2011, 312 PTAs were in effect and reported to the WTO, compared with 28 in 1990. Nearly all WTO members are members of PTAs, and many new PTAs have been formed between developed and developing countries. UNCTAD (2007) reports that 27% of PTAs are North-South PTAs. In addition to the North American Free Trade Area (NAFTA), the United States formed several bilateral and plurilateral free trade areas (FTAs) with Jordan, Morocco, and Central and Latin American countries. Japan has signed bilateral FTAs with Chile and several East Asian countries. In the 2004 and 2007 enlargements of the EU, ten Eastern European

[1] Although intra-industry trade has been well-observed in trade between developed countries, it is far from negligible in trade between developed countries and developing countries. For instance, Liao (2006) suggests that the share of the country-level intra-industry trade in the U.S. trade with Brazil, China, India, and Mexico was 0.62, 0.24, 0.41, and 0.70, respectively, in 2003. Tharakan and Kerstens (1995) found that there is horizontal intra-industry trade in a number of products between high-income countries and low-income countries.

countries, Cyprus, and Malta joined the world largest customs union (CU). The EU also has bilateral FTAs and CUs with several North African countries, Eastern European countries, and West Asian countries.

These contrasting trends in international trade agreements give rise to the question of whether the formation of North-South PTAs serves as a catalyst for technological advancements by developing countries and accelerates multilateral trade liberalization. In recent years, many researchers have confirmed the prevalence of international technology diffusion from developed to developing countries.[2] If technology diffusion narrows the technology gaps among countries over time, then it might be only a matter of time before developing countries support multilateral liberalization. However, even if technology diffusion guarantees the long-run feasibility of multilateral free trade, the time to reach the goal differs between a time path on which countries adhere to multilateralism and a time path on which countries pursue regionalism. Under this situation, the significance of PTAs should be evaluated not only in terms of their effects on the feasibility of multilateral free trade but also their effects on the speed with which countries achieve it. Although the speed issue in relation to PTAs has been recognized as one of the *dynamic time-path problems* of regionalism (Bhagwati, 1993; Srinivasan, 1998), it has been overlooked in the literature.[3]

To address this question, this paper investigates the firm-level incentives to adopt new technology and the country-level incentives to form trade agreements in a simple three-country model of international oligopoly. The model comprises two developed countries, in which a single domestic firm produces a good with an advanced technology, and one developing country, in which a single domestic firm produces a good with an old technology. Because of the technology gap, the developing country may suffer a welfare loss from reciprocal trade liberalization and oppose the formation of trade agreements. The firm in the developing country can adopt the advanced technology by paying a fixed cost, the amount of which depends on the timing of technology adoption. The firm decides the timing of technology adoption while anticipating how it affects the timing of trade agreements.

The effects of PTAs are explored by comparing different paths to multilateral free trade. The realization of multilateral free trade requires the conclusion of a multilateral trade agreement (MTA) among the three countries, which eliminates all remaining tariffs. On the MTA path, all three countries negotiate the MTA at each point of time. On the PTA path, the countries follow two-step liberalization toward multilateral free trade. Specifically, one of the developed countries and the developing country first negotiate the formation of a North-South PTA, under which the members of the PTA eliminate the tariffs on each

[2]See Keller (2004) for a survey of the literature.

[3]For instance, Srinivasan (1998) claims: "An alternative way of assessing the entire time path of equilibria is ... to rank time paths according to the time each takes to reach a global free-trading equilibrium. Thus, the answer to another of Bhagwati's questions (1993), namely, "Is regionalism quicker?" is affirmative, if one can show that a time path, based on some version of regionalism, minimizes the time to global free trade among all other feasible time paths to the same goal, including ones based on multilateralism." See Section 2 for the related literature on the dynamic time-path problems.

other's goods but maintain the tariffs on the nonmember country's good. If the North-South PTA is successfully formed, then the three countries subsequently negotiate the MTA from the next period.

On the PTA path, we consider two types of PTA: (i) an FTA under which members individually set tariffs on the nonmember; and (ii) a CU under which members jointly impose a common external tariff (CET) on the nonmember. Hence, we consider three paths to multilateral free trade: the MTA path, the FTA path, and the CU path. The focus of the paper is to specify which path will accomplish multilateral free trade the fastest.

Whereas the MTA path attains multilateral free trade in one step, the FTA path and the CU path need two-step agreements to attain multilateral free trade. This means that PTAs have their own inherent disadvantage with respect to the speed of attaining multilateral free trade. However, the comparison is not as simple as it may initially appear because different paths to multilateral free trade result in differences in the timing of the firm's technology adoption. The result of this research suggests that there is a possibility that the timing of attaining multilateral free trade and of technology adoption can be faster on the PTA path than on the MTA path. There is also a possibility that sequential trade liberalization on the PTA path can delay the timing of technology adoption and the realization of multilateral free trade. For instance, imagine the case in which the countries cannot form an MTA but can form a North-South FTA or a North-South CU before the firm in the developing country adopts the new technology. In this case, the timing of technology adoption and the timing of attaining multilateral free trade are faster on the FTA path than on the MTA path, whereas they are slower on the CU path than on the MTA path.

To understand these results, we should first describe the different natures of FTAs and CUs in their determination of external tariffs against the nonmember. Due to the *tariff-complementarity effect*, which is well known in the literature, the members of a PTA are willing to reduce their external tariffs from the pre-PTA level. In the case of the FTA, the tariff reduction is large enough to increase each member's imports from the nonmember. Because of its trade-creating nature, if the domestic firm uses the old technology and the efficiency gap between the new technology and the old technology is sufficiently large, the profit of the domestic firm is decreased by the formation of the FTA. However, in the case of CUs, there is another effect. For each member of a PTA, an increase in the partner's external tariff benefits the domestic firm because it increases the domestic firm's exports to the partner country. As member countries set a CET in CUs, they can coordinate their external tariffs such that they can internalize this positive externality effect. The effect, which is known as the *externality-internalization effect*, partly offsets the tariff-complementarity effect and limits the degree of tariff reductions. As a result, a CU always decreases each member's imports from the nonmember. This trade-diverting nature of the CU means that the firms in the member country always benefit from the CU formation.

A conflict of interests between the government and the firm in the developing country is another key to gaining insight into the results. Suppose the efficiency gap is sufficiently large so that the profit loss of the domestic firm from the MTA exceeds the consumer gains in the developing country. Under this situation, the developing country opposes the MTA until the fixed cost of technology adoption becomes sufficiently low to trigger the domestic firm's adoption of the new technology. Due to the aforementioned trade-creating nature of the FTA, there is a case where both the MTA and the North-South FTA decrease the profit of the firm in the developing country. However, the extent of the profit loss is relatively small in the case of the North-South FTA because it maintains its tariffs against the nonmember. Therefore, even if both agreements harm the domestic firm, there is a case in which the developing country rejects the MTA but supports the North-South FTA before technology adoption. As a result of the North-South FTA, the firm in the developing country earns a smaller profit with the old technology. This in turn increases the firm's gains from adopting the new technology.

Consequently, the North-South FTA accelerates the timing of technology adoption. Once the new technology is adopted, the three countries are willing to ratify the MTA. Therefore, the North-South FTA also accelerates the realization of multilateral free trade in this case. In contrast, because the North-South CU always increases the profit of the firms in the member countries, the firm in the developing country becomes reluctant to adopt the new technology because its profit using the old technology under the North-South CU is higher than its profit under no agreements. Due to these effects, the North-South FTA can be an accelerator to multilateral free trade whereas the North-South CU acts as a brake.

To evaluate the welfare consequences of the different time paths, we also calculate the discounted sum of social welfare along each path. Even if the FTA path that accelerates multilateral free trade is the most preferred and the CU path that delays multilateral free trade is the least preferred from the standpoint of world welfare, the developed country that will be a member of the PTAs may prefer the CU path because it yields short-run gains from being a member of the CU and also delays the technology adoption of the rival firm.

It is worth noting that the superiority of the FTA over the CU is reversed if the following is true: (i) the technology gap between the countries is sufficiently large that the developing country opposes both the formation of a PTA and the conclusion of an MTA before technology adoption, or (ii) the developed countries form a North-North PTA, rather than a North-South PTA. In these cases, the trade-creating nature of the FTA and the trade-diverting nature of the CU work in opposite directions.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 develops an international oligopoly model in which the firm determines the timing of technology adoption and countries negotiate trade agreements given the firm's adoption decision. Section 4 derives the equilibrium of each path separately. Section 5 compares the three paths and identifies the fastest path to attaining multilateral free trade. Welfare implications of the comparison are also discussed. Section

6 explores the results under a number of alternative setups including the PTA-paths with North-North PTAs, the case with the lower bound of the adoption cost, and the model that allows discriminatory tariffs under the case of no agreements. Section 7 summarizes the paper and offers concluding remarks. The Appendix provides proofs of the lemmas and propositions.

2 Related Literature

This paper contributes to the literature on PTAs and, in particular, to a recent stream of the literature that has investigated the relationship between regionalism and multilateralism. Some papers have investigated how PTAs affect multilateral tariff cooperation in a repeated-game framework (e.g., Bagwell and Staiger, 1999; Bond et al., 2001; Saggi, 2006). Some other papers have examined whether PTAs become "stumbling blocks" in the sense that they make an initially feasible multilateral free trade infeasible (e.g., Levy, 1997; Krishna, 1998; Ornelas, 2005a; Mukunoki and Tachi, 2006; Furusawa and Konishi, 2007). There are also a few papers that have used political-economy models to point out that PTAs can be "building blocks" in that they make an initially infeasible multilateral free trade feasible (e.g., Orneals, 2005b; Aghion et al., 2007; Saggi and Yildiz, 2010). This paper is distinct from those papers in that we focus on the speed of attaining multilateral free trade. In particular, this paper shows that even if countries can reach multilateral free trade irrespective of the path they follow, the speed at which multilateral free trade is achieved differs depending on the paths chosen, which means that the countries generate different interim welfare.[4] Furthermore, it can be shown that if the level of adoption cost does not approach zero in the long run, our model can produce both "stumbling-block" and "building-block" results (see Section **??** for details).

As is briefly discussed in the introduction, the tariff-complementarity effect of PTAs is a key mechanism behind the main results. The effect was named by Bagwell and Staiger (1999), who considered a partial equilibrium model under perfect competition. The effect has been found to exist in a variety of trade models including general equilibrium models such as those of Kennan and Riezman (1990) and Bond et al. (2004), and international oligopoly models such as those of Yi (2000), Mukunoki (2004), Ornelas (2005a, 2005b), and Saggi (2006). Estevadeordal et al. (2008) and Calvo-Pardo, et al. (2009) found the empirical existence of the tariff-complementarity effect in Latin American countries.

The paper also relates to the literature on trade and technology adoption. Miyagiwa and Ohno (1995) investigated the effects of permanent protections and temporary protections on the technology adoption of the domestic firm. Yeaple (2005) showed that a reduction in trade costs could induce firms to employ

[4]Freund (2000) compared free trade reached through expanding a PTA (which corresponds to the PTA path in our model) with free trade accomplished by multilateral negotiations. She found that, with sunk costs, the welfare of the original members of a PTA and worldwide *after the realization* of global free trade is greater if it is reached through the PTA path.

more advanced technologies, expand trade volume, and increase the wage premium paid to highly skilled workers. Crowley (2006) examined how safeguard tariffs and antidumping duties affect the outcomes of the technology adoption game between a domestic firm and a foreign firm. Ederington and McCalman (2008, 2009) explored how firm heterogeneity and a decline in the number of firms in an industry evolves as a result of international trade. A distinguishing feature of this paper is the investigation of a linkage between firm-level technology adoption and the government-level formation of trade agreements.[5] This linkage, which is uniquely explored in this paper, yields novel policy implications.

In this paper, increased competition due to trade liberalization promotes technology upgrading at the firm-level. This effect is supported by recent empirical studies, such as those of Fernandes (2007), Teshima (2009), and Utar and Ruiz (2010). López-Acevedo (2002) suggested that NAFTA appeared to have a positive and a significant impact on the likelihood of technology adoption in Mexico.

3 The Basic Model

We consider a three-country, international oligopoly model. Let $\Omega = \{A, B, C\}$ denote the set of countries. In each country, a single firm produces a good and supplies it to both home and foreign markets. The firms produce a homogeneous good, and they are named such that Firm i is the domestic firm of Country i ($i \in \Omega$). The inverse demand of the good in Country i is time-invariant and given by $p_i = a - \sum_j q_i^j$, where q_i^j is the amount of the good that Firm j ($j \in \Omega$) supplies in Country i. The markets are segmented so that each firm makes a separate decision in each market.

Let τ_i^j (≥ 0) denote the specific tariff imposed by Country i on the imports from Country j. The instantaneous profit earned by Firm i from selling the good is given by $\pi_i = \sum_{k \in \Omega}(p_k - \theta_i - \tau_k^i)q_k^i$, where θ_i is the unit production cost of Firm i. We set $\tau_k^i = 0$ if $i = k$. The instantaneous welfare of each country is denoted by W_i, which consists of consumer surplus, the domestic firm's profit, and tariff revenues.

The time t is a continuous variable defined on $t \in [0, +\infty)$. Sometime before $t = 0$, a new technology that reduced the variable costs of production became available. We assume that Firms B and C have already adopted the new technology by $t = 0$, and that their unit costs of production are given by $\theta_B = \theta_C = \underline{\theta}$ (≥ 0), respectively. In contrast, Firm A has not adopted the new technology by $t = 0$, and its unit cost of production with the old technology is given by $\theta_A = \overline{\theta}$ ($> \underline{\theta}$). For expositional convenience, we occasionally use the term "developed country" or "North" to represent Country B or Country C, and the term "developing country" or "South" to represent Country A.

By adopting the new technology, Firm A can reduce its unit production cost from $\overline{\theta}$ to $\underline{\theta}$. Firm A's decision regarding technology adoption is modeled upon a framework developed by Reinganum (1981) and

[5] Cabrales and Motta (2001) examined how firms choose product specifications if they anticipate the (exogenously given) pace of trade liberalization in a partial equilibrium model of vertical product differentiation.

Fudenberg and Tirole (1985). Firm A's adoption of the new technology at time t requires a one-time fixed cost denoted by $k(t)$. We assume that $k(0) = \overline{k}$ is sufficiently high because Country A initially lacks the specific skills to implement the new technology.[6] We also assume that $k'(t) < 0$ and $k''(t) \geq 0$. This means that the fixed cost of technology adoption declines exogenously over time, although the rate of decline of the adoption cost either remains constant or slows down. The declining fixed cost would occur because Firm A accumulates knowledge about the new technology or some complementary technologies become available.

We consider four trade regimes: (i) regime N represents **no agreements** in which all three countries impose nondiscriminatory tariffs; (ii) regime F represents a **North-South FTA** in which Countries A and B reciprocally eliminate the tariffs between them while they independently choose their tariffs against country C; (iii) regime U represents a **North-South CU** in which Countries A and B reciprocally eliminate the tariffs between them and also set a CET against country C; (iv) regime M represents **multilateral free trade** in which trade between any pair of countries is free from tariffs. Let $\Gamma = \{N, F, U, M\}$ denote the set of trade regimes. The North-North FTA and the North-North CU between Countries B and C are discussed in Section **??**.

Countries have not formed any trade agreement by $t = 0$. Trade agreements change the trade regime. The formation of the North-South FTA changes the trade regime from N to F, while the formation of the North-South CU changes it from N to U. Given that no PTAs are formed, an MTA changes the trade regime from N to M. Given that a North-South FTA or a North-South CU is in place, however, the subsequent MTA changes the trade regime from F to M and from U to M, respectively.

We assume that Firm A correctly anticipates future events and commits to the timing of technology adoption, T, at the beginning of $t = 0$. In particular, the firm takes into account how the adoption of the new technology changes the timing of future trade agreements.[7] At every point in time, there are three stages. In stage 1, countries negotiate a trade agreement if multilateral free trade has not been realized, given the technology of the firms. The trade agreement is formed if and only if it increases the intertemporal welfare of all countries involved in the trade negotiation.[8] Countries cannot arrange

[6] Keller (2004) has pointed out that a firm (or a country) needs to have a certain complementary skill to adopt foreign technology successfully.

[7] The assumption allows us to preclude possible time inconsistency problems of policy implementation that complicate the analysis. Suppose the case in which a trade agreement improves the welfare of a participating country only if Firm A adopts the new technology. If Firm A cannot commit to the timing of technology adoption, then the country may form a welfare-reducing trade agreement in the anticipation that it will accelerate the timing of future technology adoption. Such a trade agreement, however, is not credible because if the domestic firm does not adopt the new technology, it is not optimal ex post for the country to sign the agreement. See the papers by Matsuyama (1990) and Miyagiwa and Ohno (1999), which investigated the time-inconsistency problem of trade policies.

[8] We assume that the negotiation process is efficient in the sense that a trade agreement is formed whenever it benefits all countries involved in the negotiation.

international transfers of welfare, and trade agreements are immediately implemented once they are formed. After the current trade regime is determined, each country chooses the level of tariffs in stage 2, given the current technologies of the firms. In stage 3, the firms play a quantity-setting game at every point in time, given the current technologies and tariff levels.[9]

3.1 Determinations of tariff levels

Let us first investigate the determination of tariffs at each point in time. Given the current trade regime, each country sets its tariffs to maximize its instantaneous welfare. Let $\pi_i(\theta_A, \boldsymbol{\tau})$ and $W_i(\theta_A, \boldsymbol{\tau})$ respectively denote the instantaneous profit of Firm i and the instantaneous welfare of Country i in the equilibrium of the stage-3 subgame, given the unit cost of Firm A, θ_A ($\in \{\overline{\theta}, \underline{\theta}\}$), and the tariff vector, $\boldsymbol{\tau} := (\tau_A^B, \tau_A^C; \tau_B^A, \tau_B^C; \tau_C^A, \tau_C^B)$.

Under the case of no agreements, the corresponding tariff vector is denoted as $\boldsymbol{\tau}^N = (\tau_A, \tau_A; \tau_B, \tau_B; \tau_C, \tau_C)$. The tariffs have no superscripts because the Most Favored Nation (MFN) principle of the GATT/WTO prohibits countries from discriminating in terms of setting different tariffs on the same good imported from different countries. Article XXIV of the GATT/WTO allows the formation of PTAs as an exception to the MFN clause.[10] In the absence of trade agreements, the optimal MFN tariff set by country i is given by $\widetilde{\tau}_i^N(\theta_A) := \arg\max_{\tau_i} W_i\left(\theta_A, \boldsymbol{\tau}^N\right)$ ($i \in \Omega$).[11] By substituting the optimal tariffs into $\pi_i(\theta_A, \boldsymbol{\tau}^N)$ and $W_i\left(\theta_A, \boldsymbol{\tau}^N\right)$, the equilibrium profit of each firm and the equilibrium welfare of each country in the stage-2 subgame become a function of θ_A, which are respectively denoted by $\pi_i^N(\theta_A)$ and $W_i^N(\theta_A)$. We consider the situation where $\overline{\theta} < \overline{\lambda} := (a + 6\underline{\theta})/7$ is always satisfied, so each firm can sell a positive amount in all markets under the case of no agreements.[12]

If the North-South FTA is formed, then the corresponding tariff vector under the North-South FTA is given by $\boldsymbol{\tau}^F = (0, \tau_A; 0, \tau_B; \tau_C, \tau_C)$, and country i's optimal tariff is obtained by $\widetilde{\tau}_i^F(\theta_A) := \arg\max_{\tau_i} W_i\left(\theta_A, \boldsymbol{\tau}^F\right)$. Correspondingly, the instantaneous profit of each firm and the instantaneous welfare of each country in the equilibrium of the stage-2 subgame are respectively denoted by $\pi_i^F(\theta_A)$ and $W_i^F(\theta_A)$.

---

[9] We assume that countries follow Markov strategies in their choices of tariff level and that firms also follow Markov strategies in their quantity settings. With this assumption, we can ignore possible implicit collusions between countries in terms of tariff setting, as well as between firms.

[10] The generalized system of preferences of the GATT/WTO is another exception to the MFN principle. It allows countries to impose tariffs on imports from developing countries that are lower than the tariffs they impose on the imports from other countries. We discuss this case in Section **??**.

[11] Because the markets are segmented and the marginal costs of production are constant, strategic interactions in tariff setting between countries are absent. If markets are integrated, the optimal tariffs of each country depend on those of other countries. See Mukunoki (2004), for example.

[12] Let $q_i^{j,N}(\theta_A)$ denote the equilibrium sales of good i in country j given θ_A ($i, j \in \Omega$). It can be verified that $q_A^{B,N}(\theta_A) = q_A^{C,N}(\theta_A)$ will represent the smallest amount among the equilibrium sales of firms. As $q_A^{B,N}(\theta_A) = q_A^{C,N}(\theta_A) = (a - 7\underline{\theta} + 6\overline{\theta})/10$, $q_A^{B,N}(\theta_A) = q_A^{C,N}(\theta_A) > 0$ holds if and only if $\overline{\lambda} = (a + 6\underline{\theta})/7 > \overline{\theta}$ holds.

If Countries A and B form the North-South CU, then the corresponding tariff vector becomes $\boldsymbol{\tau}^U = (0, \overline{\tau}; 0, \overline{\tau}; \tau_C, \tau_C)$, where $\overline{\tau}$ is the CET of the CU. We assume the members choose the level of the CET to maximize the joint welfare of the member countries. Then, the optimal CET is obtained by $\overline{\tau}^U(\theta_A) := \arg\max_{\overline{\tau}}\{W_A\left(\theta_A, \boldsymbol{\tau}^U\right) + W_B\left(\theta_A, \boldsymbol{\tau}^U\right)\}$ and the optimal tariff of Country $C$is given by $\widetilde{\tau}_C^U(\theta_A) := \arg\max_{\tau_C} W_C\left(\theta_A, \boldsymbol{\tau}^U\right)$. By substituting these tariffs, the instantaneous profit of each firm and the instantaneous welfare of each country under the North-South CU are respectively given by $\pi_i^U(\theta_A)$ and $W_i^U(\theta_A)$.

Finally, if multilateral free trade is realized, the corresponding tariff vector is the null vector and the instantaneous profit and the instantaneous welfare in the equilibrium of the stage-2 subgame are respectively given by $\pi_i^M(\theta_A) = \pi_i(\theta_A, \mathbf{0})$ and $W_i^M(\theta_A) = W_i(\theta_A, \mathbf{0})$.

By comparing the optimal tariffs under different trade regimes, we have the following lemma.

Lemma 1 *Given θ_A, we have (i) $\widetilde{\tau}_h^N(\theta_A) > \overline{\tau}^U(\theta_A) > \widetilde{\tau}_h^F(\theta_A)$ $(h \in \Omega \backslash C)$, (ii) $\widetilde{\tau}_B^F(\theta_A) \geq \widetilde{\tau}_A^F(\theta_A)$ with equality if $\theta_A = \underline{\theta}$, and (iii) $\widetilde{\tau}_C^N(\theta_A) = \widetilde{\tau}_C^F(\theta_A) = \widetilde{\tau}_C^U(\theta_A)$.*

Given θ_A, the members' optimal external tariffs under the North-South FTA and the common external tariff of the North-South CU are both lower than their respective MFN tariffs.[13] This is due to the *tariff-complementarity effect*, which is common in the literature.[14] While the MFN tariff targets imports from all countries, the external tariffs of PTAs only target imports from the nonmember country. As a result, an increase in the external tariff only decreases the imports from the nonmember and it increases the imports from the partner country. Therefore, members' incentives to raise their tariffs to earn tariff revenues and those to protect the domestic firm become weaker after the formation of PTAs. Meanwhile, the consumer loss due to an increase in the external tariff is smaller than the consumer loss due to an equal increase in the MFN tariff. As the former effect dominates the latter, the members of a PTA autonomously set their external tariffs below the pre-PTA level.

We should mention another effect that is specific to CUs. If a member reduces its external tariff against the nonmember, ceteris paribus, it reduces the imports from the partner and worsens the welfare of the partner country. Because a CU sets a CET, it can internalize this externality. Because of this effect, which is commonly known as the *externality-internalization effect*, the optimal common external tariff of the North-South CU becomes higher than the optimal external tariffs of the North-South FTA.

The magnitude of the tariff complementarity effect is also associated with Firm A's technology level. If Firm A uses the old technology, an increase in Country B's external tariff enhances the imports from Country A by less than if firm A adopts the new technology. This implies that the tariff-complementarity

[13] Lemma **??** also means that both the FTA and the CU meet the requirement of Article XXIV of the GATT/WTO, which forbids member countries from raising their external tariffs above the pre-PTA level.

[14] See Section 2 for the literature.

effect for Country $B$is less significant than that for Country A. Therefore, the optimal external tariff of Country $B$is higher than that of Country A if $\theta_A = \overline{\theta}$.

We have shown that, given Firm A's technology choice, the formation of PTAs leads to each member's unilateral tariff reduction toward the nonmember country. However, because the trades between members are liberalized to a greater degree, it is uncertain whether the external tariff reductions are sufficient to increase the imports from the nonmember country. We have the following proposition.

Proposition 1 *Given θ_A, the formation of the North-South FTA increases each member's imports from the nonmember country, while the formation of the North-South CU decreases them.*

Because of the tariff-complementarity effect, the North-South FTA becomes *trade-creating* in that it increases not only the trade between members but also the members' imports from the nonmember. In contrast, the North-South CU is *trade-diverting* in that it decreases the members' imports from the nonmember. This is because the externality-internalization effect limits the degree of external tariff reduction. As shown in the next section, these different effects on external trade between the North-South FTA and the North-South CU play an important role in determining the timing of trade agreements and of technology adoption.

4 The timing of trade agreements and technology adoption

In this section, we investigate how countries choose the timing of trade agreements and how Firm A chooses the timing of technology adoption on each path. As shown below, there is a case where countries can reach a trade agreement only after Firm A adopts the new technology. In this case, Firm A determines the timing of technology adoption with the expectation that this adoption will change not only its marginal cost but also the post-adoption trade regime.

To identify the effects of PTAs on the timing of technology adoption and of attaining multilateral free trade, three paths toward multilateral free trade are considered: (i) **MTA path:** countries negotiate an MTA if the current trade regime involves no agreements; (ii) **FTA path**: Countries A and B negotiate a North-South FTA under the case of no agreements and then the three countries negotiate an MTA among three countries after the FTA is formed; (iii) **CU path:** Countries A and B negotiate a North-South CU under the case of no agreements, and then the CU and Country C negotiate an MTA after the CU is formed. Let R_t ($R_t \in \Gamma$) denote the trade regime in period t. Then, the discounted sum of each country's welfare is given by

$$V_i(T) = \int_0^T e^{-rt} W_i^{R_t}(\overline{\theta}) dt + \int_T^{\infty} e^{-rt} W_i^{R_t}(\underline{\theta})\, dt - \alpha_i e^{-rT} k(T) \tag{1}$$

where α_i denote the parameter that takes $\alpha_A = 1$ and $\alpha_B = \alpha_C = 0$.

In the following analysis, we assume that $\lim_{t\to+\infty} k(t) = 0$; thus, the cost of technology adoption approaches zero over time. Under this assumption, Firm A eventually adopts the new technology and multilateral free trade is always realized at some point in time, irrespective of the path countries follow.[15] It is also assumed that countries cannot form an MTA simultaneously with the formation of a North-South PTA.

4.1 The MTA path

We first investigate the MTA path. Let t_M denote the timing under which the countries form the MTA. The trade regimes along the MTA path are given by $R_t = N$ for $t \in [0, t_M)$ and $R_t = M$ for $t \in [t_M, +\infty)$. By substituting them into Eq. (**??**), we obtain the discounted sum of each country's welfare along the MTA path, which is denoted by $V_i^{MTA}(t_M, T)$.

At time t, country i supports the MTA only if it increases $V_i^{MTA}(t_M, T)$ compared with the value attained by rejecting the MTA. Because firm A commits T at $t = 0$, each country takes T as a given in deciding whether to support the MTA. Let $\widetilde{t}_M^i$ (≥ 0) denote the unilaterally optimal timing of the MTA formation, which is obtained by $\widetilde{t}_M^i = \arg\max_{t_M} V_i^{MTA}(t_M, T)$. Country i opposes the MTA for $t < \widetilde{t}_M^i$ and supports it for $t \geq \widetilde{t}_M^i$. The conclusion of the MTA requires the unilateral support of all countries. Hence, the equilibrium timing of the MTA, $\widetilde{t}_M$, coincides with the maximum $\widetilde{t}_M^i$: $\widetilde{t}_M = \max[\widetilde{t}_M^A, \widetilde{t}_M^B, \widetilde{t}_M^C]$.

By differentiating $V_i^{MTA}(t_M, T)$ with respect to t_M, we obtain:

$$\frac{\partial V_i^{MTA}(t_M, T)}{\partial t_M} = \begin{cases} -e^{-rt_M}\left[W_i^M(\overline{\theta}) - W_i^N(\overline{\theta})\right] & \text{if} \quad t_M \leq T \\ -e^{-rt_M}\left[W_i^M(\underline{\theta}) - W_i^N(\underline{\theta})\right] & \text{if} \quad t_M > T \end{cases}. \tag{2}$$

The equation represents the marginal change of country i's welfare if the timing of the MTA formation is delayed. The sign of Eq.(**??**) depends on how the MTA changes the instantaneous welfare of each country. Given θ_A, if the MTA worsens the instantaneous welfare of Country i (i.e., $W_i^M(\theta_A) - W_i^N(\theta_A) < 0$), then we have $\partial V_i^{MTA}(t_M, T)/\partial t_M > 0$, thus delaying the timing of the MTA benefits country i. Alternatively, if the MTA improves the instantaneous welfare of country i, then we have $\partial V_i^{MTA}(t_M, T)/\partial t_M < 0$, thus delaying the timing of the MTA harms Country i.

We can verify that the MTA always increases the instantaneous welfare of developed countries, irrespective of the timing of Firm A's technology adoption. Therefore, Countries B and C prefer to form the MTA as soon as possible and their unilaterally optimal timings of the MTA become $\widetilde{t}_M^B = \widetilde{t}_M^C = 0$. In contrast, the MTA may not benefit Country A before Firm A's technology adoption. Specifically, if the gap between $\overline{\theta}$ and $\underline{\theta}$ is sufficiently large, then the MTA before the technology adoption worsens the instantaneous welfare of Country A. After Firm A adopts the new technology, however, the MTA necessarily improves Country A's continuous welfare. In this case, Country A opposes the MTA until Firm A adopts

[15] We relax this assumption in Section **??**.

the new technology and the unilaterally optimal timing of the MTA for country A becomes $\widetilde{t}_M^A = T$. We therefore have the following proposition.

Lemma 2 *On the MTA path, there exists a unique threshold value, λ_N^M $(\in (\underline{\theta}, \overline{\lambda}))$, such that an MTA is formed at $t = 0$ if $\overline{\theta} < \lambda_N^M$ holds, and it is formed at $t = T$ otherwise.*

The intuition behind Lemma **??** is as follows. If Firm A has not adopted the new technology, the developing country would experience a relatively small increase in exports and a relatively large increase in imports as a result of reciprocal trade liberalization. Hence, if the cost disadvantage of Firm A is significant, then the MTA may decrease the instantaneous profit of Firm A, and the profit loss can outweigh the efficiency gains from trade liberalization. After Firm A adopts the new technology and countries become symmetric, all countries gain equally from trade liberalization. Therefore, if $\overline{\theta}$ is sufficiently large, Country A chooses to delay the conclusion of the MTA until Firm A adopts the new technology.

Having described countries' incentives to form the MTA, we now turn to the determination of the timing of technology adoption. By Lemma **??**, if the unit cost with the old technology satisfies $\overline{\theta} < \lambda_N^M$, Firm A anticipates that multilateral free trade is immediately realized at $t = 0$, independent of its technology adoption. The discounted sum of Firm A's profits, net of the cost of the technology adoption, is given by:

$$\Pi_A (T) = \int_0^T e^{-rt} \pi_A^M \left(\overline{\theta}\right) dt + \int_T^\infty e^{-rt} \pi_A^M \left(\underline{\theta}\right) dt - e^{-rT} k (T) . \tag{3}$$

If $\overline{\theta} \geq \lambda_N^M$ holds, however, the countries form the MTA at $t = T$. Then, Firm A decides T with the anticipation that its adoption of the new technology at time T triggers the MTA formation. The discounted sum of Firm A's profits is given by:

$$\Pi_A (T) = \int_0^T e^{-rt} \pi_A^N \left(\overline{\theta}\right) dt + \int_T^\infty e^{-rt} \pi_A^M \left(\underline{\theta}\right) dt - e^{-rT} k (T) . \tag{4}$$

At the beginning of $t = 0$, Firm A chooses the timing of technology adoption, T_m, that maximizes $\Pi_A (T)$. The first-order condition is given by:

$$rk(T_m) - k'(T_m) = \begin{cases} \pi_A^M \left(\underline{\theta}\right) - \pi_A^M (\overline{\theta}) & \text{if} \quad \overline{\theta} < \lambda_N^M \\ \pi_A^M \left(\underline{\theta}\right) - \pi_A^N (\overline{\theta}) & \text{if} \quad \lambda_N^M \leq \overline{\theta} \end{cases} . \tag{5}$$

As $k' (t) < 0$ and $k'' (t) \geq 0$ hold, the second-order condition is satisfied. The above condition can be interpreted as follows. By postponing the technology adoption until the next period, Firm A is able to save $rk(T)$ as the interest rate and to gain from the decline in the adoption cost by $-k'(T)$. Hence, the left-hand side of Eq. (**??**) represents the marginal opportunity cost of adopting the new technology in the current period. However, adopting the new technology also raises Firm A's instantaneous profit in the current period. Hence, the right-hand side represents the marginal gains from technology adoption. The condition requires that the optimal timing of technology adoption must make the marginal gain equivalent to the marginal cost.

4.2 The FTA path

Next, we turn to the FTA path, where Countries A and B first negotiate the North-South FTA and then the three countries negotiate an MTA after the FTA is formed. Let t_P and t_M $(> t_P)$ respectively denote the timing of the formation of the North-South FTA and of the subsequent MTA formation. The trade regimes along the FTA path are given by $R_t = N$ for $t \in [0, t_P)$, $R_t = F$ for $t \in [t_P, t_M)$, and $R_t = M$ for $t \in [t_M, +\infty)$. By substituting them into Eq. (**??**), we obtain the discounted sum of each country's welfare along the FTA path, which is denoted by $V_i^{FTA}(t_P, t_M, T)$.

Under the case of no agreements in the previous period, Countries A and B form the FTA in the current period if it increases both $V_A^{FTA}(t_P, t_M, T)$ and $V_B^{FTA}(t_P, t_M, T)$ compared with the values attained by retaining the status quo. Let $\widehat{t}_P^h$ $(h \in \Omega \backslash C)$ denote the unilaterally optimal timing of the North-South FTA that maximizes $V_h^{FTA}(t_P, t_M, T)$. Because the formation of the FTA requires the unilateral support of both countries, the equilibrium timing of the North-South FTA is given by $\widehat{t}_P = \max[\widehat{t}_P^A, \widehat{t}_P^B]$. Alternatively, after the FTA is formed, Country i $(i \in \Omega)$ supports the subsequent MTA if it increases $V_i^{FTA}(t_P, t_M, T)$. Let $\widehat{t}_M^i$ denote the unilaterally optimal timing of the MTA for country i. The equilibrium timing of the MTA is given by $\widehat{t}_M = \max[\widehat{t}_P + \varepsilon, \widehat{t}_M^A, \widehat{t}_M^B, \widehat{t}_M^C]$, where ε (> 0) is an infinitesimal length of time.

By differentiating $V_i^{FTA}(t_P, t_M, T)$ with respect to t_P, we obtain:

$$\frac{\partial V_i^{FTA}(t_P, t_M, T)}{\partial t_P} = \begin{cases} -e^{-rt_P}\left[W_i^F(\overline{\theta}) - W_i^N(\overline{\theta})\right] & \text{if} \quad t_P < T \\ -e^{-rt_P}\left[W_i^F(\underline{\theta}) - W_i^N(\underline{\theta})\right] & \text{if} \quad T \leq t_P \end{cases}. \tag{6}$$

Similarly, by differentiating $V_i^{FTA}(t_P, t_M, T)$ with respect to t_M, we have:

$$\frac{\partial V_i^{FTA}(t_P, t_M, T)}{\partial t_M} = \begin{cases} -e^{-rt_M}\left[W_i^M(\overline{\theta}) - W_i^F(\overline{\theta})\right] & \text{if} \quad t_M < T \\ -e^{-rt_M}\left[W_i^M(\underline{\theta}) - W_i^F(\underline{\theta})\right] & \text{if} \quad T \leq t_M \end{cases}. \tag{7}$$

These two equations suggest that the unilaterally optimal timing of the FTA formation and of the subsequent MTA depend on whether those agreements improve or worsen the instantaneous welfare of the negotiating countries. As with the MTA path, whether Country A prefers to form trade agreements immediately or postpone them until Firm A decides to adopt the new technology depends on the level of $\overline{\theta}$. In addition, the developed countries always seek to form trade agreements as soon as possible. The following proposition is obtained.

Lemma 3 *On the FTA path, there exist unique cutoff values, λ_N^F $(\in (\underline{\theta}, \overline{\lambda}))$ and λ_F^M $(\in (\underline{\theta}, \lambda_N^F))$, such that (i) the North-South FTA is formed at $t = 0$ and then the MTA is formed at $t = T$ if $\lambda_F^M \leq \overline{\theta} < \lambda_N^F$ holds, (ii) the North-South FTA is formed at $t = 0$ and then the MTA is formed at $t = \varepsilon$ if $\overline{\theta} < \lambda_F^M$ holds, and (iii) the North-South FTA is formed at $t = T$ and then the MTA is formed at $t = T + \varepsilon$ if $\lambda_N^F \leq \overline{\theta}$ holds.*

As the MTA following the North-South FTA eliminates the member countries' gains from preferential treatment within the FTA, each member's instantaneous welfare gains from the FTA-formation exceed those from the subsequent MTA, $W_i^F(\theta_A) - W_i^N(\theta_A) > W_i^M(\theta_A) - W_i^F(\theta_A)$. This inequality in turn means that the cutoff value λ_N^F, above which the FTA worsens the instantaneous welfare of the developing country, is larger than the cutoff value λ_F^M, above which the subsequent MTA worsens it. As a result, there exists a case in which Country A always supports the North-South FTA but rejects the subsequent MTA when Firm A employs the old technology. More specifically, when $\lambda_F^M \leq \overline{\theta} < \lambda_N^F$ holds, the North-South FTA is immediately formed at $t = 0$, but the subsequent MTA is rejected until Firm A adopts the new technology.

Now, we investigate the timing of technology adoption. If $\lambda_F^M \leq \overline{\theta} < \lambda_N^F$ holds, Firm A anticipates that a North-South FTA is immediately formed at $t = 0$, but the MTA formation is contingent on its technology adoption. The discounted sum of Firm A's profits, net of the cost of technology adoption, is given by:

$$\Pi_A(T) = \int_0^T e^{-rt}\pi_A^F\left(\overline{\theta}\right)dt + \int_T^\infty e^{-rt}\pi_A^M\left(\underline{\theta}\right)dt - e^{-rT}k(T). \tag{8}$$

If $\overline{\theta} < \lambda_N^F$ holds, then Firm A anticipates that the timing of trade agreements will be independent of its technology adoption. Namely, the North-South FTA is formed at $t = 0$ and is immediately followed by the MTA, formed at $t = \varepsilon$. Because we have assumed that $k(0) = \overline{k}$ is sufficiently high, the cost of technology adoption at $t = \varepsilon$ is sufficiently high that technology adoption at $t \leq \varepsilon$ is unprofitable for Firm A. Hence, $T > \varepsilon$ holds, and the discounted sum of Firm A's profits is given by:

$$\Pi_A(T) = \int_0^\varepsilon e^{-rt}\pi_A^F\left(\overline{\theta}\right)dt + \int_\varepsilon^T e^{-rt}\pi_A^M\left(\overline{\theta}\right)dt + \int_T^\infty e^{-rt}\pi_A^M\left(\underline{\theta}\right)dt - e^{-rT}k(T). \tag{9}$$

Finally, if $\lambda_N^F \leq \overline{\theta}$ holds, Firm A anticipates that the North-South FTA will not be formed unless it adopts the new technology. Once Firm A adopts the new technology, the FTA is formed, which is immediately followed by the subsequent MTA. The discounted sum of Firm A's profits is given by:

$$\Pi_A(T) = \int_0^T e^{-rt}\pi_A^N\left(\overline{\theta}\right)dt + \int_T^{T+\varepsilon} e^{-rt}\pi_A^F\left(\underline{\theta}\right)dt + \int_{T+\varepsilon}^\infty e^{-rt}\pi_A^M\left(\underline{\theta}\right)dt - e^{-rT}k(T). \tag{10}$$

By solving $d\Pi_A(T)/dT = 0$, the optimal timing of technology adoption, T_f, satisfies:

$$rk(T_f) - k'(T_f) = \begin{cases} \pi_A^M\left(\underline{\theta}\right) - \pi_A^M(\overline{\theta}) & \text{if} \quad \overline{\theta} < \lambda_F^M \\ \pi_A^M\left(\underline{\theta}\right) - \pi_A^F(\overline{\theta}) & \text{if} \quad \lambda_F^M \leq \overline{\theta} < \lambda_N^F \\ \pi_A^F\left(\underline{\theta}\right) - \pi_A^N\left(\overline{\theta}\right) + e^{-r\varepsilon}\left[\pi_A^M\left(\underline{\theta}\right) - \pi_A^F\left(\underline{\theta}\right)\right] & \text{if} \quad \lambda_N^F \leq \overline{\theta} \end{cases}. \tag{11}$$

Note that with $\lambda_N^F \leq \overline{\theta}$, manipulating the timing of technology adoption changes the timing of both the North-South FTA and the subsequent MTA. Hence, the marginal gain from technology adoption becomes the sum of the increases in the instantaneous profit from the North-South FTA in the current period and from the subsequent MTA in the next period.

4.3 The CU path

The analytical structure of the CU path is the same as that of the FTA path, where t_P now denotes the timing of the formation of the North-South CU. The trade regimes along the path are given by $R_t = N$ for $t \in [0, t_P)$, $R_t = U$ for $t \in [t_P, t_M)$, and $R_t = M$ for $t \in [t_M, +\infty)$, and the discounted sum of each country's welfare along the CU path is given by $V_i^{CU}(t_P, t_M, T)$.

The optimal timing of the CU formation and that of the subsequent MTA are derived by replacing $W_i^F(\theta_A)$ with $W_i^U(\theta_A)$ in Eqs. (**??**) and (**??**). As with the MTA path and the FTA path, we can confirm that the developing country's decisions determine the fate of trade agreements and that the country's support of trade agreements depends on the level of $\overline{\theta}$. We have the following lemma.

Lemma 4 *On the CU path, there exist unique cutoff values,* λ_N^U $(\in (\underline{\theta}, \overline{\lambda}))$ *and* λ_U^M $(\in (\underline{\theta}, \lambda_N^U))$*, such that (i) the North-South CU is formed at* $t = 0$ *and then the MTA is formed at* $t = T$ *if* $\lambda_U^M \leq \overline{\theta} < \lambda_N^U$ *holds, (ii) the North-South CU is formed at* $t = 0$ *and then the MTA is formed at* $t = \varepsilon$ *if* $\overline{\theta} < \lambda_U^M$ *holds, and (iii) the North-South CU is formed at* $t = T$ *and then the MTA is formed at* $t = T + \varepsilon$ *if* $\lambda_N^U \leq \overline{\theta}$ *holds.*

Analogous to the FTA path, the optimal timing of technology adoption, T_u, satisfies:

$$rk(T_u) - k'(T_u) = \begin{cases} \pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta}) & \text{if} \quad \overline{\theta} < \lambda_U^M \\ \pi_A^M(\underline{\theta}) - \pi_A^U(\overline{\theta}) & \text{if} \quad \lambda_U^M \leq \overline{\theta} < \lambda_N^U \\ \pi_A^U(\underline{\theta}) - \pi_A^N(\overline{\theta}) + e^{-r\varepsilon}\left[\pi_A^M(\underline{\theta}) - \pi_A^U(\underline{\theta})\right] & \text{if} \quad \lambda_N^U \leq \overline{\theta} \end{cases}. \tag{12}$$

5 Comparisons

We have investigated each country's decisions to form trade agreements and Firm A's private incentive to adopt the new technology. In this section, we compare the equilibrium outcomes of three paths to identify which path realizes technology adoption and multilateral free trade most quickly. We also discuss which path realizes the highest welfare for each country, as well as which path realizes the highest world welfare. It is shown that the path that maximizes an individual country's welfare may be different from the path that maximizes world welfare.

First, we order the critical values of $\overline{\theta}$ that change the equilibrium timing of technology adoption and of trade agreements.

Lemma 5 *The cutoff values of* $\overline{\theta}$ *satisfy* $\lambda_F^M < \lambda_U^M < \lambda_N^M < \lambda_N^U < \lambda_N^F$.

By this lemma, $\lambda_N^M \leq \overline{\theta}$ may hold even if $\overline{\theta} < \lambda_N^F$ or $\overline{\theta} < \lambda_N^U$ holds, which implies that the developing country may support the North-South FTA or the North-South CU, even when it opposes the MTA formation. In contrast to the one-shot elimination of all tariffs on the MTA path, the formation of the FTA as well as that of the CU only eliminates tariffs between the member countries. As a result, when trade agreements reduce Firm A's profits, the MTA on the MTA path reduces them more than the FTA-formation or the CU-formation. The effect increases the feasibility of the FTA-formation as well as that of the CU-formation compared with the MTA on the MTA path.

By contrast, the lemma suggests that $\overline{\theta} < \lambda_N^M$ holds whenever $\overline{\theta} < \lambda_F^M$ or $\overline{\theta} < \lambda_U^M$ holds. Therefore, given that Firm A uses the old technology, the MTA on the MTA path is always feasible whenever the subsequent MTA on the FTA path or on the CU path is feasible. Preferential trade liberalization by the North-South PTA gives Firm A some gains from the preferential treatment within the PTA. As the subsequent MTA that follows the FTA or CU eliminates such gains from the preferential treatment, Firm A's loss from the subsequent MTA is greater than the loss from the one-shot MTA on the MTA path.

Comparing the North-South FTA and the North-South CU, the developing country may support the FTA-formation even if it opposes the CU-formation because $\lambda_N^U \leq \overline{\theta} < \lambda_N^F$ may hold. The FTA may be preferable because it allows member countries to set external tariffs that maximize their individual welfare, whereas the CET set by the CU diverges from their individually optimal external tariffs. However, the CET has the externality-internalization effect, which improves the welfare of both member countries. When $\overline{\theta}$ is high enough that the PTA-formation becomes infeasible, the former effect dominates the latter and we have $\lambda_N^U < \lambda_N^F$. This inequality in turn means that $\lambda_F^M < \lambda_U^M$ holds as the MTA following the FTA has more difficulty in gaining the support of the developing country than the MTA that follows the CU.

5.1 The timing of technology adoption and the realization of free trade

In this subsection, we compare the equilibrium timings of technology adoption and the equilibrium timings of attaining multilateral free trade. By Lemma **??**, we have six cases, depending on the level of $\overline{\theta}$. Figure 1 depicts how the equilibrium timings of trade agreements change as $\overline{\theta}$ increases on each path.

[Insert Figure 1 around here]

Let us first compare the FTA path with the MTA path. When $\lambda_N^M \leq \overline{\theta} < \lambda_N^F$ holds, $\overline{\theta}$ is high enough to undermine Country A's support of the MTAs on both paths but low enough to maintain its support of the North-South FTA. As the MTA is formed only after Firm adopts the new technology, the timing of the MTA formation coincides with the timing of technology adoption on each path. In contrast, the North-South FTA is immediately formed at $t = 0$ on the FTA path.

By (**??**) and (**??**), the timing of the technology adoption on the MTA path and on the FTA path are respectively determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ and $rk(T_f) - k'(T_f) = \pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})$. Due to its trade-creating nature, $\pi_A^N(\overline{\theta}) > \pi_A^F(\overline{\theta})$ holds because the formation of the North-South FTA reduces Firm A's pre-adoption profit with the old technology under this range of $\overline{\theta}$. This effect makes the technology adoption more attractive for Firm A. Firm A's post-adoption profit is given by $\pi_A^M(\underline{\theta})$ on both paths because the technology adoption immediately leads to the realization of multilateral free trade on both paths, which, in turn, means that the marginal gains from technology adoption are higher on the FTA path than on the MTA path: $\pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta}) > \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$. Consequently, we have $T_f < T_m$ if $\lambda_N^M \leq \overline{\theta} < \lambda_N^F$ holds. In this case, compared with the one-shot tariff eliminations on the MTA path, the sequential tariff eliminations on the FTA path accelerate the technology adoption in the South and the realization of multilateral free trade.

When $\overline{\theta} < \lambda_N^M$ or $\lambda_N^F \leq \overline{\theta}$ holds, the MTA on the MTA path is feasible whenever the North-South FTA is feasible. Hence, the FTA formation on the FTA path cannot precede the conclusion of the MTA on the MTA path. Under this situation, the FTA path is not the quickest method to attain multilateral free trade. For instance, suppose $\lambda_F^M \leq \overline{\theta} < \lambda_N^M$ holds. In this case, the MTA is immediately formed at $t = 0$ on the MTA path whereas the conclusion of the MTA on the FTA path is delayed until Firm A adopts the new technology. The optimal timing of technology adoption on the MTA path is now determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})$, whereas it is determined by $rk(T_f) - k'(T_f) = \pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})$ on the FTA path. As $\pi_A^F(\overline{\theta}) > \pi_A^M(\overline{\theta})$ always holds in this range of $\overline{\theta}$, we have $\pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta}) > \pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})$ and thereby $T_m < T_f$ holds in equilibrium. Intuitively, because Firm A always faces multilateral free trade on the MTA path but only faces preferential trade liberalization under the FTA path until adopting the new technology, Firm A's profit before technology adoption on the MTA path is lower than its profit on the FTA path. The smaller profit before technology adoption means that Firm A's incentive to increase its profit to $\pi_A^M(\underline{\theta})$ by adopting the new technology is higher on the MTA path. If $\overline{\theta} < \lambda_F^M$ holds, multilateral free trade is immediately realized on all three paths. In this case, the three paths lead to the same timing of technology adoption. However, the MTA path realizes multilateral free trade faster because it reaches the technology adoption in one step, whereas the FTA path needs a two-step agreement. Finally, if $\overline{\theta}$ is sufficiently high to satisfy $\lambda_N^F \leq \overline{\theta}$, the trade regime stays at no agreements on all paths until Firm A adopts the new technology. As the FTA path requires a two-step agreement, it delays both the timing of technology adoption and the realization of multilateral free trade. The following proposition summarizes the comparison between the MTA path and the FTA path.

Proposition 2 *If the FTA is formed before technology adoption and an MTA is formed after technology adoption on both the MTA path and the FTA path, then the timing of technology adoption and the timing of attaining multilateral free trade are faster on the FTA path than on the MTA path. Otherwise, they cannot be faster on the FTA path.*

This proposition suggests that when a multilateral trade negotiation is undermined by the opposition of a developing country, an FTA that includes the developing country can foster multilateral trade liberalization because it accelerates the country's technological advancement.

By contrast, even if the North-South CU is formed before technology adoption, it never accelerates the timing of technology adoption and delays the realization of multilateral free trade. Suppose $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$ holds, under which the North-South CU is formed at $t = 0$, and the MTA formation needs Firm A's technology adoption on both the MTA path and the CU path. The optimal timing of technology adoption on the MTA path and on the CU path are respectively determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ and $rk(T_u) - k'(T_u) = \pi_A^M(\underline{\theta}) - \pi_A^U(\overline{\theta})$. Because of the trade-diverting nature of CU, Firm A earns profit under the North-South CU that is higher than the profit it earns under the case of no agreements (i.e., $\pi_A^U(\overline{\theta}) > \pi_A^N(\overline{\theta})$) in this range of $\overline{\theta}$. The higher profit before technology adoption on the CU path means that $T_m < T_u$ holds. Hence, even if the realization of multilateral free trade is contingent on technology adoption, the formation of a CU cannot accelerate the adoption. Rather, it delays both the realization of multilateral free trade and the timing of technology adoption.

Figure 2 compares the equilibrium timing of technology adoption when $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$ holds. In this range of $\overline{\theta}$, the realization of multilateral free trade is contingent on technology adoption on all three paths. The marginal opportunity cost of technology adoption, $rk(t) - k'(t)$, is downward sloping because we have assumed that $k'(t) < 0$, $k''(t) \geq 0$. The marginal gains of technology adoption on each path do not depend on t and they are drawn as a horizontal line. The timing of technology adoption is determined at the intersection of these two curves.[16]

[Insert Figure 2 around here]

Because $\pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta}) > \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta}) > \pi_A^M(\underline{\theta}) - \pi_A^U(\overline{\theta})$ holds in this case, the order of the equilibrium timing of technology adoption becomes $T_f < T_m < T_u$. Because the timing of technology adoption coincides with the timing of the MTA formation on all three paths, the speed of attaining multilateral free trade is fastest on the FTA path, followed by the MTA path and then the CU path.

We have shown that the formation of the North-South CU before technology adoption cannot accelerate technology adoption and the realization of multilateral free trade if the technology gap is sufficiently large that the CU is formed *after* technology adoption, then the CU path accelerates technology adoption and may also accelerate the realization of multilateral free trade. Suppose $\lambda_N^U \leq \overline{\theta}$ holds, in which case not only does the MTA formation require Firm A's technology adoption on the MTA path, but also the CU formation requires the technology adoption on the CU path. The optimal timing of technology adoption on the MTA path and on the CU path is respectively determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ and $rk(T_u) - k'(T_u) = \pi_A^U(\underline{\theta}) - \pi_A^N\left(\overline{\theta}\right) + e^{-r\varepsilon}\left[\pi_A^M(\underline{\theta}) - \pi_A^U(\underline{\theta})\right]$. We can confirm that $\pi_A^U(\underline{\theta}) > \pi_A^M(\underline{\theta})$ holds. Due to the trade-diverting nature of the CU, Firm A's post-adoption profit is larger under the case of the North-South CU than under the case of multilateral free trade. Because $\{rk(T_m) - k'(T_m)\} -$

[16]Because we have assumed that $k(0) = \overline{k}$ is large enough and that $\lim_{t\to+\infty} k(t) = 0$ holds, a unique interior solution always exists. We relax this assumption in the next section.

$\{rk(T_u) - k'(T_u)\} = (1 - e^{-r\varepsilon}) \left[\pi_A^M(\underline{\theta}) - \pi_A^U(\underline{\theta})\right] < 0$, $T_u < T_m$ holds. Intuitively, compared to the one-shot multilateral liberalization on the MTA path, Firm A can earn a higher interim profit from the temporary formation of the CU on the CU path. This higher profit increases firm A's gains from adopting the new technology. Because multilateral free trade is realized at $t = T_m$ on the MTA path and at $t = T_u + \varepsilon$ on the CU path, the countries may reach multilateral free trade faster on the CU path if $T_m - T_u > \varepsilon$ holds.

Proposition 3 *I If the CU is not formed before technology adoption, then the timing of technology adoption is faster and the timing of attaining multilateral free trade can be faster on the CU path than on the MTA path. Otherwise, they cannot be faster on the CU path.*

The comparisons of the timing of technology adoption and the timing of attaining multilateral free trade are summarized in Table 1.

[**Insert Table 1 around here**]

5.2 World welfare

We have shown that the formation of PTAs can expedite the timing of realizing multilateral free trade if it accelerates the timing of technology adoption. Therefore, an important question is whether faster technology adoption and earlier realizations of multilateral free trade are desirable from the viewpoint of world welfare. Let $WW^{R_t}(\theta_A) = \sum_{i \in \Omega} W_i^{R_t}(\theta_A)$ denote the instantaneous world welfare in regime R_t ($R_t \in \Gamma$), given θ_A ($\in \{\underline{\theta}, \overline{\theta}\}$). The discounted sum of world welfare is calculated by the discounted sum of $WW^{R_t}(\theta_A)$ for $t \in [0, +\infty)$ subtracted by the present value of the fixed cost of technology adoption, $e^{-rT}k(T)$. We have the following lemma.

Lemma 6 *(i) Given θ_A, $WW^M(\theta_A) > WW^F(\theta_A) > WW^U(\theta_A) > WW^N(\theta_A)$ holds. (ii) Given R_t, $WW^{R_t}(\underline{\theta}) > WW^{R_t}(\overline{\theta})$ holds.*

The lemma implies that any trade agreement improves instantaneous world welfare. Even if a trade agreement harms the developing country, its loss is outweighed by the sum of the developed countries' gains. As the equilibrium external tariffs of the North-South FTA are lower than the CET of the North-South CU, world welfare under the FTA is higher than world welfare under the CU.

The lemma also suggests that Firm A's technology adoption increases the instantaneous world welfare given the trade regime. This is because both the cost-reduction and the enhanced product market competition generate efficiency gains. As technology adoption may trigger and never deters, trade agreements, faster technology adoption increases the discounted sum of $WW^{R_t}(\theta_A)$. In addition, faster technology adoption increases the fixed cost of technology adoption. In general, whether faster technology adoption increases the discounted sum of world welfare depends on the relative magnitudes of these two effects.

By comparing the discounted sum of world welfare along the FTA path or the CU path with that along the MTA path, we obtain the following proposition.

Proposition 4 *If the timing of attaining multilateral free trade on the MTA path is the fastest among the three paths, then the MTA path generates the highest discounted sum of world welfare. If the FTA path or the CU path reaches multilateral free trade faster than the MTA path, then it also generates a higher discounted sum of world welfare.*

Thus, sequential trade liberalization through the FTA path or through the CU path is beneficial if it accelerates the realization of multilateral free trade, but it is harmful if neither path can expedite multilateral free trade.

Regarding the comparison between the FTA path and the CU path, we have the following proposition.

Proposition 5 *If the timing of technology adoption on the FTA path is faster than or the same as that on the CU path, then the FTA path generates the higher discounted sum of world welfare.*

Table 1 suggests that if the timing of technology adoption is faster on the FTA path than on the CU path, the timing of attaining multilateral free trade is also faster on the FTA path. Because Lemma **??** suggests that the interim world welfare before technology adoption is higher on the FTA path, the FTA path generates a higher discounted sum of world welfare if it accelerates the adoption of new technology in the developing country. In contrast, even if the CU path accelerates technology adoption faster than the FTA path, it may not realize a higher discounted sum of world welfare because the interim world welfare before technology adoption can be lower on the CU path.

5.3 The welfare of each country

In the above subsection, we have specified the path that maximizes the discounted sum of world welfare. However, the result does not automatically mean the same path is also preferable to individual countries. For instance, consider the case with $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$, in which the timing of technology adoption and of concluding the MTA is ordered by $T_f < T_m < T_u$. In this case, the FTA path attains the highest world welfare. Figure 3 illustrates the changes in each country's individual welfare given that $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$ holds.

[Insert Figure 3 around here]

As the figure shows, Country A generates the highest welfare, gross of the adoption cost, on the FTA path at every point in time. Although the late adoption of the new technology on the MTA path and on the CU path saves adoption costs, the gross welfare gain always outweighs the cost-saving effect. Hence, Country A always prefers the FTA path to other paths under $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$. With regard to Country B, which is Country A's partner in North-South PTAs, it is evident that the CU path generates higher welfare than the other paths at every point in time. As the technology adoption of the foreign rival firm harms the domestic firm, Country B prefers the CU path, which delays technology adoption and sustains the domestic firm's technological advantage over the foreign rival for longer periods.[17] Country C necessarily prefers the FTA path because it generates the highest welfare at every point in time, although it cannot affect the type of PTA that Countries A and B will form. The following proposition summarizes the result.

Proposition 6 *The developed country may prefer the CU path, even if the other paths realize multilateral free trade faster and generate higher world welfare.*

Suppose countries can choose the path they will follow at the beginning of $t = 0$. If Country B has strong bargaining power in the determination of the path, then it may select the CU path even if the other paths realize higher world welfare.

6 Further discussion

Given that multilateral free trade is eventually realized by countries, we have shown that the temporary formation of North-South PTAs may change the speed of technology adoption and of attaining multilateral free trade. This section reviews the result under a number of alternative setups.

6.1 North-North PTAs

Thus far, we have not considered North-North PTAs between Countries B and C. We can confirm that the result of Proposition **??** is maintained in the case of North-North PTAs. Namely, the formation of the FTA benefits the nonmember firm and improves the welfare of the nonmember country, whereas the formation of the CU decreases the firm's profit and worsens welfare in the nonmember country. However, the trade-creating nature of the FTA and the trade-diverting nature of the CU have different impacts on the timing of technology adoption, as well as those of the subsequent MTA.

For instance, consider the case where the MTA is contingent upon Firm A's technology adoption on all three paths. Firm A anticipates that its technology adoption will immediately lead to multilateral free trade. As the North-North FTA increases Firm A's profit, maintaining the status quo as the nonmember of the FTA becomes more attractive for Firm A. As a result, Firm A's marginal gains from the technology

[17] Even though $W_B^U(\overline{\theta}) > W_B^F(\overline{\theta}) > W_B^M(\underline{\theta})$ means that the subsequent MTA following the FTA or the CU reduces the welfare of Country C on the equilibrium path, Country B always signs the MTA because $W_B^M(\underline{\theta}) > W_B^U(\underline{\theta}) > W_B^F(\underline{\theta})$ means that the MTA always improves the welfare of Country B once the new technology is adopted by Firm A.

adoption decline under the North-North FTA. Furthermore, the trade-creating nature of the FTA reduces Country A's potential welfare gains from the subsequent MTA formation. Hence, an initially feasible MTA prior to technology adoption may become infeasible after the formation of the FTA.[18] Because of these effects, the timing of technology adoption and the realization of multilateral free trade on the FTA path are always delayed compared with the MTA path.

In contrast, the trade-diverting nature of the CU harms Firm A and thereby increases its incentive to adopt the new technology. Furthermore, the CU increases Country A's welfare gains from the subsequent MTA formation. As a result, there is a case where the timing of technology adoption and the realization of multilateral free trade are accelerated by the North-North CU. The next proposition summarizes the results.

Proposition 7 *Suppose a PTA is formed between developed countries. Compared with one-shot trade liberalization on the MTA path, sequential trade liberalization on the CU path may accelerate the technology adoption and the realization of multilateral free trade, whereas liberalization via the FTA path never accelerates the technology adoption and always delays the realization of multilateral free trade.*

This proposition suggests that the possible superiority of the FTA over the other paths with respect to the speed of attaining multilateral free trade, which exists in the case of North-South PTAs, disappears if the developing country becomes the nonmember. Table 2 summarizes all possible cases. In the table, μ_F^M and μ_U^M respectively represent the cutoff values under which the subsequent MTA following the North-North FTA and the North-North CU increases the instantaneous welfare of Country A.[19]

[Insert Table 2 around here]

6.2 The lower bound of the adoption cost

We have assumed that $\lim_{t\to+\infty} k(t) = 0$ holds, so that the cost of technology adoption eventually becomes zero. Under this assumption, the new technology is eventually adopted by Firm A, irrespective of the liberalization path followed by countries. However, owing to the limited availability of skilled labor or a lack of absorptive capacity in the developing country, the level of the adoption cost may have a lower bound. If this is true, the new technology may not be adopted and multilateral free trade may not be attained in a particular path of trade liberalization.

Let $\underline{k}$ denote the lower bound of the adoption cost. The marginal opportunity cost of the technology adoption approaches $r\underline{k}$ over time. If the marginal gains in profits from adopting the new technology are smaller than $r\underline{k}$ on a particular path, then Firm A never adopts the new technology on that path. For

[18]Yi (2000) and Ornelas (2005a) have pointed out this effect.

[19]See the proof of Proposition **??** for details.

instance, consider the case in which $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$ holds so that the technology adoption is a prerequisite for the endorsement of the MTA. If $\pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta}) > r\underline{k} > \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ holds, the new technology is adopted and multilateral free trade is realized only on the FTA path.

Alternatively, consider the case with $\lambda_U^M \leq \overline{\theta} < \lambda_N^M$. The MTA is immediately formed on the MTA path, whereas on the FTA and CU paths, it is formed only if the new technology is adopted. If $r\underline{k} > \pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})$ holds, multilateral free trade is realized only on the MTA path.

These results suggest that the formation of the North-South PTA can be a stepping-stone to multilateral free trade, in that it can make an initially infeasible MTA feasible. The formation of the PTA can also undermine multilateral free trade, in that it can make an initially feasible MTA infeasible. In Bhagwati (1991)'s words, PTAs can be either "building blocks" or "stumbling blocks" depending on the degree of the technology gap between developed and developing countries. In contrast to the existing studies that focus on political-economic factors for trade agreements, this result points out a role of technology diffusion in considering the relationship between regionalism and multilateralism.

6.3 Discriminatory tariffs under the case of no agreements

It has been assumed that countries impose MFN tariffs under the case of no agreements. The WTO has a generalized system of preference (GSP) that allows countries to impose lower tariffs on products originating in developing countries. If the system is applied to our model, Countries B and C are able to set lower tariffs on their imports from Country A under the case of no agreements. In fact, they are willing to do so because the optimal trade policy for each country is to impose a lower tariff on the product from the less cost-efficient country under international oligopoly.[20]

A remaining question is whether allowing discriminatory tariffs accelerates technology adoption and the realization of multilateral free trade. Because lower tariffs are applied to Firm A as long as it uses the old technology, $\pi_A^N(\overline{\theta})$, $\pi_A^F(\overline{\theta})$, and $\pi_A^U(\overline{\theta})$ become larger in the case of allowing GSP than in the case of the MFN tariff.[21] By Eqs. (**??**), (**??**), and (**??**), these increases in the profit reduce Firm A's gains from technology adoption and thereby delays the timing of technology adoption. If the MTA is contingent on technology adoption, the system also delays the realization of multilateral free trade. Therefore, although the GSP benefits developing countries in the short run, it may harm them if it substantially delays their technological advancement and the realization of multilateral free trade. The result points out a new impact of committing to the MFN principle that has been overlooked in the literature.[22]

[20] This is because a country can extract more rent from the lower-cost exporter by increasing a tariff. See Gatsios (1990) and Hwang and Mai (1991).

[21] See Appendix for details.

[22] Choi (1995) suggests that committing to the MFN principle is important to overcome a time-inconsistency problem of ex ante technology adoption and ex post trade policy and to promote the adoption of a lower marginal cost technology.

7 Conclusion

Is regionalism a faster path to multilateral free trade? Answering this question is important to evaluate the proliferation of PTAs under the sluggish WTO negotiations. To our knowledge, this is the first theoretical paper that formally analyzes the issue.

By developing an international oligopoly model with two developed countries and one developing country, we have investigated how PTAs change the timing of technology adoption and the speed of attaining multilateral free trade. If a PTA is formed between a developing country and a developed country prior to technology adoption, then a two-step liberalization via an FTA accelerates the technological advancement of the developing country and the realization of multilateral free trade compared with the one-shot multilateral agreement. This case emerges when the success of multilateral trade negotiations is contingent on the diffusion of the advanced technology to the developing country. In contrast, sequential trade liberalization via a CU delays the timing of both technology adoption and realizing multilateral free trade in the same situation.

Even if the faster attainment of multilateral free trade and the faster technology adoption both improve world welfare, the developed country may prefer the CU path to the FTA path or the MTA path because faster adoption of the new technology by the rival firm does not always benefit the developed country. We have also shown that the advantage of an FTA over a CU is reversed if a PTA cannot be formed prior to technology adoption or if the members of PTAs are developed countries.

Although the model is highly stylized, the paper provides new insight into the recent strand of literature that investigates the relationship between regionalism and multilateralism. There remains room for further research, and this paper should be seen as a first step toward deeper investigation of the issue. For example, including more than two developing countries will allow us to investigate technology adoption games between firms and possible South-South PTAs. Incorporating firm relocation via foreign direct investments will be another interesting extension.

Appendix

Proof of Lemma ??

Since markets are segmented, neither A-B FTA nor A-B CU changes the effects of Country C's tariffs on its instantaneous welfare. Hence, it is evident that $\widetilde{\tau}_C^N(\theta_A) = \widetilde{\tau}_C^F(\theta_A) = \widetilde{\tau}_C^U(\theta_A)$ holds. With regard to Countries A and B, we can calculate that $\widetilde{\tau}_A^N(\overline{\theta}) - \overline{\tau}^U(\overline{\theta}) = \widetilde{\tau}_B^N(\overline{\theta}) - \overline{\tau}^U(\overline{\theta}) = \left(7a - 29\overline{\theta} + 22\underline{\theta}\right)/190 > \left(7a - 29\overline{\lambda} + 22\underline{\theta}\right)/190 = 2\left(a - \underline{\theta}\right)/133 > 0$, $\overline{\tau}^U(\overline{\theta}) - \widetilde{\tau}_A^F(\overline{\theta}) = 8\{6\left(a - \overline{\theta}\right) + 11(\overline{\theta} - \underline{\theta})\}/399 > 0$, $\overline{\tau}^U(\overline{\theta}) - \widetilde{\tau}_B^F(\overline{\theta}) = 16\left(3a - 7\overline{\theta} + 4\underline{\theta}\right)/399 > 16\left(3a - 7\overline{\lambda} + 4\underline{\theta}\right)/399 = 32\left(a - \underline{\theta}\right)/399 > 0$, $\widetilde{\tau}_A^N(\underline{\theta}) - \overline{\tau}^U(\underline{\theta}) = \widetilde{\tau}_B^N(\underline{\theta}) -$

$\overline{\tau}^U(\underline{\theta}) = 7(a-\underline{\theta})/190 > 0$, and $\overline{\tau}^U(\underline{\theta}) - \widetilde{\tau}_A^F(\underline{\theta}) = \overline{\tau}^U(\underline{\theta}) - \widetilde{\tau}_B^F(\underline{\theta}) = 16(a-\underline{\theta})/133 > 0$ hold. These equalities prove that $\widetilde{\tau}_h^N(\theta_A) > \overline{\tau}^U(\theta_A) > \widetilde{\tau}_h^F(\theta_A)$ $(h \in \Omega \backslash C)$ holds. In addition, we have $\widetilde{\tau}_B^F(\theta_A) - \widetilde{\tau}_A^F(\theta_A) = 8(\theta_A - \underline{\theta})/21$. This equation means that $\widetilde{\tau}_B^F(\theta_A) \geq \widetilde{\tau}_A^F(\theta_A)$ holds, with equality if $\theta_A = \underline{\theta}$. ■

Proof of Propostion ??

Let $q_C^{j,N}(\theta_A)$, $q_C^{j,F}(\theta_A)$ and $q_C^{j,U}(\theta_A)$ denote the equilibrium sales of Firm C in Country j given θ_A under the case of no agreements, of North-South FTA, and of North-South CU, respectively. We have $q_C^{A,F}(\theta_A) - q_C^{A,N}(\theta_A) = \{3(a-\theta_A) + 2(\theta_A - \underline{\theta})\}/70 > 0$ and $q_C^{B,F}(\theta_A) - q_C^{B,N}(\theta_A) = 3(\overline{\lambda} - \overline{\theta})/10 > 0$. Besides that, we have $q_C^{A,U}(\theta_A) - q_C^{A,N}(\theta_A) = q_C^{B,U}(\theta_A) - q_C^{B,N}(\theta_A) = -\{9(a-\theta_A) + 26(\theta_A - \underline{\theta})\}/190 < 0$. Thus, North-South FTA increases but North-South CU decreases each member's import from the nonmember country. ■

Proof of Lemma ??

Because we have $W_j^M(\overline{\theta}) - W_j^N(\overline{\theta}) = \{13(a-\overline{\theta}) + 62(\overline{\theta} - \underline{\theta})\}\{3(a-\overline{\theta}) + 2(\overline{\theta} - \underline{\theta})\}/800 > 0$ $(j \in \Omega \backslash A)$ and $W_i^M(\underline{\theta}) - W_i^N(\underline{\theta}) = 39(a-\underline{\theta})^2/800 > 0$ $(i \in \Omega)$, $\partial V_j^{MTA}(t_M, T)/\partial t_M < 0$ $(j \in \Omega \backslash A)$ always holds. We also have $W_A^M(\overline{\theta}) - W_A^N(\overline{\theta}) = \left(3a - \overline{\theta} - 2\underline{\theta}\right)(13a - 111\overline{\theta} + 98\underline{\theta})$. If $\overline{\theta} \leq \lambda_N^M := (13a + 98\underline{\theta})/111$ is satisfied, then $W_A^M(\overline{\theta}) \geq W_A^N(\overline{\theta})$ and $\partial V_A^{MTA}(t_M, T)/\partial t_M \leq 0$ hold. Note that λ_N^M is smaller than the upper bound of $\overline{\theta}$, $\overline{\lambda}$. In this case, it is optimal for all countries to form the MTA at $t = 0$ on the MTA path, and $t_M = 0$ holds in equilibrium. If $\overline{\theta} < \lambda_N^M$ is satisfied, however, then $W_A^M(\overline{\theta}) < W_A^N(\overline{\theta})$ and $\partial V_A^{MTA}(t_M, T)/\partial t_M > 0$ hold for $0 \leq t_M \leq T$ and $\partial V_A^{MTA}(t_M, T)/\partial t_M < 0$ holds for $T < t_M$. In this case, it is optimal for Country A to form the MTA at $t = T$ while Countries B and C prefer to form the MTA as early as possible. As a result, the equilibrium timing of the MTA satisfies $\widetilde{t}_M = T$. ■

Proof of Lemma ??

Because $d\{W_B^F(\overline{\theta}) - W_B^N(\overline{\theta})\}/d\overline{\theta} = \{3882\left(a - \overline{\theta}\right) + 5353(\overline{\theta} - \underline{\theta})\}/22050 > 0$ and $W_B^F(\underline{\theta}) - W_B^N(\underline{\theta}) = 141(a-\underline{\theta})^2/4900 > 0$ hold, $W_B^F(\overline{\theta}) - W_B^N(\overline{\theta}) > 0$ always holds. Besides that, $d\{W_B^M(\overline{\theta}) - W_B^F(\overline{\theta})\}/d\overline{\theta} < 0$ and $W_B^M(\overline{\lambda}) - W_B^F(\overline{\lambda}) = 727(a-\underline{\theta})^2/43218 > 0$ mean that $W_B^M(\underline{\theta}) - W_B^F(\underline{\theta}) > W_B^M(\overline{\theta}) - W_B^F(\overline{\theta}) > 0$ holds. Furthermore, $d\{W_C^M(\overline{\theta}) - W_C^F(\overline{\theta})\}/d\theta_A = 3\{269\left(a - \overline{\theta}\right) + 366(\overline{\theta} - \underline{\theta})\}/3920 > 0$ and $W_C^M(\underline{\theta}) - W_C^F(\underline{\theta}) = 219(a-\underline{\theta})^2/7840$ mean that $W_C^M(\overline{\theta}) - W_C^F(\overline{\theta}) > 0$ holds. Thus, $\partial V_j^{FTA}(t_P, t_M, T)/\partial t_P < 0$ and $\partial V_j^{FTA}(t_P, t_M, T)/\partial t_M < 0$ $(j \in \Omega \backslash A)$ hold irrespective of the timing of technology adoption. This implies that Country B prefers to form North-South FTA as early as possible, and Countries B and C intend to form the subsequent MTA as early as possible.

We next turn to welfare changes in Country A. After Firm A adopts the new technology, both A-B FTA and the subsequent MTA improve the instantaneous welfare of Country A because $W_A^F(\underline{\theta}) - W_A^N(\underline{\theta}) =$

$141\left(a-\underline{\theta}\right)^2/4900 > 0$ and $W_A^M(\underline{\theta}) - W_A^F(\underline{\theta}) = 783\left(a-\underline{\theta}\right)^2/39200 > 0$ hold. Because $d\{W_A^F(\overline{\theta}) - W_A^N(\overline{\theta})\}/d\overline{\theta} = -\{669(a-\overline{\theta}) + 986(\overline{\theta}-\underline{\theta})\}/3150 < 0$, $W_A^F(\underline{\theta}) - W_A^N(\underline{\theta}) > 0$, and $W_A^F(\overline{\lambda}) - W_A^N(\overline{\lambda}) = -8\left(a-\underline{\theta}\right)^2/3087 < 0$ hold, there exists a unique cut-off value, λ_N^F ($\in \{\underline{\theta}, \overline{\lambda}\}$), such that $W_A^F(\overline{\theta}) < W_A^N(\overline{\theta})$ holds if $\overline{\theta} > \lambda_N^F$ and $W_A^F(\overline{\theta}) \geq W_A^N(\overline{\theta})$ holds otherwise. Similarly, $d\{W_A^M(\overline{\theta}) - W_A^F(\overline{\theta})\}/d\overline{\theta} = -\left(5547a - 9529\overline{\theta} + 3982\underline{\theta}\right)/25200 < -(5547a - 9529\overline{\lambda} + 3982\underline{\theta})/25200 = -293\left(a-\underline{\theta}\right)/1764 < 0$, $W_A^M(\underline{\theta}) - W_A^F(\underline{\theta}) > 0$, and $W_A^M(\overline{\lambda}) - W_A^F(\overline{\lambda}) = -47\left(a-\underline{\theta}\right)^2/6174 < 0$ hold. These inequalities mean that there exists a unique cut-off value, λ_F^M ($\in (\underline{\theta}, \overline{\lambda})$), such that $W_A^M(\overline{\theta}) < W_A^F(\overline{\theta})$ holds if $\overline{\theta} > \lambda_F^M$ and $W_A^M(\overline{\theta}) \geq W_A^F(\overline{\theta})$ holds otherwise. Because $W_A^F(\lambda_N^M) - W_A^N(\lambda_N^M) = 17\,452\left(a-\underline{\theta}\right)^2/5433561 > 0 > W_A^M(\lambda_N^M) - W_A^F(\lambda_N^M) = -17452\left(a-\underline{\theta}\right)^2/5433561$ holds, the cut-off values are ordered as $\lambda_F^M < \lambda_N^M < \lambda_N^F$.

When $\overline{\theta} \leq \lambda_F^M$, $\partial V_A^{FTA}\left(t_P, t_M, T\right)/\partial t_P < 0$ and $\partial V_A^{FTA}\left(t_P, t_M, T\right)/\partial t_M < 0$ always hold. In this case, Countries A and B prefer the formation of North-South FTA as soon as possible, and all three countries support the subsequent MTA formation. Therefore, $\widehat{t}_P = 0$ and $\widehat{t}_M = \varepsilon$ hold in equilibrium, where ε (> 0) is the infinitesimal length of time. When $\lambda_F^M < \overline{\theta} \leq \lambda_N^F$ is satisfied, $\partial V_A^{FTA}\left(t_P, t_M, T\right)/\partial t_P < 0$ always holds while $\partial V_A^{FTA}\left(t_P, t_M, T\right)/\partial t_M > 0$ holds if $t_M \leq T$ and $\partial V_A^{FTA}\left(t_P, t_M, T\right)/\partial t_M < 0$ holds if $t_M > T$. In this case, the equilibrium timing of the FTA formation and of the subsequent MTA formation are respectively given by $\widehat{t}_P = 0$ and $\widehat{t}_M = T$. When $\lambda_N^F < \overline{\theta}$, $\partial V_A^{FTA}(t_P, t_M, T)/\partial t_P > 0$ and $\partial V_A^{FTA}(t_P, t_M, T)/\partial t_M > 0$ hold for $t_M \leq T$ while $\partial V_A^{FTA}(t_P, t_M, T)/\partial t_P < 0$ and $\partial V_A^{FTA}(t_P, t_M, T)/\partial t_M < 0$ hold for $t_M > T$. In this case, the equilibrium timing of the FTA formation and of the subsequent MTA formation are respectively given by $\widehat{t}_P = T$ and $\widehat{t}_M = T + \varepsilon$. ■

Proof of Lemma ??

Because $d\{W_B^U(\overline{\theta}) - W_B^N(\overline{\theta})\}/d\overline{\theta} = \{238\left(a-\overline{\theta}\right) + 227(\overline{\theta}-\underline{\theta})\}/950 > 0$ and $W_B^U(\underline{\theta}) - W_B^N(\underline{\theta}) = 71\left(a-\underline{\theta}\right)^2/1900 > 0$ hold, $W_B^U(\overline{\theta}) - W_B^N(\overline{\theta}) > 0$ always holds. These inequalities mean that $\partial V_B\left(t_F, t_M, T_u\right)/\partial t_P < 0$ always holds. We can verify that $d\{W_A^U(\overline{\theta}) - W_A^N(\overline{\theta})\}/d\overline{\theta} = -(1133a - 3081\overline{\theta} + 1948\underline{\theta})/950 < -(1133a - 3081\overline{\lambda} + 1948\underline{\theta})/950 = -97\left(a-\underline{\theta}\right)/133 < 0$, $W_A^U(\underline{\theta}) - W_A^N(\underline{\theta}) = 71\left(a-\underline{\theta}\right)^2/1900 > 0$, and $W_A^U(\overline{\lambda}) - W_A^N(\overline{\lambda}) = -6\left(a-\underline{\theta}\right)^2/931 < 0$. These inequalities mean that there exists a unique cut-off value, λ_N^U ($\in (\underline{\theta}, \overline{\lambda})$), such that $W_A^U(\overline{\theta}) < W_A^N(\overline{\theta})$ holds if $\overline{\theta} > \lambda_N^U$ is satisfied and $W_A^U(\overline{\theta}) \geq W_A^N(\overline{\theta})$ holds otherwise. Regarding the subsequent MTA, because $d\{W_B^M(\overline{\theta}) - W_B^F(\overline{\theta})\}/d\overline{\theta} = -\{631\left(a-\overline{\theta}\right) + 1474(\overline{\theta}-\underline{\theta})\}/7600 < 0$, $W_B^M(\underline{\theta}) - W_B^F(\underline{\theta}) = 173\left(a-\underline{\theta}\right)^2/15200 > 0$, and $W_B^M(\overline{\lambda}) - W_B^F(\overline{\lambda}) = -3\left(a-\underline{\theta}\right)^2/1862 < 0$ hold, there exists a unique cut-off value, $\lambda_{U,B}^M$ ($\in (\underline{\theta}, \overline{\lambda})$), such that $W_B^M(\overline{\theta}) < W_B^F(\overline{\theta})$ holds if $\overline{\theta} > \lambda_{U,B}^M$ and $W_B^M(\overline{\theta}) \geq W_B^F(\overline{\theta})$ holds otherwise. In addition, $d\{W_A^M(\overline{\theta}) - W_A^U(\overline{\theta})\}/d\overline{\theta} = -(911a - 1477\overline{\theta} + 566\underline{\theta})/7600 < -(911a - 1477\overline{\lambda} + 566\underline{\theta})/7600 = -7\left(a-\underline{\theta}\right)/76 < 0$, $W_A^M(\underline{\theta}) - W_A^U(\underline{\theta}) > 0$, and $W_A^M(\overline{\lambda}) - W_A^U(\overline{\lambda}) = -\left(a-\underline{\theta}\right)^2/266 < 0$ mean that there exists a unique cut-off value, λ_U^M ($\in (\underline{\theta}, \overline{\lambda})$), such that $W_A^M(\overline{\theta}) < W_A^U(\overline{\theta})$ holds if and only if $\overline{\theta} > \lambda_U^M$ is satisfied and $W_A^M(\overline{\theta}) \geq W_A^U(\overline{\theta})$ holds otherwise.

We can calculate that $W_A^U(\lambda_N^M) - W_A^N(\lambda_N^M) = 310\,(a-\underline{\theta})^2/234099 > 0 > W_A^M(\lambda_N^M) - W_A^U(\lambda_N^M) = -310\,(a-\underline{\theta})^2/234099$. Hence, $\lambda_U^M < \lambda_N^M < \lambda_N^U$ holds. In addition, because $W_B^M(\lambda_N^M) - W_B^U(\lambda_N^M) = 70\,(a-\underline{\theta})^2/78033 > 0$ holds, we have $\lambda_U^M < \lambda_{U,B}^M$, which means that whenever the MTA formation increases the instantaneous welfare of Country A, it also increases the instantaneous welfare of Country B.

Let t_p^* and t_M^* denote the equilibrium timing of the formation of North-South CU and of the subsequent MTA. If $\overline{\theta} \le \lambda_U^M$ holds, $\partial V_h^{CU}(t_P, t_M, T)/\partial t_P < 0$ $(h \in \Omega \backslash C)$ and $\partial V_i^{CU}(t_P, t_M, T)/\partial t_M < 0$ $(i \in \Omega)$ always hold. In this case, $t_p^* = 0$ and $t_M^* = \varepsilon$ hold in equilibrium. When $\lambda_U^M < \overline{\theta} \le \lambda_N^U$ is satisfied, $\partial V_h^{CU}(t_P, t_M, T)/\partial t_P < 0$ $(h \in \Omega \backslash C)$ always holds while $\partial V_A^{CU}(t_P, t_M, T)/\partial t_M > 0$ holds if $t_M \le T$ and $\partial V_i^{CU}(t_P, t_M, T)/\partial t_M < 0$ $(i \in \Omega)$) holds if $t_M > T$. In this case, Country A prefers to delay the timing of the MTA until T_u, but Countries A and B support the formation of North-South CU at $t = 0$. Hence, the optimal timings of trade agreements become $t_p^* = 0$ and $t_M^* = T$. When $\lambda_N^F < \overline{\theta}$ is satisfied, $\partial V_A^{CU}(t_P, t_M, T)/\partial t_P > 0$ and $\partial V_A^{CU}(t_P, t_M, T)/\partial t_M > 0$ hold if $t_M \le T$ while $\partial V_A^{CU}(t_P, t_M, T)/\partial t_P < 0$ and $\partial V_A^{CU}(t_P, t_M, T)/\partial t_M < 0$ hold if $t_M > T$. In this case, the equilibrium timing of the FTA formation and of the MTA are respectively given by $t_p^* = T$ and $t_M^* = T + \varepsilon$. ■

Proof of Lemma ??

By the proofs of Lemma **??** and Lemma **??**, $\lambda_F^M < \lambda_N^M < \lambda_N^F$ and $\lambda_U^M < \lambda_N^M < \lambda_N^U$ hold. Because $\{W_A^F(\lambda_N^M) - W_A^N(\lambda_N^M)\} - \{W_A^U(\lambda_N^M) - W_A^N(\lambda_N^M)\} = 0.0018877\,(a-\underline{\theta})^2 > 0$ and $d\{W_A^U(\overline{\theta}) - W_A^N(\overline{\theta})\}/d\overline{\theta} - d\{W_A^F(\overline{\theta}) - W_A^N(\overline{\theta})\}/d\overline{\theta} = -20[42(\overline{\lambda} - \overline{\theta}) + 31(\overline{\theta} - \underline{\theta})]/1197 < 0$, we have $\lambda_N^U < \lambda_N^F$. Besides that, $W_A^M(\lambda_F^M) - W_A^U(\lambda_F^M) = 0.00044727\,(a-\underline{\theta})^2 > 0$ means that $\lambda_F^M < \lambda_U^M$ holds. By these comparisons, we have $\lambda_F^M < \lambda_U^M < \lambda_N^M < \lambda_N^U < \lambda_N^F$. ■

Proof of Propostion ??

(i) Suppose $\lambda_N^M \le \overline{\theta} < \lambda_N^F$ holds. By (**??**) and (**??**), the optimal timing of technology adoption on the MTA path and on the FTA path are respectively determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ and $rk(T_f) - k'(T_f) = \pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})$. We can calculate that $\{rk(T_m) - k'(T_m)\} - \{rk(T_f) - k'(T_f)\} = \pi_A^F(\overline{\theta}) - \pi_A^N(\overline{\theta})$. Because $d\{\pi_A^F(\overline{\theta}) - \pi_A^N(\overline{\theta})\}/d\overline{\theta} = -[801(a-\overline{\theta}) + 5434(\overline{\theta} - \underline{\theta})]/22050 < 0$ and $\pi_A^F(\lambda_N^M) - \pi_A^N(\lambda_N^M) = -67540\,(a-\underline{\theta})^2/5433561 < 0$, $\pi_A^F(\overline{\theta}) - \pi_A^N(\overline{\theta}) < 0$ always holds under $\lambda_N^M \le \overline{\theta} < \lambda_N^F$. Hence, $\{rk(T_m) - k'(T_m)\} < \{rk(T_f) - k'(T_f)\}$ holds. Since $rk(T) - k'(T)$ is decreasing in T, we have $T_f < T_m$. By Lemma **??** and Lemma **??**, the optimal timing of technology adoption coincides with the timing of the MTA on each path .In this case, the FTA path realizes technology adoption and multilateral free trade faster than the MTA path does.

(ii) Suppose $\lambda_F^M \le \overline{\theta} < \lambda_N^M$ holds in which T_m and T_f are respectively determined by $rk(T_m) - k'(T_m) =$

$\pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})$ and $rk(T_f) - k'(T_f) = \pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})$. Because we can verify that $\{rk(T_m) - k'(T_m)\} - \{rk(T_f) - k'(T_f)\} = \pi_A^F(\overline{\theta}) - \pi_A^M(\overline{\theta})$, $d\{\pi_A^F(\overline{\theta}) - \pi_A^M(\overline{\theta})\}/d\overline{\theta} > 0$, and $\pi_A^F(\lambda_F^M) - \pi_A^M(\lambda_F^M) = 7.9912 \times 10^{-4}(a-\underline{\theta})^2 > 0$, $\pi_A^F(\overline{\theta}) - \pi_A^M(\overline{\theta}) > 0$ holds in this case. This inequality means that $T_m < T_f$ holds. By Lemma **??** and Lemma **??**, multilateral free trade is realized at $t = 0$ on the MTA path and at $t = T_f > 0$ on the FTA path.

(iii) Suppose $\overline{\theta} < \lambda_F^M$ holds. By Lemma **??** and Lemma **??**, multilateral free trade is realized at $t = 0$ on the MTA path and at $t = \varepsilon$ on the FTA path. Hence, the FTA path delays the realization of multilateral free trade. Firm A determines the timing of technology adoption given that multilateral free trade prevails both before and after technology adoption. It is determined $rk(T_m) - k'(T_m) = rk(T_f) - k'(T_f) = \pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})$, which means $T_f = T_m$.

(iv) Suppose $\lambda_N^F \leq \overline{\theta}$ holds, under which T_m is determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ while T_f is determined by $rk(T_f) - k'(T_f) = \pi_A^F(\underline{\theta}) - \pi_A^N(\overline{\theta}) + e^{-r\varepsilon}\left[\pi_A^M(\underline{\theta}) - \pi_A^F(\underline{\theta})\right]$. By comparing these equations, we have $\{rk(T_m) - k'(T_m)\} - \{rk(T_f) - k'(T_f)\} = (1 - e^{-r\varepsilon})\left[\pi_A^M(\underline{\theta}) - \pi_A^F(\underline{\theta})\right] = 279(1 - e^{-r\varepsilon})(a-\underline{\theta})^2/19600 > 0$. Hence, $T_m < T_f$ holds. In this case, multilateral free trade is realized at T_m on the MTA path and at $T_f + \varepsilon$ on the FTA path. Therefore, the FTA path delays or does not change the timing of technology adoption and delays the realization of multilateral free trade. ■

Proof of Propostion ??

(i) Suppose $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$ holds. By (**??**) and (**??**), the optimal timing of technology adoption on the MTA path and of the CU path are respectively determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ and $.rk(T_u) - k'(T_u) = \pi_A^M(\underline{\theta}) - \pi_A^U(\overline{\theta})$. We can calculate that $\{rk(T_m) - k'(T_m)\} - \{rk(T_f) - k'(T_f)\} = \pi_A^U(\overline{\theta}) - \pi_A^N(\overline{\theta})$. Because $d\{\pi_A^U(\overline{\theta}) - \pi_A^N(\overline{\theta})\}/d\overline{\theta} = -3[907(a-\overline{\theta}) + 1438(\overline{\theta}-\underline{\theta})]/18050 < 0$ and $\pi_A^F(\overline{\lambda}) - \pi_A^N(\overline{\lambda}) = 124(a-\underline{\theta})^2/17689 > 0$ hold, $\pi_A^U(\overline{\theta}) - \pi_A^N(\overline{\theta}).> 0$ always holds for all $\overline{\theta}$. As a result, we have $T_u > T_m$ because $\{rk(T_U) - k'(T_U)\} > \{rk(T_f) - k'(T_f)\}$ holds. By Lemma **??** and Lemma **??**, the optimal timing of technology adoption coincides with the timing of the MTA on each path. Therefore, the MTA path realizes technology adoption and multilateral free trade faster than the CU path does.

(ii) Suppose $\lambda_U^M \leq \overline{\theta} < \lambda_N^M$ holds, under which T_m and T_u are respectively determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})$ and $rk(T_u) - k'(T_u) = \pi_A^M(\underline{\theta}) - \pi_A^U(\overline{\theta})$. We can verify that $\{rk(T_m) - k'(T_m)\} - \{rk(T_u) - k'(T_u)\} = \pi_A^U(\overline{\theta}) - \pi_A^M(\overline{\theta})$, $\partial\{\pi_A^U(\overline{\theta}) - \pi_A^M(\overline{\theta})\}/\partial\overline{\theta} > 0$, and $\pi_A^U(\underline{\theta}) - \pi_A^M(\underline{\theta}) = 3169(a-\underline{\theta})^2/144400 > 0$, $\pi_A^U(\overline{\theta}) - \pi_A^M(\overline{\theta}) > 0$ always holds. These inequalities mean that $T_m < T_u$ holds. By Lemma **??** and Lemma **??**, multilateral free trade is realized at $t = 0$ on the MTA path and at $t = T_U > 0$ on the CU path.

(iii) Suppose $\overline{\theta} < \lambda_N^M$ holds where multilateral free trade is realized at $t = 0$ on the MTA path and at $t = \varepsilon$ on the CU path. Hence, the CU path delays the realization of multilateral free trade. Firm A

determines the timing of technology adoption given that multilateral free trade prevails both before and after the technology adoption. It is determined $rk(T_m) - k'(T_m) = rk(T_u) - k'(T_u) = \pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})$, which means $T_u = T_m$.

(iv) Suppose $\lambda_N^U \leq \overline{\theta}$ holds, under which T_m and T_u are respectively determined by $rk(T_m) - k'(T_m) = \pi_A^M(\underline{\theta}) - \pi_A^N(\overline{\theta})$ and $rk(T_u) - k'(T_u) = \pi_A^U(\underline{\theta}) - \pi_A^N(\overline{\theta}) + e^{-r\varepsilon}\left[\pi_A^M(\underline{\theta}) - \pi_A^U(\underline{\theta})\right]$. Because we have $\{rk(T_m) - k'(T_m)\} - \{rk(T_u) - k'(T_u)\} = (1 - e^{-r\varepsilon})\left[\pi_A^M(\underline{\theta}) - \pi_A^U(\underline{\theta})\right] < 0$, $T_u < T_m$ holds. Multilateral free trade is realized at $t = T_m$ on the MTA path and at $t = T_u + \varepsilon$ on the CU path. If $T_m - T_u > \varepsilon$ holds, the realization of multilateral free trade is faster on the CU path than on the MTA path. ■

Proof of Lemma ??

(i) By the proof of Lemmas **??**, **??** and **??**, any trade agreement increases instantaneous world welfare after Firm A adopts the new technology. Hence, $WW^M(\underline{\theta}) > WW^F(\underline{\theta}) > WW^N(\underline{\theta})$ and $WW^M(\underline{\theta}) > WW^U(\underline{\theta}) > WW^N(\underline{\theta})$ hold. We also have $WW^F(\underline{\theta}) - WW^U(\underline{\theta}) = 320\,(a - \underline{\theta})^2/17689 > 0$.

Suppose $\theta_A = \overline{\theta}$ holds, under which $WW^M(\overline{\theta}) - WW^F(\overline{\theta})$ is an U-shaped function in $\overline{\theta}$ and takes the minimum value at $\overline{\theta} = \eta := (4023a + 42\,398\underline{\theta})/46421$. Because we have $WW^M(\eta) - WW^F(\eta) = 0.066894\,(a - \underline{\theta})^2 > 0$, $WW^M(\overline{\theta}) > WW^F(\overline{\theta})$ holds for any $\overline{\theta} \in (\underline{\theta}, \overline{\lambda})$. We can also calculate that $d\{WW^F(\overline{\theta}) - WW^U(\overline{\theta})\}/d\overline{\theta} = 4\{17\,577(\overline{\lambda} - \overline{\theta}) + 12\,820(\overline{\theta} - \underline{\theta})\}/159201 > 0$. Since $WW^F(\underline{\theta}) > WW^U(\underline{\theta})$ holds, $WW^F(\overline{\theta}) > WW^U(\overline{\theta})$ also holds. Furthermore, we have $d\{WW^U(\overline{\theta}) - WW^N(\overline{\theta})\}/d\overline{\theta} = -\{2487(a - \overline{\theta}) + 4493(\overline{\theta} - \underline{\theta})\}/18050 < 0$ and $WW^U(\overline{\lambda}) - WW^N(\overline{\lambda}) = 698\,(a - \underline{\theta})^2/17689 > 0$, which means that $WW^U(\overline{\theta}) > WW^N(\overline{\theta})$ holds. Therefore, $WW^M(\theta_A) > WW^F(\theta_A) > WW^U(\theta_A) > WW^N(\theta_A)$ holds given θ_A.

(ii) By using $\overline{\theta} < \overline{\lambda}$, we have $dWW^N(\overline{\theta})/d\overline{\theta} = -(21a - 107\overline{\theta} + 86\underline{\theta})/25 < -(21a - 107\overline{\lambda} + 86\underline{\theta})/25 = -8(a - \underline{\theta})/35 < 0$, $dWW^M(\overline{\theta})/d\overline{\theta} = -3(5a - 23\overline{\theta} + 18\underline{\theta})/16 < -3(5a - 23\overline{\lambda} + 18\underline{\theta})/16 = -9(a - \underline{\theta})/28 < 0$, $dWW^F(\overline{\theta})/d\overline{\theta} = -(20169a - 89288\overline{\theta} + 69119\underline{\theta})/22050 < -(20169a - 89288\overline{\lambda} + 69119\underline{\theta})/22050 = -10379\,(a - \underline{\theta})/30870 < 0$, and $dWW^U(\overline{\theta})/d\overline{\theta} = -(17649a - 75248\overline{\theta} + 57599\underline{\theta})/18050 < -(17649a - 75248\overline{\lambda} + 57599\underline{\theta})/18050 = -9659(a - \underline{\theta})/25270$. Hence, $WW^{R_t}(\underline{\theta}) > WW^{R_t}(\overline{\theta})$ holds given R_t. ■

Proof of Proposition ??

Given T, the discounted sum of world welfare along the MTA path, the FTA path, and the CU path are respectively given by $WV^{MTA}(T) = \sum_{i\in\Omega} V_i^{MTA}(\widetilde{t}_M, T)$, $WV^{FTA}(T) = \sum_{i\in\Omega} V_i^{FTA}(\widehat{t}_P, \widehat{t}_M, T)$, and $WV^{CU}(T) = \sum_{i\in\Omega} V_i^{CU}(t_P^*, t_M^*, T)$.

(i) Suppose $\overline{\theta} < \lambda_F^M$ holds. In this case, we have $T_m = T_f = T_u$ and $\widetilde{t}_M = 0 < \widehat{t}_M = t_M^* = \varepsilon$, which means that the MTA path realizes multilateral free trade faster than the other paths. By Lemma **??**, $WV^{MTA}(T_m) > WV^{FTA}(T_f) > WV^{CU}(T_u)$ holds.

(ii) Suppose $\lambda_F^M \leq \overline{\theta} < \lambda_U^M$ holds. In this case, the speed of attaining multilateral free trade is the fastest on the MTA path ($t = 0$). Since the CU path yields the same timing of technology adoption while it delays the realization of multilateral free trade, it is evident that $WV^{MTA}(T_m) > WV^{CU}(T_u)$ holds. On the FTA path, the timing of technology adoption coincides with the timing of the MTA formation. Given this property, we have

$$\frac{dWV^{FTA}(T)}{dT} = -e^{-rT}[\{WW^M(\underline{\theta}) - WW^F(\overline{\theta})\} - \{rk(T) - k'(T)\}]. \tag{A.1}$$

Because $\pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta}) = rk(T_m) - k'(T_m)$ holds in this case, $(dWV^{FTA}(T)/dT)\big|_{T=T_m} = -e^{-rT_m}[\{WW^M(\underline{\theta}) - WW^F(\overline{\theta})\} - \{\pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})\}]$ holds. Because $d[\{WW^M(\underline{\theta}) - WW^F(\overline{\theta})\} - \{\pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})\}]/d\overline{\theta} = -\{18549(a - \overline{\theta}) + 78\,026(\overline{\theta} - \underline{\theta})\}/88200 < 0$, $WW^M(\underline{\theta}) > WW^F(\underline{\theta})$, and $\{WW^M(\underline{\theta}) - WW^F(\overline{\lambda})\} - \{\pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\lambda})\} = 21407\,(a - \underline{\theta})^2/691488 > 0$ hold, we have $WW^M(\underline{\theta}) - WW^F(\overline{\theta}) > \pi_A^M(\underline{\theta}) - \pi_A^M(\overline{\theta})$. This inequality means that $dWV^{FTA}(T)/dT < 0$ holds for $T \geq T_m$. Because $T_f > T_m$ holds, we have $WV^{FTA}(T_f) < WV^{FTA}(T_m)$. By using this inequality and Lemma **??**, we can calculate that $WV^{MTA}(T_m) - WV^{FTA}(T_f) > WV^{MTA}(T_m) - WV^{FTA}(T_m) = [(1 - e^{-rT_m})\{WW^M(\overline{\theta}) - WW^N(\overline{\theta})\}]/r > 0$. Hence, the MTA path realizes the highest world welfare.

(iii) Suppose $\lambda_U^M \leq \overline{\theta} < \lambda_N^M$ holds. In this case, the speed of attaining multilateral free trade is the fastest on the MTA path. By the same calculation as case (ii), $WV^{MTA}(T_m) > WV^{FTA}(T_f)$ holds because $dWV^{FTA}(T)/dT < 0$ holds for $T \geq T_m$. By Lemma **??** and $T_f < T_u$, we can confirm that $WV^{FTA}(T_f) - WV^{CU}(T_u) > WV^{FTA}(T_u) - WV^{CU}(T_u) = (1 - e^{-rT_u})\{WW^F(\overline{\theta}) - WW^U(\overline{\theta})\}/r > 0$ holds. Hence, the MTA path realizes the highest world welfare.

(iv) Suppose $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$ holds. In this case, the speed of attaining multilateral free trade is the fastest on the FTA path. Since $dWV^{FTA}(T)/dT$ is given by Eq. (A.1) and $\pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta}) = rk(T_f) - k'(T_f)$ holds in this case, $(dWV^{FTA}(T)/dT)\big|_{T=T_f} = -e^{-rT_f}[\{WW^M(\underline{\theta}) - WW^F(\overline{\theta})\} - \{\pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})\}]$ holds. We can verify that $d[\{WW^M(\underline{\theta}) - WW^F(\overline{\theta})\} - \{\pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\theta})\}]/d\overline{\theta} = -\{306(a - \overline{\theta}) + 7799(\overline{\theta} - \underline{\theta})\}/7350 < 0$ and $\{WW^M(\underline{\theta}) - WW^F(\overline{\lambda})\} - \{\pi_A^M(\underline{\theta}) - \pi_A^F(\overline{\lambda})\} = 8597\,(a - \underline{\theta})^2/23490 > 0$ hold, which means that $dWV^{FTA}(T)/dT < 0$ holds for $T \geq T_f$. By using this property and $T_f < T_m < T_u$, we have $WV^{FTA}(T_f) - WV^{MTA}(T_m) > WV^{FTA}(T_m) - WV^{MTA}(T_m) = (1 - e^{-rT_m})\{WW^F(\overline{\theta}) - WW^N(\overline{\theta})\}/r > 0$ and $WV^{FTA}(T_f) - WV^{CU}(T_u) > WV^{FTA}(T_u) - WV^{CU}(T_u) = (1 - e^{-rT_u})\{WW^F(\overline{\theta}) - WW^U(\overline{\theta})\}/r > 0$. These inequalities imply that the FTA path realizes the highest world welfare.

(v) Suppose $\lambda_N^U \leq \overline{\theta} < \lambda_N^F$, under which the speed of attaining multilateral free trade is faster on the FTA path than on the MTA path. By the same reason as case (iv), $WV^{FTA}(T_f) > WV^{MTA}(T_m)$ holds. If $T_u + \varepsilon < T_m$ holds, the timing of attaining multilateral free trade is faster on the CU path than on the MTA path. Following the same procedure as above, we can verify that $\partial WV^{CU}(T)/\partial T < 0$ holds for $T \geq T_u$.[23] Given $T_u + \varepsilon < T_m$, we have $WV^{CU}(T_u) - WV^{MTA}(T_m) > WV^{CU}(T_m - \varepsilon) - WV^{MTA}(T_m) =$

[23]The detailed calculation will be provided upon request.

$e^{-r(T_m-\varepsilon)}[(1-e^{-r\varepsilon})\{WW^U(\underline{\theta})-WW^N(\overline{\theta})\}/r-\{k(T_m-\varepsilon)-e^{-r\varepsilon}k(T_m)\}]>\lim_{\varepsilon\to 0}[WV^{CU}(T_m-\varepsilon)-WV^{MTA}(T_m)]=0$. Therefore, $WV^{CU}(T_u)>WV^{MTA}(T_m)$ is satisfied if $T_u+\varepsilon<T_m$ holds.

(vi) Suppose $\lambda_N^F\le\overline{\theta}$, under which the speed of attaining multilateral free trade is the fastest on the MTA path ($t=T_m$). We can calculate that $\partial[\{WW^M(\underline{\theta})-WW^N(\overline{\theta})\}-\{\pi_A^M(\underline{\theta})-\pi_A^N(\overline{\theta})\}]/\partial\overline{\theta}=-2\{(a-\overline{\theta})+14(\overline{\theta}-\underline{\theta})\}/25<0$ and $WW^M(\underline{\theta})-WW^N(\overline{\lambda})\}-\{\pi_A^M(\underline{\theta})-\pi_A^N(\overline{\lambda})\}=183(a-\underline{\theta})^2/1568>0$ hold. These inequalities means that $\partial WV^{MTA}(T)/\partial T<0$ holds for $T\ge T_m$. By using this property, we have $WV^{MTA}(T_m)-WV^{FTA}(T_f)>WV^{MTA}(T_f)-WV^{FTA}(T_f)=e^{-rT_f}(1-e^{-r\varepsilon})\left[WW^M(\underline{\theta})-WW^F(\underline{\theta})\right]/r>0$ and $WV^{MTA}(T_m)-WV^{CU}(T_u)>WV^{MTA}(T_u)-WV^{CU}(T_u)=e^{-rT_u}(1-e^{-r\varepsilon})\{WW^M(\underline{\theta})-WW^U(\underline{\theta})\}/r>0$. Thus, the MTA path realizes the highest world welfare.

By these comparisons, we have $WV^{MTA}(T_m)>\max[WV^{FTA}(T_f),WV^{CU}(T_u)]$ if $\widetilde{t}_M<\min[\widehat{t}_M,t_M^*]$ holds, $WV^{FTA}(T_f)>WV^{MTA}(T_m)$ if $\widehat{t}_M<\widetilde{t}_M$ holds, and $WV^{CU}(T_u)>WV^{MTA}(T_m)$ if $t_M^*<\widetilde{t}_M$ holds. ■

Proof of Proposition ??

The timing of technology adoption on the FTA path is weakly faster than that on the CU path (i.e., $\widehat{t}_M\le t_M^*$) if either (a) $\overline{\theta}<\lambda_F^M$ or (b) $\lambda_U^M\le\overline{\theta}<\lambda_N^U$ holds, and it can be faster if (c) $\lambda_N^U\le\overline{\theta}<\lambda_N^F$ holds.

(a) Suppose $\overline{\theta}<\lambda_F^M$ holds under which $T_f=T_u$ holds. By (i) in the proof of Proposition **??**, $WV^{FTA}(T_f)>WV^{CU}(T_u)$ holds. (b) Suppose $\lambda_U^M\le\overline{\theta}<\lambda_N^U$ holds under which $T_f<T_u$ holds. By (iii) and (iv) in the proof of Proposition **??**, $WV^{FTA}(T_f)>WV^{CU}(T_u)$ holds. (c) Suppose $\lambda_N^U\le\overline{\theta}<\lambda_N^F$ holds. In this case, $T_f<T_u$ is satisfied if $\pi_A^M(\underline{\theta})-\pi_A^F(\overline{\theta})>\pi_A^U(\underline{\theta})-\pi_A^N(\overline{\theta})+e^{-r\varepsilon}\left[\pi_A^M(\underline{\theta})-\pi_A^U(\underline{\theta})\right]$ holds. Since $\partial WV^{FTA}(T)/\partial T<0$ holds for $T\ge T_f$, we have $WV^{FTA}(T_f)-WV^{CU}(T_u)>WV^{FTA}(T_u)-WV^{CU}(T_u)=[\left(1-e^{-rT_u}\right)\{WW^F(\overline{\theta})-WW^N(\overline{\theta})\}+e^{-rT_u}(1-e^{-r\varepsilon})\{WW^M(\underline{\theta})-WW^U(\underline{\theta})\}]/r>0$.

By these comparisons, we have $WV^{FTA}(T_f)>WV^{CU}(T_u)$ if $T_f\le T_u$ holds. ■

Proof of Proposition ??

Suppose $\lambda_N^M\le\overline{\theta}<\lambda_N^U$ under which, compared to the MTA path, the FTA path accelerates the realization of multilateral free trade and the CU path delays it. By Proposition **??**, the highest world welfare is realized on the FTA path. By comparing $V_B^{FTA}(\widehat{t}_P,\widehat{t}_M,T_f)$ with $V_B^{MTA}\left(\widetilde{t}_M,T_m\right)$, we have $V_B^{FTA}(\widehat{t}_P,\widehat{t}_M,T_f)-V_B^{MTA}\left(\widetilde{t}_M,T_m\right)=[\left(1-e^{-rT_m}\right)\{W_B^F(\overline{\theta})-W_B^N(\overline{\theta})\}+(e^{-rT_f}-e^{-rT_m})\{W_B^M(\underline{\theta})-W_B^F(\overline{\theta})\}]/r$. By the proof of Lemma **??**, $W_B^F(\overline{\theta})>W_B^N(\overline{\theta})$ holds. Besides that, because $d\{W_B^M(\underline{\theta})-W_B^F(\overline{\theta})\}/d\overline{\theta}=-\{8523(a-\overline{\theta})+19892(\overline{\theta}-\underline{\theta})\}/22050<0$ and $W_B^M(\underline{\theta})-W_B^F(\overline{\lambda})=162503(a-\underline{\theta})^2/691488>0$, we have $W_B^M(\underline{\theta})>W_B^F(\overline{\theta})$. Therefore, $V_B^{FTA}(\widehat{t}_P,\widehat{t}_M,T_f)>V_B^{MTA}\left(\widetilde{t}_M,T_m\right)$ holds.

By comparing $V_B^{CU}(t_P^*,t_M^*,T_u)$ with $V_B^{FTA}(\widehat{t}_P,\widehat{t}_M,T_f)$, we have $V_B^{CU}(t_P^*,t_M^*,T_u)-V_B^{FTA}(\widehat{t}_P,\widehat{t}_M,T_f)=[\left(1-e^{-rT_f}\right)\{W_B^U(\overline{\theta})-W_B^F(\overline{\theta})\}+\left(e^{-rT_f}-e^{-rT_u}\right)\{\left(W_B^U(\overline{\theta})-W_B^M(\underline{\theta})\right)\}]/r$. Because $d\{W_B^U(\overline{\theta})-W_B^F(\overline{\theta})\}$

$/d\overline{\theta} = 16\{273(\overline{\lambda} - \overline{\theta}) + 232(\overline{\theta} - \underline{\theta})\}/8379 > 0$ and $W_B^U(\underline{\theta}) - W_B^F(\underline{\theta}) = 8(a-\underline{\theta})^2/931 > 0$ hold, $W_B^U(\overline{\theta}) > W_B^F(\overline{\theta})$ holds. Besides that, because $d\{W_B^U(\overline{\theta}) - W_B^M(\underline{\theta})\}/d\overline{\theta} = \{257(a-\overline{\theta}) + 778(\overline{\theta} - \underline{\theta})\}/950 > 0$ and $W_B^U(\lambda_N^M) - W_B^M(\underline{\theta}) = 180259(a-\underline{\theta})^2/7491168 > 0$ hold, we also have $W_B^U(\overline{\theta}) > W_B^M(\underline{\theta})$. Therefore, $V_B^{CU}(t_P^*, t_M^*, T_u) > V_B^{FTA}(\widehat{t}_P, \widehat{t}_M, T_f)$ and Country B attains the highest present value of welfare on the CU path if $\lambda_N^M \leq \overline{\theta} < \lambda_N^U$ holds. ■

Proof of Proposition ??

By calculating the optimal external tariffs under the North-North FTA and the North-North CU, and substituting them into welfare functions, the equilibrium welfare of country i ($i \in \Omega$) in the stage-2 subgame under the North-North FTA and under the North-North CU are respectively given by $W_i^{FN}(\theta_A)$ and $W_i^{UN}(\theta_A)$.

In the case of the North-North FTA, positive bilateral trade between countries are guaranteed as long as $\overline{\theta} < \overline{\lambda}$ holds. By using $\overline{\theta} < \overline{\lambda}$, it can be calculated that $d\{W_j^{NFTA}(\overline{\theta}) - W_j^N(\overline{\theta})\}/d\overline{\theta} = -(52a - 219\overline{\theta} + 167\underline{\theta})/2450 < -(52a - 219\overline{\lambda} + 167\underline{\theta})/2450 = -29(a-\underline{\theta})/3430 < 0$ ($j \in \Omega\backslash A$) holds. By using this property, we have $W_j^{FN}(\overline{\theta}) - W_j^N(\overline{\theta}) > W_j^{FN}(\overline{\lambda}) - W_j^N(\overline{\lambda}) = 64(a-\underline{\theta})^2/2401 > 0$. In this case, $d\{W_j^M(\overline{\theta}) - W_j^{FN}(\overline{\theta})\}/d\overline{\theta} = (3699a - 4153\overline{\theta} + 454\underline{\theta})/19600 > (3699a - 4153\overline{\lambda} + 454\underline{\theta})/19600 = 1087(a-\underline{\theta})/6860 > 0$ ($j \in \Omega\backslash A$) holds and so we have $W_j^M(\overline{\theta}) - W_j^{FN}(\overline{\theta}) > W_j^{FN}(\underline{\theta}) - W_j^N(\underline{\theta}) = 783(a-\underline{\theta})^2/39200 > 0$. Hence, $W_j^M(\theta_A) > W_j^{FN}(\theta_A) > W_j^N(\theta_A)$ ($j \in \Omega\backslash A$) holds given $\theta_A \in \{\underline{\theta}, \overline{\theta}\}$.

In the case with the North-North CU, there are no exports from Country A to Countries B and C if $\overline{\theta} \geq \overline{\mu} := (a + 8\underline{\theta})/9$ holds where $\overline{\mu} < \overline{\lambda}$. Given $\overline{\theta} < \overline{\mu}$, it is calculated that $d\{W_j^{UN}(\overline{\theta}) - W_j^N(\overline{\theta})\}/d\overline{\theta} = -(12a - 89\overline{\theta} + 77\underline{\theta})/950 < -(12a - 89\overline{\mu} + 77\underline{\theta})/950 = -(a-\underline{\theta})/450 < 0$ ($j \in \Omega\backslash A$) and then we have $W_j^{UN}(\overline{\theta}) - W_j^N(\overline{\theta}) > W_j^{UN}(\overline{\mu}) - W_j^N(\overline{\mu}) = 74(a-\underline{\theta})^2/2025 > 0$. Besides that, $d\{W_j^M(\overline{\theta}) - W_j^{UN}(\overline{\theta})\}/d\overline{\theta} = (1369a - 1643\overline{\theta} + 274\underline{\theta})/7600 > (1369a - 1643\overline{\mu} + 274\underline{\theta})/7600 = 281(a-\underline{\theta})/1800 > 0$ ($j \in \Omega\backslash A$) holds and so we have $W_j^M(\overline{\theta}) - W_j^{UN}(\overline{\theta}) > W_j^M(\underline{\theta}) - W_j^{UN}(\underline{\theta}) = 173(a-\underline{\theta})^2/15200 > 0$. Because there are no exports from Country A for $\overline{\mu} \leq \overline{\theta} < \overline{\lambda}$, $W_j^{UN}(\theta_A)$ becomes independent of θ_A and the instantaneous welfare is given by $W_j^{UN}(\overline{\mu})$ under this range of the parameterization. We have $W_j^{UN}(\overline{\mu}) - W_j^N(\overline{\theta}) = (31a - 57\overline{\theta} + 26\underline{\theta})\{a - \overline{\theta} + 4(\overline{\theta} - \underline{\theta})\}/900 > (31a - 57\overline{\lambda} + 26\underline{\theta})\{a - \overline{\theta} + 4(\overline{\theta} - \underline{\theta})\}/900 = 8(a-\underline{\theta})\{a - \overline{\theta} + 4(\overline{\theta} - \underline{\theta})\}/315 > 0$. Besides that, we have $\{W_j^M(\overline{\theta}) - W_j^{UN}(\overline{\mu})\}/d\overline{\theta} = \{51(a-\overline{\theta}) + 154(\overline{\theta} - \underline{\theta})\}/400 > 0$ and $W_j^M(\overline{\theta}) - W_j^{UN}(\overline{\mu}) > W_j^M(\overline{\mu}) - W_j^{UN}(\overline{\mu}) = 487(a-\underline{\theta})^2/16200 > 0$. Hence, $W_j^M(\theta_A) > W_j^{UN}(\theta_A) > W_j^N(\theta_A)$ ($j \in \Omega\backslash A$) holds given θ_A.

By conducting the same analysis as the baseline model, we can verify that each country supports a trade agreement if the agreement increases its instantaneous welfare given θ_A. The above inequalities imply that Countries B and C form the North-North FTA and the North-North CU immediately at $t = 0$ on the FTA path and on the CU path respectively, and they always support the MTA on all three paths.

Hence, the MTA is concluded if and only if Country A supports the MTA.

With regard to Country A's welfare, we have $d\{W_A^M(\overline{\theta})-W_A^{FN}(\overline{\theta})\}/d\overline{\theta} = -(1833a-5891\overline{\theta}+4058\underline{\theta})/3920 < -(1833a - 5891\overline{\lambda} + 4058\underline{\theta})/3920 = -347\left(a-\underline{\theta}\right)/1372 < 0$, $W_A^M(\underline{\theta}) - W_A^{FN}(\underline{\theta}) = 219\left(a-\underline{\theta}\right)^2/7840 > 0$, and $W_A^M(\lambda_N^M) - W_A^{FN}(\lambda_N^M) = -9976\left(a-\underline{\theta}\right)^2/603729 < 0$. These inequalities mean that there exists a unique cut-off value, μ_F^M $(\in (\underline{\theta}, \lambda_N^M))$, such that $W_A^M(\overline{\theta}) < W_A^{FN}(\overline{\theta})$ holds if $\overline{\theta} > \mu_F^M$ is satisfied and $W_A^M(\overline{\theta}) \geq W_A^{FN}(\overline{\theta})$ holds otherwise. For $\overline{\theta} < \overline{\mu}$, we have $d\{W_A^M(\overline{\theta}) - W_A^{UN}(\overline{\theta})\}d\partial\overline{\theta} = -\left(17697a - 38699\overline{\theta} + 21002\underline{\theta}\right)/28880 < -\left(17697a - 38699\overline{\mu} + 21002\underline{\theta}\right)/28880 = -167\left(a-\underline{\theta}\right)/360 < 0$ and $W_A^M(\overline{\mu}) - W_A^{UN}(\overline{\mu}) = 11\left(a-\underline{\theta}\right)^2/3240 > 0$. For $\overline{\theta} \geq \overline{\mu}$, $W_A^{UN}(\overline{\theta})$ is calculated given $q_B^A = q_C^A = 0$. In this case, we have $d\{W_A^M(\overline{\theta}) - W_A^{UN}(\overline{\theta})\}d\overline{\theta} = -\left(57a - 179\overline{\theta} + 122\underline{\theta}\right)/80 < -\left(57a - 179\overline{\lambda} + 122\underline{\theta}\right)/80 = -11\left(a-\underline{\theta}\right)/28 < 0$ and $W_A^M(\overline{\lambda}) - W_A^{UN}(\overline{\lambda}) = -\left(a-\underline{\theta}\right)^2/98 < 0$. These inequalities mean that there exists a unique cut-off value, μ_U^M $(\in (\lambda_N^M, \overline{\lambda}))$, such that $W_A^M(\overline{\theta}) < W_A^{UN}(\overline{\theta})$ holds if $\overline{\theta} > \mu_U^M$ is satisfied and $W_A^M(\overline{\theta}) \geq W_A^{UN}(\overline{\theta})$ holds otherwise. Note that $\overline{\mu} < \mu_U^M$ holds.

Let us compare the equilibrium timing of technology adoption and of attaining multilateral free trade on the CU path with those of the MTA path. (i) Suppose $\underline{\theta} < \overline{\theta} < \lambda_N^M$ holds under which multilateral free trade is realized at $t = 0$ on the MTA path and it is realized at $t = \varepsilon$ on the CU path. Hence, the CU path delays the realization of multilateral free trade. The timing of technology adoption is determined by $rk(T_m) - k'(T_m) = rk(T_u) - k'(T_u) = \pi_A^M\left(\underline{\theta}\right) - \pi_A^M(\overline{\theta})$, which means $T_m = T_u$. (ii) Suppose $\lambda_N^M \leq \overline{\theta} < \mu_U^M$ holds under which T_m is determined by $rk(T_m) - k'(T_m) = \pi_A^M\left(\underline{\theta}\right) - \pi_A^N(\overline{\theta})$ while T_u is determined by $rk(T_u) - k'(T_u) = \pi_A^M\left(\underline{\theta}\right) - \pi_A^M\left(\overline{\theta}\right)$. We have $\{\pi_A^M\left(\underline{\theta}\right) - \pi_A^N(\overline{\theta})\} - \{\pi_A^M\left(\underline{\theta}\right) - \pi_A^M\left(\overline{\theta}\right)\} = \pi_A^M\left(\overline{\theta}\right) - \pi_A^N(\overline{\theta}) = \left(a - 27\overline{\theta} + 26\underline{\theta}\right)\{3\left(a-\overline{\theta}\right) + 2\left(\overline{\theta}-\underline{\theta}\right)\}/400 \leq (a - 27\lambda_N^M + 26\underline{\theta})\{3\left(a-\overline{\theta}\right) + 2\left(\overline{\theta}-\underline{\theta}\right)\}/400 = \left(a-\overline{\theta}\right)\{\left(a-\overline{\theta}\right) + 2\left(\overline{\theta}-\underline{\theta}\right)\}/185 < 0$. The inequality means that $T_u < T_m$ holds. Since multilateral free trade is realized at $t = T_m$ on the MTA path and at $t = \varepsilon$ on the CU path, the CU path accelerates both the timing of technology adoption and the realization of multilateral free trade. (iii) Suppose $\mu_U^M \leq \overline{\theta} < \overline{\lambda}$ holds under which T_m is determined by $rk(T_m) - k'(T_m) = \pi_A^M\left(\underline{\theta}\right) - \pi_A^N(\overline{\theta})$ while T_u is determined by $rk(T_u) - k'(T_u) = \pi_A^M\left(\underline{\theta}\right) - \pi_A^U\left(\overline{\theta}\right)$. Note that $\mu_U^M > \overline{\mu}$ means that $\pi_A^U\left(\overline{\theta}\right)$ is calculated given $q_B^A = q_C^A = 0$. Because $d\{\pi_A^U\left(\overline{\theta}\right) - \pi_A^N(\overline{\theta})\}/d\overline{\theta} = -\left(491a - 1737\overline{\theta} + 1246\underline{\theta}\right)/800 < -\left(491a - 1737\overline{\lambda} + 1246\underline{\theta}\right)/800 = -17\left(a-\underline{\theta}\right)/56 < 0$ holds, we have $\{\pi_A^M\left(\underline{\theta}\right) - \pi_A^N(\overline{\theta})\} - \{\pi_A^M\left(\underline{\theta}\right) - \pi_A^U\left(\overline{\theta}\right)\} = \pi_A^U\left(\overline{\theta}\right) - \pi_A^N(\overline{\theta}) > \pi_A^U\left(\overline{\lambda}\right) - \pi_A^N(\overline{\lambda}) = \left(a-\underline{\theta}\right)^2/196 > 0$. As a result, $T_u < T_m$ holds in this case. Because multilateral free trade is realized at $t = T_m$ on the MTA path and at $t = T_u$ on the CU path, the CU path accelerates the timing of technology adoption and the realization of multilateral free trade.

Let us next compare the equilibrium timing of technology adoption and of attaining multilateral free trade on the FTA path with those of the MTA path. (i) Suppose $\underline{\theta} < \overline{\theta} < \mu_F^M$ holds under which multilateral free trade is realized at $t = 0$ on the MTA path and at $t = \varepsilon$ on the FTA path. Hence, the FTA path delays the realization of multilateral free trade. The timing of technology adoption is determined by $rk(T_m)-k'(T_m)=rk(T_f)-k'(T_f)=\pi_A^M(\underline{\theta})-\pi_A^M(\overline{\theta})$, which means $T_m=T_f$. (ii) Suppose $\mu_F^M \le \overline{\theta} < \lambda_N^M$ holds under which multilateral free trade is attained at $t=0$ on the MTA path and at $t=T_f$ on the FTA path. In this case, T_m is determined by $rk(T_m)-k'(T_m)=\pi_A^M(\underline{\theta})-\pi_A^M(\overline{\theta})$ while T_f is determined by $rk(T_f)-k'(T_f)=\pi_A^M(\underline{\theta})-\pi_A^F(\overline{\theta})$. Since $d\{\pi_A^F(\overline{\theta})-\pi_A^M(\overline{\theta})\}/d\overline{\theta}=(2353a-13\,331\overline{\theta}+10\,978\underline{\theta})/9800>2197\,(a-\underline{\theta})\,/27195>0$ holds, we have $\{\pi_A^M(\underline{\theta})-\pi_A^M(\overline{\theta})\}-\{\pi_A^M(\underline{\theta})-\pi_A^F(\overline{\theta})\}=\pi_A^F(\overline{\theta})-\pi_A^M(\overline{\theta})>\pi_A^F(\mu_F^M)-\pi_A^M(\mu_F^M)>\pi_A^F(\underline{\theta})-\pi_A^M(\underline{\theta})=261\,(a-\underline{\theta})^2\,/19600>0$. This means that $T_m<T_f$ holds. (iii) Suppose $\lambda_N^M \le \overline{\theta} < \overline{\lambda}$ under which multilateral free trade is realized at $t=T_m$ on the MTA path and at $t=T_f$ on the FTA path. The timing of technology adoption on the MTA path and on the FTA path are respectively determined by $rk(T_m)-k'(T_m)=\pi_A^M(\underline{\theta})-\pi_A^N(\overline{\theta})$ and $rk(T_f)-k'(T_f)=\pi_A^M(\underline{\theta})-\pi_A^F(\overline{\theta})$. Because $\pi_A^M(\underline{\theta})-\pi_A^N(\overline{\theta})-\{\pi_A^M(\underline{\theta})-\pi_A^F(\overline{\theta})\}=\pi_A^F(\overline{\theta})-\pi_A^N(\overline{\theta})=(17a-79\overline{\theta}+62\underline{\theta})\{3\left(a-\overline{\theta}\right)+22\left(\overline{\theta}-\underline{\theta}\right)\}/2450>4(a-\underline{\theta})\{3\left(a-\overline{\theta}\right)+22\left(\overline{\theta}-\underline{\theta}\right)\}/1715>0$ holds, we have $T_m<T_f$ in this case.

In sum, compared with one-shot trade liberalization on the MTA path, two-step liberalization on the CU path may accelerate the technology adoption and the realization of multilateral free trade. However, two-step liberalization on the FTA path never accelerates the technology adoption and always delays the realization of multilateral free trade. ■

The discriminatory tariffs under no agreements

If we allow countries to impose discriminatory tariffs, the tariff vector under no agreements is given by $\boldsymbol{\tau}^N=(\tau_A^B,\tau_A^C;\tau_B^A,\tau_B^C;\tau_C^A,\tau_C^B)$. The optimal tariffs set by Country i to Country j $(j\neq i)$ is given by $\widehat{\tau}_i^{j,N}(\theta_A):=\arg\max_{\tau_i} W_i\left(\theta_A,\boldsymbol{\tau}^N\right)$. We can confirm that $\widehat{\tau}_i^{j,N}(\underline{\theta})=\widehat{\tau}_i^N(\underline{\theta})$, which means that each country levies the same level of tariff as the MFN tariff after technology adoption. Before technology adoption, on the other hand, we have $\widehat{\tau}_A^{B,N}(\overline{\theta})=\widehat{\tau}_A^{C,N}(\overline{\theta})=\widetilde{\tau}_A^N(\overline{\theta})$, $\widehat{\tau}_B^{A,N}(\overline{\theta})=\widehat{\tau}_C^{A,N}(\overline{\theta})$, $\widehat{\tau}_B^{C,N}(\overline{\theta})=\widehat{\tau}_C^{B,N}(\overline{\theta})$, $\widehat{\tau}_B^{A,N}(\overline{\theta})-\widetilde{\tau}_B^N(\overline{\theta})=\widehat{\tau}_C^{A,N}(\overline{\theta})-\widetilde{\tau}_C^N(\overline{\theta})=-\left(\overline{\theta}-\underline{\theta}\right)/4<0$ and $\widehat{\tau}_B^{C,N}(\overline{\theta})-\widetilde{\tau}_B^N(\overline{\theta})=\widehat{\tau}_C^{B,N}(\overline{\theta})-\widetilde{\tau}_C^N(\overline{\theta})=\left(\overline{\theta}-\underline{\theta}\right)/4>0$. Hence, we have $\widehat{\tau}_B^{C,N}(\overline{\theta})=\widehat{\tau}_C^{B,N}(\overline{\theta})>\widetilde{\tau}_B^N(\overline{\theta})=\widetilde{\tau}_C^N(\overline{\theta})>\widehat{\tau}_B^{A,N}(\overline{\theta})=\widehat{\tau}_C^{A,N}(\overline{\theta})$. Because $\partial\pi_i(\theta_A,\boldsymbol{\tau}^N)/\partial\tau_j^i<0$ $(i\neq j)$ and $\partial\pi_i(\theta_A,\boldsymbol{\tau}^N)/\partial\tau_j^k>0$ $(i\neq j,\, i\neq k)$ hold, $\pi_A^N(\overline{\theta})$ becomes larger if we allow the discriminatory tariffs, while $\pi_A^N(\underline{\theta})$ remains unchanged.

After the formation of the North-South FTA or the North-South CU, the tariff vectors become $\boldsymbol{\tau}^F=(0,\tau_A;0,\tau_B;\tau_C^A,\tau_C^B)$ and $\boldsymbol{\tau}^U=(0,\overline{\tau};0,\overline{\tau};\tau_C^A,\tau_C^B)$ respectively. It can be easily verify that Countries A and B impose the same external tariffs as the MFN tariff case because they choose only tariffs against Country C in both cases. In addition, Country C's optimal tariffs are unchanged by the formation of the North-North FTA or of the North-North CU. Because $\partial\pi_A(\overline{\theta},\boldsymbol{\tau}^F)/\partial\tau_C^A=\partial\pi_A(\overline{\theta},\boldsymbol{\tau}^U)/\partial\tau_C^A<0$, $\widetilde{\tau}_C^N(\overline{\theta})>\widehat{\tau}_C^{A,N}(\overline{\theta})$ and $\widetilde{\tau}_C^N(\underline{\theta})>\widehat{\tau}_C^{A,N}(\underline{\theta})$ mean that both $\pi_A^F(\overline{\theta})$ and $\pi_A^U(\overline{\theta})$ are higher if we allow the discriminatory tariffs, while $\pi_A^F(\underline{\theta})$ and $\pi_A^U(\underline{\theta})$ remain unchanged.

References

[1] Aghion, P., P. Antras, and E. Helpman (2007) "Negotiating Free Trade", *Journal of International Economics* 73(1), pp.1–30.

[2] Bagwell, K. and R.W. Staiger (1999) "Regionalism and Multilateral Tariff Cooperation", in J. Pigott and A. Woodland (eds.), *International Trade Policy and the Pacific Rim,* London: Macmillan, pp.157-185.

[3] Bhagwati, J. (1991) *The World Trading System at Risk*, Princeton, NJ: Princeton University Press.

[4] Bhagwati, J. (1993) "Regionalism and Multilateralism: An Overview," in J. de Melo and A. Panagariya (eds.), *New Dimensions in Regional Integration*. Cambridge: Cambridge University Press pp.22–51.

[5] Bond, E.W., R.G. Riezman, and C. Syropoulos (2004). "A Strategic and Welfare Theoretic Analysis of Free Trade Areas", *Journal of International Economics* 64(1), pp.1–27.

[6] Bond, E.W., C. Syropoulos, and L.A. Winters (2001) "Deepening of Regional Integration and Multilateral Trade Agreements", *Journal of International Economics* 53(2), pp.335–361.

[7] Cabrales, A. and M. Motta (2001) "Country Asymmetries, Endogenous Product Choice and the Timing of Trade Liberalization", *European Economic Review* 45(1), pp.87–107.

[8] Calvo-Pardo, H., Freund, C. and Ornelas, E (2009) "The ASEAN Free Trade Agreement: Impact on Trade Flows and External Trade Barriers", CEP Discussion Papers No.0930.

[9] Choi, J.P. (1995) "Optimal Tariffs and the Choice of Technology: Discriminatory Tariffs vs. the 'Most Favored Nation' Clause", *Journal of International Economics* 38(1–2), pp.143–160.

[10] Crowley, M.A. (2006) "Do Safeguard Tariffs and Antidumping Duties Open or Close Technology Gaps?", *Journal of International Economics* 68, pp.469–484.

[11] Ederington, J. and P. McCalman (2008) "Endogenous Firm Heterogeneity and the Dynamics of Trade Liberalization", *Journal of International Economics* 74(2), pp.422–440.

[12] Ederington, J. and P. McCalman (2009) "International Trade and Industrial Dynamics", *International Economic Review* 50(3), pp.961–989.

[13] Estevadeordal, A., C. Freund, and E. Orneals (2008) "Does Regionalism Affect Trade Liberalization Toward Nonmembers?", *Quarterly Journal of Economics* 123(4), pp.1531–1575.

[14] Fernandes, A.M. (2007) “Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries”, *Journal of International Economics* 71(1), pp.52–71.

[15] Freund, C. (2000) “Different Paths to Free Trade: The Gains from Regionalism”, *Quarterly Journal of Economics* 115(4), pp.1317–1341.

[16] Fudenberg, D. and J. Tirole (1985) “Preemption and Rent Equalization in the Adoption of New Technology”, *Review of Economic Studies* 52(3), pp.383–401.

[17] Furusawa, T. and H. Konishi (2007). “Free Trade Networks”, *Journal of International Economics* 72(2), pp.310–335.

[18] Gatsios, K. (1990) “Preferential Tariffs and the ‘Most Favored Nation’ Principle: A Note”, *Journal of International Economics* 28(3–4), pp.365–373.

[19] Hwang, H. and C.-C. Mai (1991) “Optimum Discriminatory Tariffs under Oligopolistic Competition”, *Canadian Journal of Economics* 24(3), pp.693–702.

[20] Keller, W. (2004) “International Technology Diffusion”, *Journal of Economic Literature* 42(3), pp.752–782.

[21] Kennan, J. and R. Riezman (1990) “Optimal Tariff Equilibria with Customs Unions”, *Canadian Journal of Economics* 23(1), pp.70–83.

[22] Krishna, P. (1998) “Regionalism and Multilateralism: A Political Economy Approach”, *Quarterly Journal of Economics* 113(1), pp.227–250.

[23] Levy, P.I. (1997) “A Political-Economic Analysis of Free-Trade Agreements”, *American Economic Review* 87(4), pp.506–519.

[24] Liao, C.-H. (2006) ”Heterogeneity of North-South Sectorial Intra-industry Trade: A Demand-Side Approach”, mimeo.

[25] López-Acevedo, G. (2002) ”Determinants of Technology Adoption in Mexico”, Policy Research working paper series, WPS 2780, The World Bank.

[26] Matsuyama, K. (1990) “Perfect Equilibria in a Trade Liberalization Game”, *American Economic Review* 80(3), pp.480–492.

[27] Miyagiwa, K. and Y. Ohno (1995) “Closing the Technology Gap Under Protection”, *American Economic Review* 85(4), pp.755–770.

[28] Miyagiwa, K. and Y. Ohno (1999) "Credibility of Protection and Incentives to Innovate", *International Economic Review* 40(1), pp.143–163.

[29] Mukunoki, H. (2004) "On the Optimal Tariff of a Free Trade Area with Internal Market Integration", *Japan and the World Economy* 16(4), pp.431–448.

[30] Mukunoki, H. and K. Tachi (2006) "Multilateralism and Hub-and-Spoke Bilateralism", *Review of International Economics* 14(4), pp.658–674.

[31] Ornelas, E. (2005a) "Trade Creating Free Trade Areas and the Undermining of Multilateralism", *European Economic Review* 49(7), pp.1717–1735.

[32] Ornelas, E. (2005b) "Endogenous Free Trade Agreements and the Multilateral Trading System", *Journal of International Economics* 67(2), pp.471–497".

[33] Reinganum, J. (1981) "On the Diffusion of New Technology: A Game Theoretic Approach", *Review of Economic Studies* 48(3), pp.395–405.

[34] Saggi, K. (2006) "Preferential Trade Agreements and Multilateral Tariff Cooperation", *International Economic Review* 47(1), pp.29–57.

[35] Saggi, K. and H.M. Yildiz (2010) "Bilateralism, Multilateralism, and the Quest for Global Free Trade", *Journal of International Economics* 81(1), pp.26–37.

[36] Srinivasan, T.N. (1998) "Regionalism and the WTO: Is Nondiscrimination Passé?", in A.O. Krueger (ed.), *The WTO as an International Organization*, Chicago: University of Chicago Press, pp.330–345.

[37] Teshima, K. (2009) "Import Competition and Innovation at the Plant Level: Evidence from Mexico", mimeo.

[38] Tharakan, P.K.M. and Kerstens, B. (1995) "Does North-South Horizontal Intra-Industry Trade Really Exist?: An Analysis of the Toy Industry", *Review of World Economics (Weltwirtschaftliches Archiv)* 131(1), pp.86–105.

[39] UNCTAD (2007) *Trade and Development Report 2007*, United Nations: New York and Geneva.

[40] Utar, H. and L.B.T. Ruiz (2010) "The Impact of Chinese Competition on Mexican Maquiladoras: Evidence from Plant-level Panel Data", mimeo.

[41] Yeaple, S. (2005) "A Simple Model of Firm Heterogeneity, International Trade and Wages", *Journal of International Economics* 65(1), pp.1–20.

[42] Yi, S.-S. (2000) "Free-Trade Areas and Welfare: An Equilibrium Analysis", *Review of International Economics* 8(2), pp.336–347.

Tables and Figures

Table 1: The timing of technology adoption and trade agreements

(Country A: The member of PTAs)

	The timing of	The timing of the MTA			The timing of the PTA	
	technology adoption	MTA path	FTA path	CU path	FTA path	CU path
$\overline{\theta} < \lambda_F^M$	$T_m = T_f = T_u$	0	ε	ε	0	0
$\lambda_F^M \leq \overline{\theta} < \lambda_U^M$	$T_m = T_u < T_f$	0	T_f	ε	0	0
$\lambda_U^M \leq \overline{\theta} < \lambda_N^M$	$T_m < T_f < T_u$	0	T_f	T_u	0	0
$\lambda_N^M \leq \overline{\theta} < \lambda_N^U$	$T_f < T_m < T_u$	T_m	T_f	T_u	0	0
$\lambda_N^U \leq \overline{\theta} < \lambda_N^F$	$T_f < T_m$ and $T_u < T_m$	T_m	T_f	$T_u + \varepsilon$	0	T_u
$\lambda_N^F \leq \overline{\theta}$	$T_u < T_m < T_f$	T_m	$T_f + \varepsilon$	$T_u + \varepsilon$	T_f	T_u

Table 2: The timing of technology adoption and trade agreements

(Country A: The nonmember of PTAs)

	The timing of	The timing of the MTA			The timing of the PTA	
	technology adoption	MTA path	FTA path	CU path	FTA path	CU path
$\overline{\theta} < \mu_F^M$	$T_m = T_f = T_u$	0	ε	ε	0	0
$\mu_F^M \leq \overline{\theta} < \lambda_N^M$	$T_m = T_u < T_f$	0	T_f	ε	0	0
$\lambda_N^M \leq \overline{\theta} < \mu_U^M$	$T_u < T_m < T_f$	T_m	T_f	ε	0	0
$\mu_U^M \leq \overline{\theta}$	$T_u < T_m < T_f$	T_m	T_f	T_u	0	0

Figure 1: The technology gap and the equilibrium timing of trade agreements

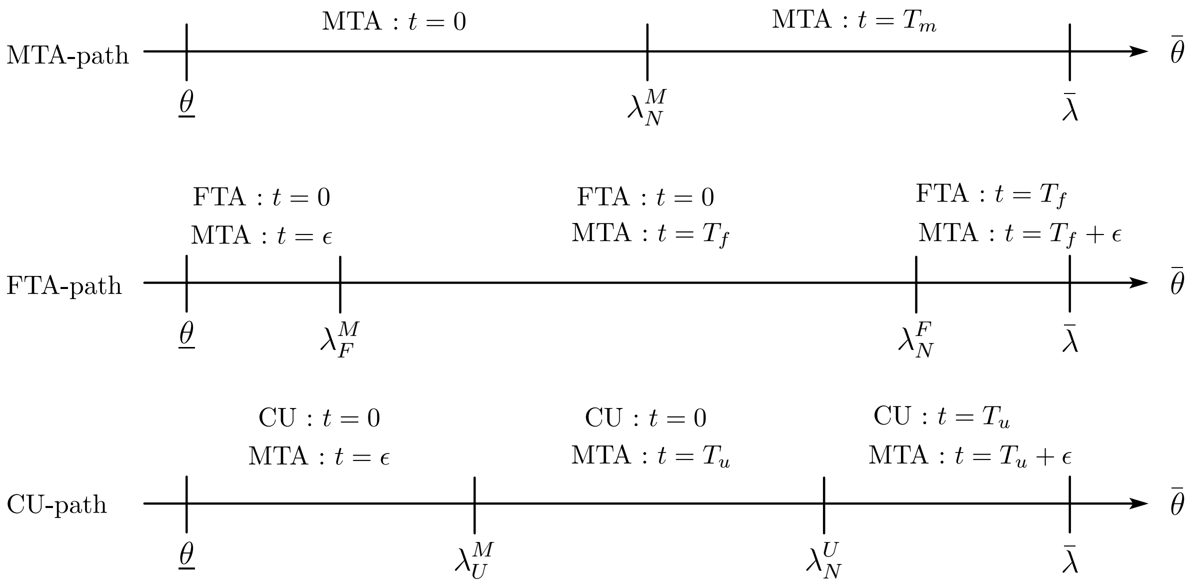

Figure 2: The optimal timing of technology adoption ($\lambda_N^M \leq \overline{\theta} < \lambda_N^U$)

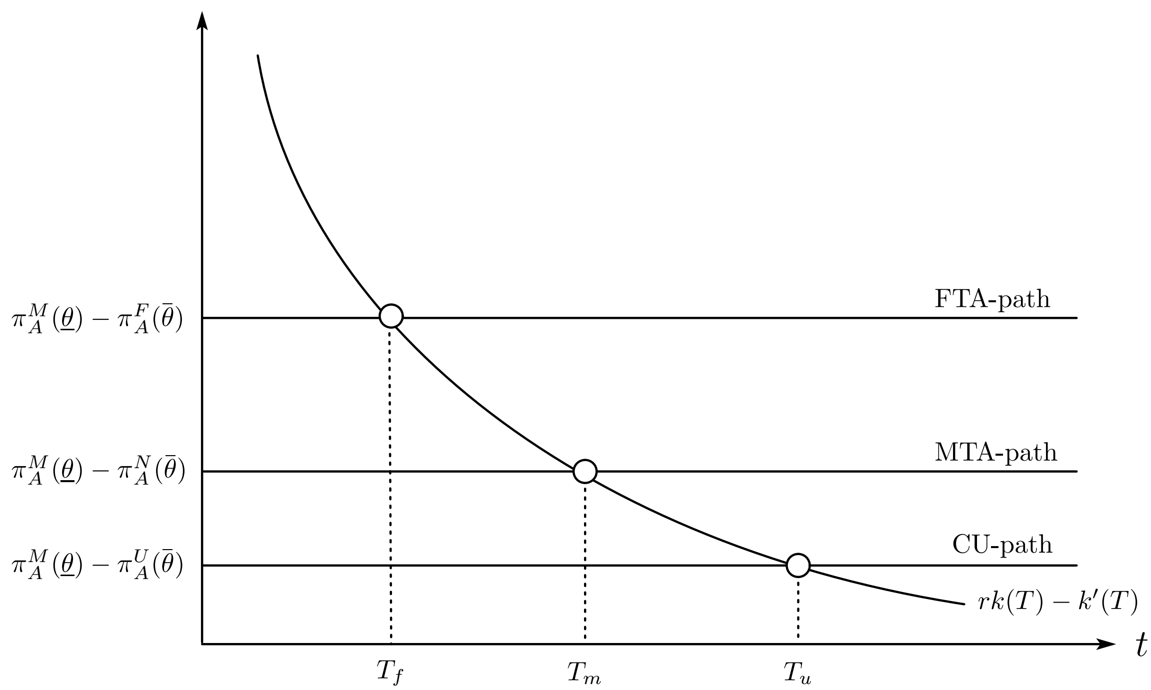

Figure 3: Trade agreements and welfare changes ($\lambda_N^M \leq \overline{\theta} < \lambda_N^U$)

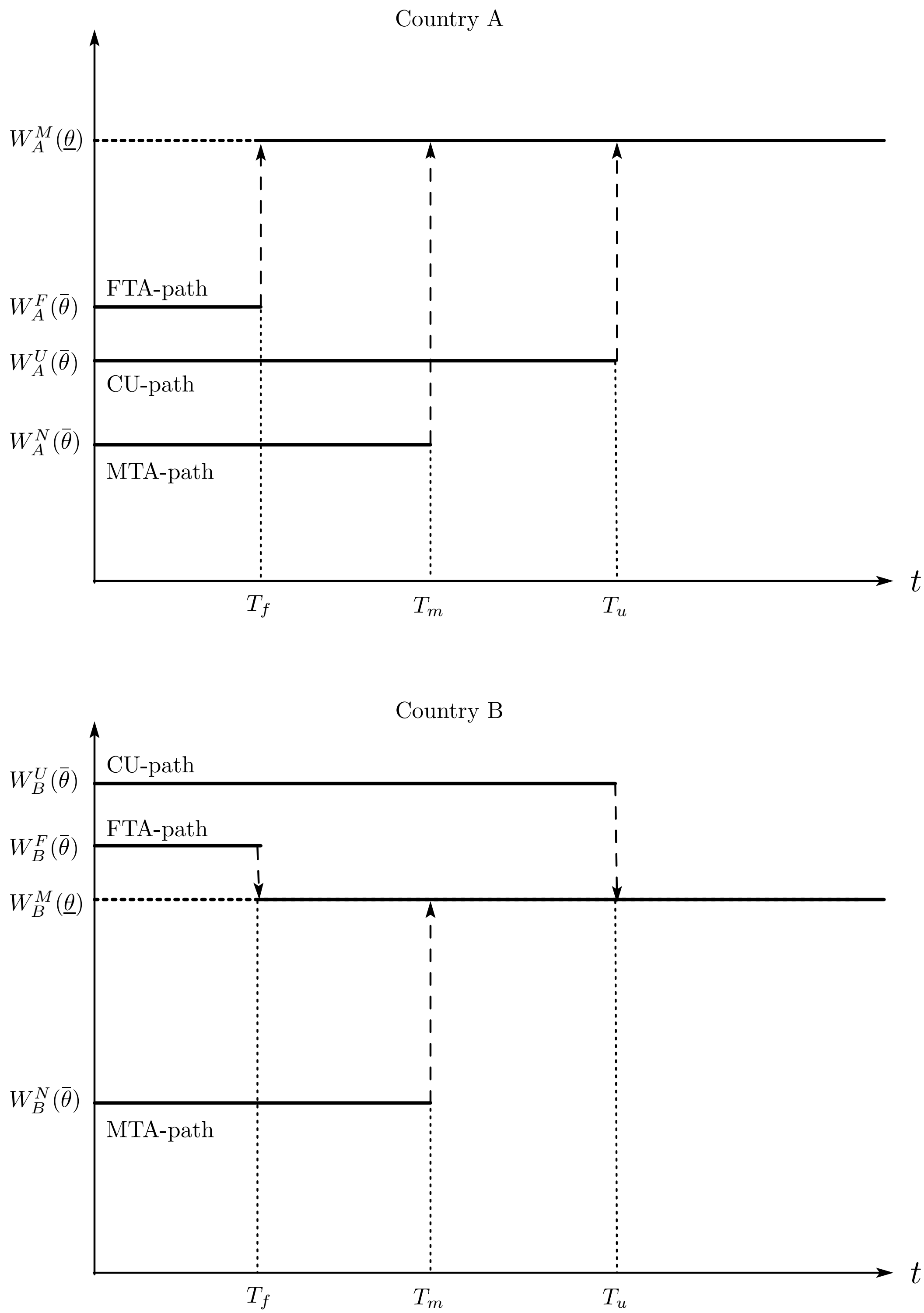

Figure 3 (contiuned)

Country C