Discussion Paper Series

RIEB

Kobe University

DP2020-22

Endogenous Timing in Tax Competition: The Effect of Asymmetric Information*

RIEB Junior Research Fellow

Takaaki HAMADA

Revised September 5, 2022

*This Discussion Paper won the Kanematsu Prize (FY 2019).

Research Institute for Economics and Business Administration

Kobe University

2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Endogenous Timing in Tax Competition: The Effect of Asymmetric Information*

Takaaki Hamada †

September 5, 2022

Abstract

This study explores the effects of asymmetric information on endogenous leadership in a simple tax competition environment (Ogawa, 2013). The study models a two-country economy where one country is informed about its own and opponent's productivity of private goods, while the other country knows only its own productivity. The results show that each type of informed country has an incentive to pretend to be the other type, which leads to a Stackelberg outcome endogenously, while the simultaneous move is the unique outcome under complete information. Under the Stackelberg outcome, the uninformed country moves first and the informed country moves second. Moreover, ex-post social welfare under asymmetric information can become larger than that under complete information, because the uninformed country chooses a less aggressive tax rate under asymmetric information. These results depend on the type of uncertainty, and capital ownership and share.

Keywords: Tax Competition, Endogenous Leadership, Asymmetric Information, Pooling Equilibrium, Welfare Improvement

JEL classification: D82, H30, H87

*The previous version of this paper (entitled "Welfare-Improving Asymmetric Information in an Endogenous Timing Game of Tax Competition") won the Kanematsu Prize from the Research Institute for Economics and Business Administration (RIEB) at the Kobe University in 2019. I am grateful to the three anonymous referees of the Kanematsu Prize for carefully reviewing the paper and providing useful comments. I am also thankful for the valuable comments of Akihiko Matsui, Akira Ogawa, Daiki Kishishita, Hikaru Ogawa, Susumu Sato and the participants of the 2018 Spring Meeting of the Japanese Economic Association at the University of Hyogo. This project was financially supported by JSPS Grand-in-Aid for JSPS Research Fellows (17J07819). All remaining errors are my own.

†Faculty of Economics, Toyo University, 5-28-20, Hakusan, Bunkyo-ku, Tokyo 112-8606 Japan/Research Institute for Economics and Business Administration, Kobe University, 2-1 Rokkodai, Nada, Kobe 657-8501, Japan. Email: t.hamada.econ@gmail.com

1 Introduction

Recently, there have been studies on the endogenous timing in tax competition models. In the simple tax competition environment and Hamilton and Slutsky (1990)'s observable delay manner, the equilibrium outcomes are characterized by production technology, capital ownership, and the distribution of the capital across regions (Kempf and Rota-Graziosi, 2010, 2015; Ogawa, 2013; Hindriks and Nishimura, 2017).

In a two-country economy, assuming the absence of capital ownership, Kempf and Rota-Graziosi (2010) conclude that the subgame perfect equilibrium (SPE) yields two sequential-move outcomes. Ogawa (2013) shows that the simultaneous move is a unique SPE outcome in economies with non-absentee capital ownership and an equal capital share between two countries. Moreover, Kempf and Rota-Graziosi (2015) characterize the equilibrium outcome in the non-absentee capital ownership model, which allows an unequal share of the capital, and Hindriks and Nishimura (2017) develop more flexible capital ownership models. These studies have adequately addressed the timing decisions of fiscal authorities in the complete information environments.

This study provides a new perspective on the endogenous leadership in a simple tax competition by introducing asymmetric information. Using Ogawa (2013)'s setting, the study analyzes information asymmetry regarding a country's productivity level of private goods in a two-country economy, where one country is informed about its own and opponent's productivity, while the other country knows only its own productivity. The results show that a Stackelberg outcome emerges endogenously, while the simultaneous move is the unique outcome under complete information. Moreover, we find that ex-post social welfare under asymmetric information can become larger than that under complete information.

In Ogawa (2013)'s setting, information asymmetry can provide each type of the informed country with a novel incentive of pretending to be the other type, which leads to a pooling equilibrium. The mechanism behind that is as follows. In Ogawa (2013), as the capital is fully owned by the residents in the two countries and their productivity levels are heterogeneous, one country becomes a capital importer and the other becomes an exporter. The capital importer strategically chooses its tax rate to decrease the price of capital, while the exporter chooses a tax rate to increase it. Now, suppose the uninformed country does not know whether its opponent has higher or lower productivity than itself, which implies that the country is unaware of whether it is a capital importer or exporter. The informed country with higher productivity (henceforth, high type) wants the uninformed country to choose a tax rate that decreases the capital price, because it is a capital importer given its higher productivity. If the uninformed country incorrectly expects itself to be a capital importer (i.e., to be more productive than the opponent), then it will lower the capital price, which will actually benefit the high type. In this situation, the high type has no incentive to resolve the uncertainty. Similarly, the low type also has an incentive to use the uncertainty such that the uninformed country is misled into increasing the capital price.

These two-way concealment incentives can make each type choose the same choice of timing and the same tax rate such that information on its productivity level is not given to the uninformed country. Such incentives consequently lead to a Stackelberg outcome endogenously in which the uninformed country moves first and both types of the informed country move second, while the simultaneous move is the unique outcome under complete information.

These concealment incentives depend on the assumption of the type of uncertainty, and capital ownership and share. First, it is essential that the uninformed country does not know whether its opponent has higher or lower productivity than itself. Unlike this assumption, we can consider a different type of uncertainty such that the uninformed country knows that its opponent has higher or lower productivity than itself but does not understand the degree of the gap. In such a case, the uninformed country is aware of its actual role in the capital trade, and either type of the informed country wants to differentiate itself from the other type, contributing to a separating equilibrium. Second, in the case of no-capital ownership (Kempf and Rota-Graziosi, 2010) or in some cases of unequal capital share across countries (Kempf and Rota-Graziosi, 2015), two-way concealment incentives disappear and the pooling equilibrium cannot emerge. The results in this study can be brought when the incentives for the manipulation of the capital price differ between the two types.[1]

Some examples of information asymmetry in this study include some exogenous events, such as productivity shocks or disasters. After a country experiences an exogenous shock, it may hold more information on itself than others, even though most of the information had been gathered until then. Theoretical studies incorporating the risk of disasters include Wildasin (2011) and Goodspeed and Haughwout (2012).[2] Especially, Wildasin (2011) considers a similar situation with two regions: one region (Coast) is exposed to the risk of disasters, while the other (Inland) is not exposed. Unlike

[1] In other words, one type wants to lower the capital price, while the other wants to raise the price.

[2] See the survey of Goodspeed (2013).

this study, they do not focus on the information asymmetry between jurisdictions.

The aforementioned examples may seem limited to a specific scenario. Nevertheless, this model is worth analyzing for several reasons. First, this model suggests that the informed country is willing to use the uncertainty to pretend to be the other type. This implies that each country may have an incentive to "create" uncertainty by manipulating or concealing information.[3] Thus, our environment may be realized not only through some exogenous shocks but also through countries' decisions endogenously. Moreover, this study shows that such incentives significantly change the equilibrium outcome compared with that under complete information, which suggests that information asymmetry may also have non-negligible effects in related works. The driving mechanism in this study is helpful to introduce information asymmetry to others' studies.

Second, the study finds that social welfare in the equilibrium under asymmetric information can become larger than that under complete information, where social welfare is defined as the sum of ex-post utilities. The positive effect of information asymmetry on welfare is also shown in the analysis with a fixed timing structure. This occurs because the uncertainty makes the uninformed country choose less aggressive tax rates toward both types. This finding contributes to uncovering the mechanisms driving the positive effects of incomplete information on social welfare,[4] and, to my best knowledge, this is the first study to show welfare improvement by asymmetric information in endogenous timing models. However, the study does not detect a Pareto improvement by asymmetric information. Under our setup, while each type of the informed country can improve its utility by pretending to be the other type, the uninformed country suffers a loss because such a country may choose a tax rate with misperceptions of the actual flows of capital.

Finally, our model presents some interesting points from the game theoretical aspect. The first point is about the signaling game. A signaling game in our model with continuous action space (tax choice) does not satisfy the commonly used single-crossing property, but holds the double crossing. In the game, we find that there uniquely exists a pooling equilibrium (or hybrid equilibrium) outcome (Lemma 1). Interestingly, the study does not use refinement criteria to exclude any outcomes. Generally, in the games with double-crossing property, multiple outcomes emerge

[3]This has an opposite implication for the information sharing among governments (Bacchetta and Espinosa, 1995, 2000; Huizinga and Nielsen, 2003).

[4]For example, the notable work by Morris and Shin (2002) shows that more (public) information can decrease welfare in game theoretical situations. In other examples, in the insurance market models with asymmetric information, de GaridelThoron (2005) finds that sharing information about past accidents among insurers decreases welfare in dynamic environments. Koufopoulos and Kozhan (2014) show that an increase in ambiguity about the probability of an accident can lead to a Pareto improvement. The mechanism behind welfare improvement in our model is different from that in these studies.

in perfect Bayesian equilibrium (PBE), and some unstable outcomes are often eliminated using the refinement criteria (Kolev and Prusa, 1999, 2002; Daley and Green, 2014; Chen, Ishida and Suen, 2021). The uniqueness of our model lies in the two-way concealment incentives, which are characterized by opposing directions of each type's indifference curve for higher utility (Figure 2). Proofs for non-equilibrium outcomes are based on this property.

The second point is about the endogenous timing game with asymmetric information. Although the literature on tax competition lacks studies on endogenous leadership with asymmetric information, there exist some studies on the duopoly models. For example, Mailath (1993) and Normann (1997) analyze the effect of asymmetric information using Hamilton and Slutsky (1990)'s action commitment manner, and Normann (2002) uses Hamilton and Slutsky (1990)'s observable delay manner. In terms of the methodology, this paper is more related to Normann (2002)'s study. In these studies, a privately informed player may have a first-mover disadvantage owing to the over-investments often observed in signaling games. This is attributed to players' incentives for differentiating themselves from others: the main focus is separating equilibria. However, our central mechanism lies in pooling equilibria, which suggests a different mechanism for the second-mover advantage from that discussed in existing studies.

This paper is organized as follows. Section 2 describes the endogenous timing model with asymmetric information, based on Ogawa (2013). Section 3 presents the equilibrium analysis, and Section 4 analyzes the effect of information asymmetry on social welfare. Section 5 discusses the assumptions of the type of uncertainty, and capital ownership and share. Section 6 concludes the study.

2 The Model

Our basic setup follows Ogawa (2013). In a two-country economy, each country $i \in \{1, 2\}$ has homogeneous residents normalized by 1. A homogeneous private good is locally produced using capital and labor. The total capital used for the production in this economy is $2\bar{k}$. We assume that, in each country, a resident is endowed with capital $\bar{k}$, which implies a non-absentee capital ownership environment. We assume that the labor supply is fixed and immobile, and that capital is perfectly mobile.

We assume the CRS technology in each country, and specify the production per capita in country

i as $f_i(k_i) = (A_i - k_i)k_i$, where k_i is the capital per capita used for the production in country i, and A_i represents the productivity level of country i.

Country 1's productivity A_1 is either A^H or A^L ($A^H > A^L > 0$). The productivity level is determined by nature, which randomly assigns type $s \in \{H, L\}$ to country 1, following the probability distribution $\overline{\rho} = (\overline{\rho}_H, \overline{\rho}_L)$, where $\overline{\rho}_s$ is the probability of country 1 being type s and $\overline{\rho}_H + \overline{\rho}_L = 1$. The country 1 assigned type s is often called "type s." On country 2's productivity level, we assume the following.

Assumption 1. *Country 2's productivity level satisfies* $A_2 \in (A^L, A^H)$.

Under Assumption 1, country 2 does not know whether its role in the capital trade is as an importer or an exporter, which induces country 1's concealment incentive. This is the essential assumption for our results. Moreover, we assume the following for an analysis of information asymmetry.

Assumption 2. *Country 1 can observe both A_1 and A_2, but country 2 can only observe its own productivity A_2. This fact and the probability distribution $\overline{\rho}$ are common knowledge among countries.*

Let $\Lambda^s \equiv A^s - A_2$, and its expected value with prior distribution be $\overline{\Lambda}$ ($= \overline{\rho}_H \Lambda^H + \overline{\rho}_L \Lambda^L$). Under Assumption 1, $\Lambda^H > 0 > \Lambda^L$ and $\overline{\Lambda} \in [\Lambda^L, \Lambda^H]$ hold.

The resident's utility in country i is represented by $u_i(c_i) = c_i$, where c_i is the consumption of a private numeraire good. The resident in country i receives the marginal productivity of labor $f_i(k_i) - k_i f_i'(k_i)$ as labor income and rent from capital $r\overline{k}$, where r is the economy-wide net capital return. The lump-sum transfer or tax by the government in country i is denoted by g_i.[5] Then, the consumption c_i is represented by

$$c_i = f_i(k_i) - k_i f_i'(k_i) + r\overline{k} + g_i. \tag{1}$$

The government in each county i can impose a unit tax or negative unit tax (subsidy) on mobile capital, denoted as t_i. Then, the government's budget constraint is

$$g_i = t_i k_i. \tag{2}$$

By the assumption that capital is perfectly mobile and total capital is $2\overline{k}$, the market clearing

[5] g_i is the lump-sum transfer when $g_i > 0$, and it is the lump-sum tax when $g_i < 0$.

yields

$$f_1'(k_1) - t_1 = r = f_2'(k_2) - t_2, \tag{3}$$

$$k_1 + k_2 = 2\overline{k}. \tag{4}$$

By (1) − (3), the resident's utility can be rewritten as

$$u_i = f_i(k_i) + r(\overline{k} - k_i). \tag{5}$$

Moreover, (3) and (4) imply that

$$k_1^s = \overline{k} + \frac{1}{4}(\Lambda^s - t_1^s + t_2), \tag{6}$$

$$k_2^s = \overline{k} - \frac{1}{4}(\Lambda^s - t_1^s + t_2), \tag{7}$$

$$r^s = \frac{\Omega^s}{2} - 2\overline{k} - \frac{1}{2}(t_1^s + t_2), \tag{8}$$

where k_i^s and t_1^s are country i's production capital and country 1's tax rate when country 1 is type s, respectively, and $\Omega^s \equiv A^s + A_2$.

Country i is a capital exporter when $\overline{k} - k_i > 0$, and a capital importer when $\overline{k} - k_i < 0$. Since no capital comes from outside of the two countries, one country becomes a capital exporter and the other becomes a capital importer. Then, the utility function (5) implies that the two countries always have different incentives to manipulate capital return r. In particular, a capital importer wants to increase its tax rate, and a capital exporter wants to decrease its tax rate, given $\frac{\partial r}{\partial t_i} < 0$ for all i.

3 Equilibrium Analysis

3.1 Timing Game and Equilibrium Concept

We assume that the government of each country is benevolent, so that each government makes decisions to maximize its resident's utility. In the subsequent sections, the government in country i is called "country i." As in Ogawa (2013) and the literature on the leadership of tax competition, we use the two-stage timing game with observable delay introduced by Hamilton and Slutsky (1990). Taking the information asymmetry into account, the timing game is described as follows:

Stage 0 : (i) Nature assigns a type to country 1.

(ii) Only country 1 observes its type.

Stage 1 : (i) Each country simultaneously announces when it sets the tax rate, *e(arly)* or *l(ate)*.

(ii) Both countries observe each other's timing choice.

Stage 2 : Both countries commit to the timing announced in Stage 1 and set their tax rate. A situation emerges from the following:[6]

(a): the situation G_N where both countries set their tax rates simultaneously,

(b): the situation G_1 where country 1 sets the tax rate first and country 2 follows, and

(c): the situation G_2 where country 2 sets the tax rate first and country 1 follows.

Stage 3 : Capital level for production and utility are determined in both countries.

A pure strategy of country i is a pair of action strategies (a_i, t_i), where $a_i : \Theta_i \to \{e, l\}$ is country i's timing choice; $t_1 : \Theta_1 \times \{(e, e), (e, l), (l, e) \times \mathbb{R}, (l, l)\} \to \mathbb{R}$ is country 1's tax choice; and $t_2 : \{(e, e), (e, l) \times \mathbb{R}, (l, e), (l, l)\} \to \mathbb{R}$ is country 2's tax choice with $\Theta_1 = \{H, L\}$ and Θ_2 singleton. Throughout the study, we denote type s's timing and tax choices as a_1^s and t_1^s, respectively. A mixed strategy of country i is denoted by σ_i. In this model, countries may adopt mixed action strategies both in timing and tax choices. Let $p_i : \Theta_i \times \{e, l\} \to [0, 1]$ be country i's mixed action strategy on the timing choice, and $q_i : \Theta_i \times T \to [0, 1]$ be country i's mixed action strategy on the tax choice, with finite support T on $\mathbb{R}$. A mixed strategy σ_i is a pair (p_i, q_i).

Let (ρ^a, ρ^t) be a posterior belief of country 2 (the uninformed country) such that $\rho^a = (\rho_H^a, \rho_L^a)$ with $\rho_H^a + \rho_L^a = 1$, where ρ_s^a is the probability of country 1 being type s after country 1's timing choice $a \in \{e, l\}$ is observed; $\rho^t = (\rho_H^t, \rho_L^t)$ with $\rho_H^t + \rho_L^t = 1$, where ρ_s^t is the probability of country 1 being type s after country 1's tax rate t is observed. We can consider that ρ^a is updated to ρ^t by country 2, because it observes the tax rate t after the timing action a. We assume that country 2 has the same belief ρ^a, regardless of its own timing choices. In the subsequent sections, we often call the notations $\hat{\Lambda}^a$ and $\hat{\Lambda}^t$ "belief," which are the expected values of Λ^s driven by ρ^a and ρ^t, respectively.

The equilibrium concept we use is a perfect Bayesian equilibrium (PBE), defined as follows.[7]

[6]Note that Stage 2 is not a "game," because it is not a proper subgame of the whole timing game owing to the asymmetric information. However, as in Ogawa (2013) or related works, we use the notation "G_N," "G_1" and "G_2" for each situation, for convenience.

[7]The definition is similar to that of Normann (2002).

Definition 1. *A profile* $((\sigma_1, \sigma_2), (\rho^a, \rho^t))$ *is a perfect Bayesian equilibrium (PBE), if the following holds:*

I. Given the belief (ρ^a, ρ^t)*, for all* a_i *chosen with positive probability in* p_i*,*

(i) : a_1^s *maximizes type s's expected payoffs, given* p_1^{-s}*,* p_2*,* q_1*, and* q_2 *for any* $s \in \{H, L\}$*;*

(ii) : a_2 *maximizes country 2's expected payoffs, given* p_1*,* q_1*, and* q_2*;*

II. Given the belief (ρ^a, ρ^t)*, for all* t_i *chosen with positive probability in* q_i*,*

(i) : t_1^s *maximizes type s's expected payoffs, given the realized timing,* q_1^{-s} *and* q_2 *for any* $s \in \{H, L\}$*;*

(ii) : t_2 *maximizes country 2's expected payoffs, given the realized timing and* q_1*.*

III. The belief (ρ^a, ρ^t) *of country 2 must be obtained using Bayes' rule and the equilibrium strategies* (σ_1, σ_2)*, whenever calculated.*

$-s$ is the type other than s, and p_1^{-s} and q_1^{-s} are type $-s$'s mixed action strategies on timing and tax choices, respectively. A PBE does not impose any restrictions on belief of off-the-equilibrium actions. Hence, PBEs are often supported by unrealistic beliefs. To ensure that our equilibrium is not supported by unreasonable beliefs, we apply the equilibrium criterion $D1$ developed by Cho and Kreps (1987), as in the timing games of duopoly (Mailath, 1993; Normann, 1997, 2002). In this game, there are two kinds of off-the-equilibrium paths. One is country 1's timing choice, and the other is country 1's tax choice. We only apply $D1$ to the first one because the signaling game on tax choices has a unique outcome without any refinement (Lemma 1). *A perfect Bayesian equilibrium is said to survive the* $D1$ *criterion* if no one deviates from the equilibrium by imposing $D1$ to any off-the-equilibrium timing choices.[8]

Again, we denote G_N as the situation where both the countries choose their tax rates simultaneously, and G_i as the situation where country $i \in \{1, 2\}$ is a leader and country $j(\neq i)$ is a follower. In the next section, we analyze each situation. In the analysis, as in Ogawa (2013), we assume that $\overline{k}$ is sufficiently large to avoid capital concentration in either country.

[8] D1 is reminded in Appendix E.

3.2 Simultaneous Move (G_N)

When country 2 observes $a_1 = a$ $(= a_2)$, the posterior belief is ρ^a. Then, each type s and country 2's problems are as follows:

$$\max_{t_1^s} (A^s - k_1^s)k_1^s + r^s(\overline{k} - k_1^s), \tag{9}$$

$$\max_{t_2} \sum_{s\in\{H,L\}} \rho_s^a \left[(A_2 - k_2^s)k_2^s + r^s(\overline{k} - k_2^s)\right]. \tag{10}$$

Using (6)(7)(8), both countries' best response functions are

$$t_1^s = \frac{1}{3}(\Lambda^s + t_2), \tag{11}$$

$$t_2 = \frac{1}{3}(-\hat{\Lambda}^a + \hat{t_1^a}), \tag{12}$$

where $\hat{\Lambda}^a \equiv \sum_s \rho_s^a \Lambda^s$ and $\hat{t_1^a} \equiv \sum_s \rho_s^a t_1^s$. By (11)(12), their optimal action strategies are

$$(t_1^s, t_2) = \left(\frac{1}{3}\left(\Lambda^s - \frac{1}{4}\hat{\Lambda}^a\right), -\frac{1}{4}\hat{\Lambda}^a\right). \tag{13}$$

Then, by (6)(7)(8)(13), we have

$$(k_1^s, k_2^s) = \left(\overline{k} + \frac{1}{6}\left(\Lambda^s - \frac{1}{4}\hat{\Lambda}^a\right), \overline{k} - \frac{1}{6}\left(\Lambda^s - \frac{1}{4}\hat{\Lambda}^a\right)\right), \tag{14}$$

$$r^s = \frac{\Omega^s}{2} - 2\overline{k} - \frac{1}{6}(\Lambda^s - \hat{\Lambda}^a). \tag{15}$$

3.3 Sequential Move (G_2)

Country 1's best response is the same as that under the situation G_N. Country 2 maximizes its expected payoff, given (11) and the posterior belief ρ^l. Then, their optimal action strategies are

$$(t_1^s, t_2) = \left(\frac{1}{3}\left(\Lambda^s - \frac{2}{5}\hat{\Lambda}^l\right), -\frac{2}{5}\hat{\Lambda}^l\right). \tag{16}$$

Then, by (6)(7)(8)(16), we have

$$(k_1^s, k_2^s) = \left(\overline{k} + \frac{1}{6}\left(\Lambda^s - \frac{2}{5}\hat{\Lambda}^l\right), \overline{k} - \frac{1}{6}\left(\Lambda^s - \frac{2}{5}\hat{\Lambda}^l\right)\right), \tag{17}$$

$$r^s = \frac{\Omega^s}{2} - 2\overline{k} - \frac{1}{6}\left(\Lambda^s - \frac{8}{5}\hat{\Lambda}^l\right). \tag{18}$$

3.4 Sequential Move (G_1)

Country 2's best response after observing country 1's tax rate t is

$$t_2 = \frac{1}{3}\left(t - \hat{\Lambda}^t\right), \tag{19}$$

where $\hat{\Lambda}^t \equiv \sum_s \rho_s^t \Lambda^s$.[9] Given t, using (6)(7)(8)(19), we have

$$(k_1^s, k_2^s) = \left(\overline{k} + \frac{1}{6}\left(\frac{3}{2}\Lambda^s - t - \frac{1}{2}\hat{\Lambda}^t\right), \overline{k} - \frac{1}{6}\left(\frac{3}{2}\Lambda^s - t - \frac{1}{2}\hat{\Lambda}^t\right)\right) \tag{20}$$

$$r^s = \frac{\Omega^s}{2} - 2\overline{k} - \frac{1}{6}\left(4t - \hat{\Lambda}^t\right). \tag{21}$$

We should consider two cases separately: (i) $a_1^H \neq a_1^L$ and (ii) $a_1^H = a_1^L$.

3.4.1 Case (i): $a_1^H \neq a_1^L$ in an equilibrium

In this case, the type is revealed before Stage 2, because each type's timing choice is separated. Thus, the outcome is same as that in the sequential-move game under complete information, so that we have

$$(t_1^s, t_2) = \left(\frac{2}{5}\Lambda^s, -\frac{1}{5}\Lambda^s\right), \tag{22}$$

$$(k_1^s, k_2^s) = \left(\overline{k} + \frac{1}{6}\left(\frac{3}{5}\Lambda^s\right), \overline{k} - \frac{1}{6}\left(\frac{3}{5}\Lambda^s\right)\right), \tag{23}$$

$$r^s = \frac{\Omega^s}{2} - 2\overline{k} - \frac{1}{6}\left(-\frac{3}{5}\Lambda^s\right). \tag{24}$$

3.4.2 Case (ii): $a_1^H = a_1^L$ in an equilibrium

In this case, country 1's choice of timing does not change country 2's belief, because both types choose the same timing. However, country 2 can update its belief before its tax choice, because country 1 is a leader and country 2 can observe country 1's tax rate before making a choice. Thus, we must analyze a signaling game on tax choices where country 1 is a sender, choosing a tax rate, and country 2 is a receiver with a prior belief $\overline{\Lambda}$, observing country 1's tax rate and subsequently choosing a tax rate. Lemma 1 states that, in the signaling game, either a pooling or

[9] As mentioned before, the belief ρ^e is updated to ρ^t, since country 2 observes country 1's tax rate t after observing the timing choice e.

hybrid equilibrium outcome emerges, depending on the prior belief $\overline{\Lambda}$.[10]

Lemma 1. *Suppose situation G_1 is realized and $a_1^H = a_1^L = e$ holds in an equilibrium. Let $\tilde{\Lambda}^s \equiv 3 \cdot \left(1 - \frac{2\sqrt{5}}{5}\right) \Lambda^s$ for each $s \in \{H, L\}$. Each type's tax choice in situation G_1 is as follows:*

(i) If $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$ holds, then type H adopts a mixed action strategy between $t_1^H = 0$ and $\frac{2}{5}\Lambda^H$, and type L chooses $t_1^L = 0$.

(ii) If $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$ holds, both types choose 0.

(iii) If $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ holds, type L adopts a mixed action strategy between $t_1^L = 0$ and $\frac{2}{5}\Lambda^L$, and type H chooses 0.

Moreover, the equilibrium outcome is unique for each $\overline{\Lambda}$.

Proof. See Appendix B. □

In our environment, since the capital is fully owned by the residents in the two countries, one country becomes a capital exporter ($\overline{k} - k_i > 0$) and the other becomes a capital importer ($\overline{k} - k_i < 0$). While capital flows from the low-productive country into the high-productive one so that the equilibrium condition (3) yields, country 2 is uncertain of the actual capital flow because it does not know whether country 1's productivity is higher ($A^H > A_2$) or lower ($A_2 > A^L$) than its own. If country 2 misperceives country 1's productivity level, it will incorrectly interpret its own role in the capital trade and choose the tax rate that is beneficial for country 1.[11] Thus, each type has an incentive to pretend to be the other type using the uncertainty, which leads to a pooling equilibrium ($t_1^H = t_1^L = 0$).

To understand Lemma 1, let us use each type's indifference curve on t-$\hat{\Lambda}^t$ plane.[12] First, Figure 1 illustrates each type's indifference curve at the equilibrium under complete information in the situation G_1, that is (22).

Type H (L)'s utility improves when its indifference curve moves downwards (upwards), since its utility improves with a decline (increase) in country 2's belief $\hat{\Lambda}^t$ because of the concealment

[10] Generally, a hybrid equilibrium in a signaling game with two types, T_1 and T_2, is an equilibrium where type T_1 chooses an action X with probability 1, and type T_2 mixes the action X chosen by type T_1 with another action Y.

[11] Suppose $A_1 = A^H$. In this case, country 1 is a capital importer, and country 2 is a capital exporter. However, if country 2 considers country 1's productivity to be A^L, then it incorrectly interprets its own role as a capital importer. Then, country 2 will increase the tax rate to decrease the capital price ($\frac{\partial r}{\partial t_i} < 0$). In reality, since country 1 is a real importer, this manipulation of the capital price will benefit country 1.

[12] For the detail of the indifference curves, see the explanation right after Lemma 3 in Appendix B.

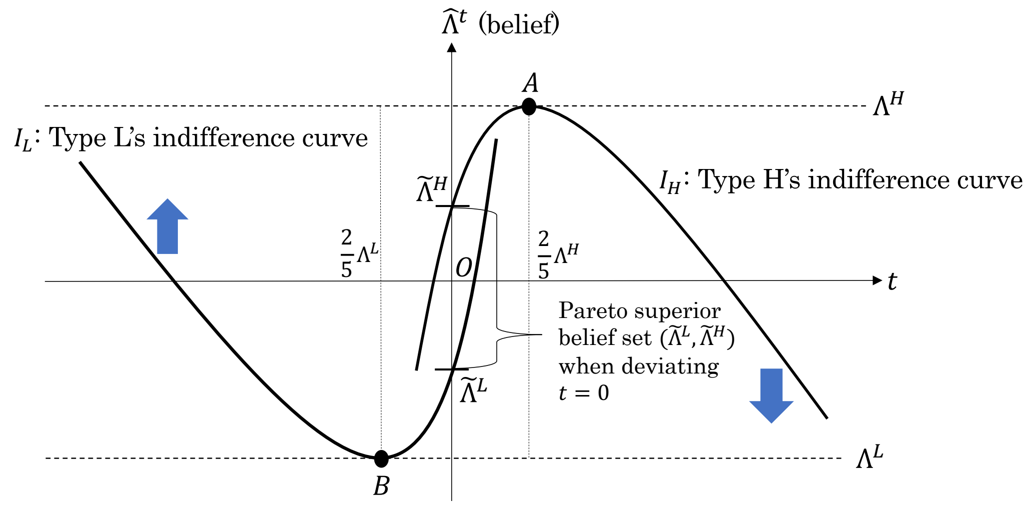

Figure 1: Two types' indifference curves at the equilibrium under complete information in situation G_1

incentive.[13] Point A is the optimal point of type H under complete information, where its utility is maximized under the belief constraint of $\hat{\Lambda}^t = \Lambda^H$. Indifference curve I_H in Figure 1 shows type H's utility level at point A. Similarly, point B is type L's optimal point under complete information with its belief constraint of $\hat{\Lambda}^t = \Lambda^L$. I_L shows type L's utility level at point B.

Here, let us define $\tilde{\Lambda}^s$ as the belief $\hat{\Lambda}^t$ such that points $(\frac{2}{5}\Lambda^s, \Lambda^s)$ and $(0, \hat{\Lambda}^t)$ on t-$\hat{\Lambda}^t$ plane are indifferent for type $s \in \{H, L\}$, which is designated at the intersection point of each indifference curve and vertical axis in Figure 1. $\tilde{\Lambda}^s$ is calculated as $3 \cdot \left(1 - \frac{2\sqrt{5}}{5}\right)\Lambda^s$ for each $s \in \{H, L\}$. From Figure 1, we can see that type H (L) has an incentive to deviate from point A (B) to point $(0, \hat{\Lambda}^0)$ if $\hat{\Lambda}^0 \in (\tilde{\Lambda}^L, \tilde{\Lambda}^H)$ holds, which reflects the concealment incentive. Based on this fact, when $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$, we can build a pooling equilibrium as in Figure 2.

Figure 2 provides one example of case (ii) in Lemma 1 where $\overline{\Lambda} = 0$. Since $t_1^H = t_1^L = 0$, country 2's belief is not updated, that is, $\hat{\Lambda}^0 = \overline{\Lambda}$ $(= 0)$ in the equilibrium. Type H and L's utility levels at the origin $(0, \overline{\Lambda})$ are shown by I_H^{pool} and I_L^{pool}, respectively. No type has an incentive to deviate from $t = 0$ when the equilibrium belief is the separating line between I_H^{pool} and I_L^{pool}, as illustrated in Figure 2. The essential point is that $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$; otherwise, either type has an incentive to deviate.

[13] By (3),(5),(20) and (21), $\frac{\partial u_1^s}{\partial \hat{\Lambda}^t} = \frac{\partial k_1^s}{\partial \hat{\Lambda}^t} f_1'(k_1^s) + \frac{\partial r^s}{\partial \hat{\Lambda}^t}(\overline{k} - k_1^s) - \frac{\partial k_1^s}{\partial \hat{\Lambda}^t} r^s = \frac{\partial k_1^s}{\partial \hat{\Lambda}^t} t_1^s + \frac{\partial r^s}{\partial \hat{\Lambda}^t}(\overline{k} - k_1^s) = -\frac{1}{18}t_1^s - \frac{1}{24}\Lambda^s + \frac{1}{72}\hat{\Lambda}^t$. Then, we have $\frac{\partial u_1^H}{\partial \hat{\Lambda}^t} < 0$ when $t_1^H > -\frac{3}{4}(\Lambda^H - \frac{1}{3}\hat{\Lambda}^t)$, and $\frac{\partial u_1^L}{\partial \hat{\Lambda}^t} > 0$ when $t_1^L < -\frac{3}{4}(\Lambda^L - \frac{1}{3}\hat{\Lambda}^t)$. We do not have to consider the tax rates such that $t_1^H \leq -\frac{3}{4}(\Lambda^H - \frac{1}{3}\hat{\Lambda}^t)$ and $t_1^L \geq -\frac{3}{4}(\Lambda^L - \frac{1}{3}\hat{\Lambda}^t)$. For type H, tax rates such that $t_1^H \leq -\frac{3}{4}(\Lambda^H - \frac{1}{3}\hat{\Lambda}^t)$ are dominated by various higher tax rates for any $\hat{\Lambda}^t$, because k_1^H is too large under those tax rates: $f'(k_1^H) < 0$ holds and the payment for capital $r(k_1^H - \overline{k})$ is too large. Similarly, for type L, tax rates such that $t_1^L \geq -\frac{3}{4}(\Lambda^L - \frac{1}{3}\hat{\Lambda}^t)$ are dominated by various lower tax rates, because k_1^L is too low under those tax rates.

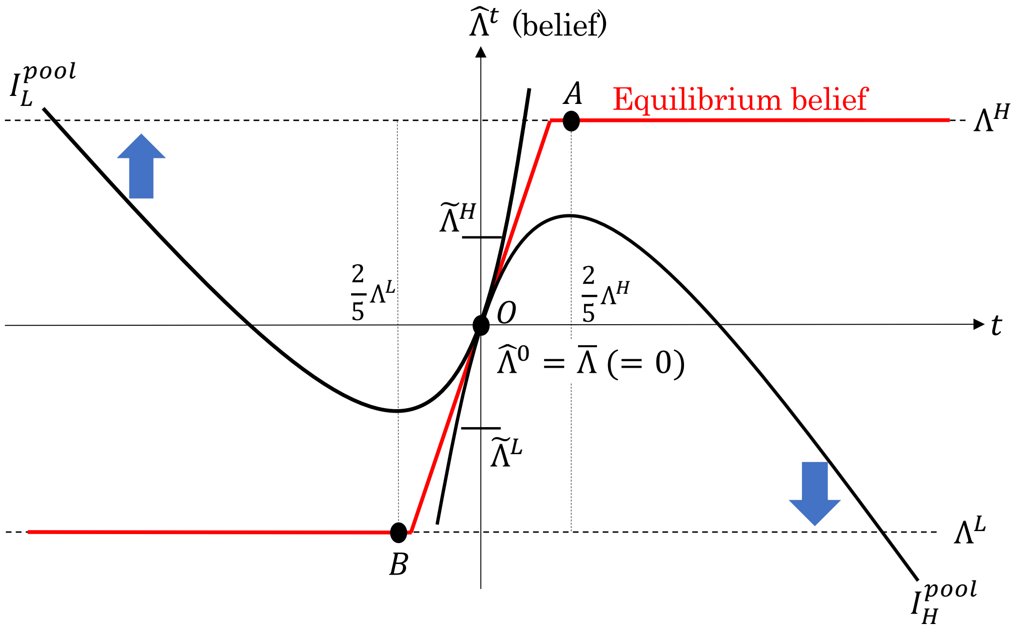

Figure 2: An example of case (ii) : $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$ in Lemma 1 with $\overline{\Lambda} = 0$

Figure 3 illustrates an example of the deviation when $\overline{\Lambda} \notin [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$, and presents an equilibrium outcome of case (i) : $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$ in Lemma 1.

Suppose to the contrary that $t_1^H = t_1^L = 0$ is an equilibrium outcome in Figure 3. Then, type H's equilibrium payoff is shown by $I_H{}'$ in Figure 3, which passes through point $C : (0, \overline{\Lambda})$. In this case, as presented in the figure, type H has an incentive to deviate from $t = 0$ to $\frac{2}{5}\Lambda^H$ for any off-the-equilibrium belief $\hat{\Lambda}^{\frac{2}{5}\Lambda^H} \in [\Lambda^L, \Lambda^H]$.[14] This deviation occurs when $\overline{\Lambda} > \tilde{\Lambda}^H$.

In the case of $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$, type H resorts to a mixed action strategy between $t_1^H = 0$ and $\frac{2}{5}\Lambda^H$, while type L chooses $t_1^L = 0$ surely. Type H and L's equilibrium payoffs at the strategy profile are shown by I_H^{hybrid} and I_L^{hybrid} in Figure 3, respectively. In the hybrid equilibrium, $\hat{\Lambda}^{\frac{2}{5}\Lambda^H} = \Lambda^H$ holds, since $t = \frac{2}{5}\Lambda^H$ is not chosen by type L. Moreover, the hybrid equilibrium is built so that $\hat{\Lambda}^0 = \tilde{\Lambda}^H (< \overline{\Lambda})$ is satisfied,[15] where type H's payoff from $t = 0$ is equal to that from $t = \frac{2}{5}\Lambda^H$.[16] Then, I_L^{hybrid} passes through point D, and I_H^{hybrid} passes through point D and A. In this situation,

[14] Since $\tilde{\Lambda}^H < \overline{\Lambda}$ and $\hat{\Lambda}^{\frac{2}{5}\Lambda^H} \leq \Lambda^H$, for type H, $(0, \overline{\Lambda}) \prec (0, \tilde{\Lambda}^H) \sim (\frac{2}{5}\Lambda^H, \Lambda^H) \precsim (\frac{2}{5}\Lambda^H, \hat{\Lambda}^{\frac{2}{5}\Lambda^H})$ holds.

[15] Suppose type H chooses 0 with probability $\tilde{q}$ and $\frac{2}{5}\Lambda^H$ with probability $1 - \tilde{q}$. Since the belief $\hat{\Lambda}^0$ must be calculated using Bayes' rule and the equilibrium strategies, we have $\hat{\Lambda}^0 = \frac{\tilde{q}\overline{\rho}_H}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^H + \frac{\overline{\rho}_L}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^L$ in the hybrid equilibrium. When $\tilde{\Lambda}^H < \overline{\Lambda}$, there exists $\tilde{q}$ such that $\hat{\Lambda}^0 = \tilde{\Lambda}^H$, since $\hat{\Lambda}^0$ is $\overline{\Lambda}(= \overline{\rho}^H\Lambda^H + \overline{\rho}^L\Lambda^L)$ when $\tilde{q} = 1$ and Λ^L when $\tilde{q} = 0$, and continuous function of $\tilde{q}$.

[16] This must hold because, otherwise, type H wants to increase the probability of the tax rate that gives higher utility.

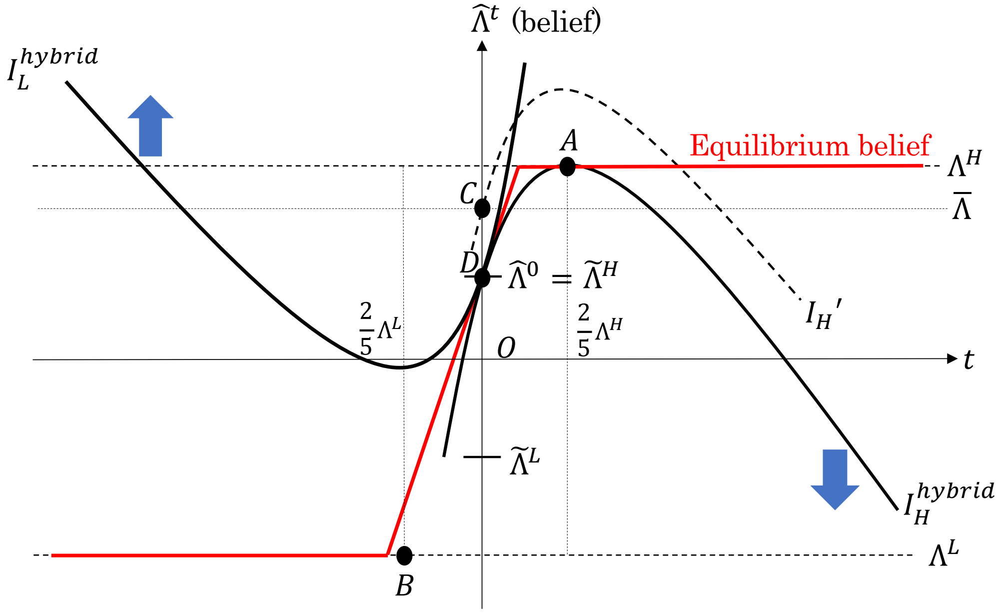

Figure 3: An example of case $(i) : \overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$ in Lemma 1

we can see from Figure 3 that no type deviates from the hybrid equilibrium if the equilibrium belief is the separating line between I_H^{hybrid} and I_L^{hybrid}. The important idea is that the mixed action strategy decreases the value $\hat{\Lambda}^0$ for the pooling behavior at $t = 0$ to be sustained. If $\hat{\Lambda}^0 > \tilde{\Lambda}^H$ holds, type H has an incentive to deviate to $\frac{2}{5}\Lambda^H$ with prob 1, as explained above.

The mechanism behind case $(iii) : \overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ is similar to case (ii). The results of Lemma 1 are summarized in figure 4.

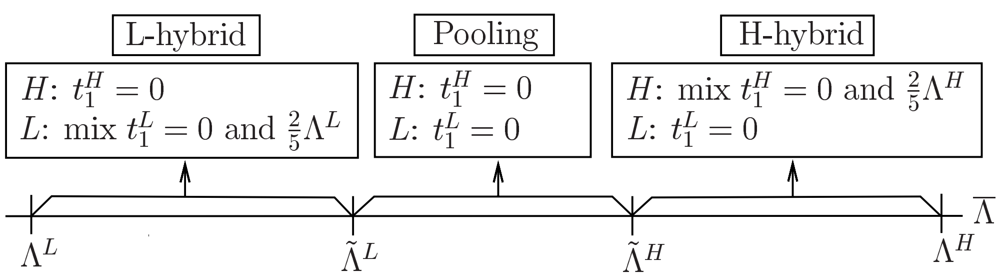

Figure 4: Summary of Lemma 1

This signaling model satisfies the double-crossing property. Generally, in models with the double-crossing property, there can be multiple PBE outcomes supported by unrealistic beliefs (Kolev and Prusa, 1999, 2002; Daley and Green, 2014; Chen et al., 2021). Surprisingly, we can show the uniqueness without using refinement criteria to exclude the unstable outcomes in Lemma 1, which derives from the two-way concealment incentives. See Appendix B for more detail.

3.5 Endogenous Sequencing

Here, we will identify which timing outcomes are attained under PBEs. First, let us confirm the result under complete information.

Proposition 1. *(Ogawa, 2013) When there is no uncertainty regarding productivity and $A_1 = A^H$ or A^L holds, $(a_1, a_2) = (e, e)$ is the unique outcome under SPE. In other words, both countries choose early and the simultaneous game is realized.*

As in Ogawa (2013), Proposition 1 can be captured by Hamilton and Slutsky (1990)'s mechanism. Figure 5 illustrates both cases where $s = H$ and $s = L$. Under complete information, by (11) and (12), reaction functions of type $s \in \{H, L\}$ and country 2 are $R_s : t_1^s = \frac{1}{3}(\Lambda^s + t_2)$ and $R_2(s) : t_2 = \frac{1}{3}(-\Lambda^s + t_1^s)$ in the figure, respectively. $\overline{u}_s$ and $\overline{u}_2$ are indifference curves of type s and country 2 at the simultaneous-move equilibrium point A when $s = H$ and point B when $s = L$.

Now, suppose $s = H$. Type H, a capital importer, achieves higher utility at a higher tax rate of country 2 given t_1^H, since $\frac{\partial r}{\partial t_i} < 0$. In contrast, country 2, a capital exporter, achieves higher utility at a lower tax rate of type H given t_2. Their origin is the assumption of a non-absentee capital ownership environment and asymmetric productivity. The area illustrated with vertical stripes depicts the Pareto superior set relative to the simultaneous-move equilibrium point A, and the point A' (A'') is the sequential-move equilibrium point when country 2 (1) is a Stackelberg leader. Now, there is no reaction function that enters the Pareto superior set, which implies that no country has an incentive to become a Stackelberg follower.[17]

Similarly, when $s = L$, the area depicted with horizontal stripes is the Pareto superior set relative to the simultaneous-move equilibrium point B, and the point B' (B'') is the sequential-move equilibrium point when country 2 (1) is a Stackelberg leader. This case has the same mechanism as in $s = H$.

[17] Consider a sequential-move equilibrium. For example, if country 2 (1) behaves as a leader, because it chooses its tax rate in anticipation of the reaction of country 1 (2), it chooses point A' (A''). However, country 1 (2) prefers simultaneous-move equilibrium point A to the sequential-move equilibrium point A' (A''). The sequential-move equilibrium puts the Stackelberg follower at a disadvantage relative to the simultaneous-move equilibrium A.

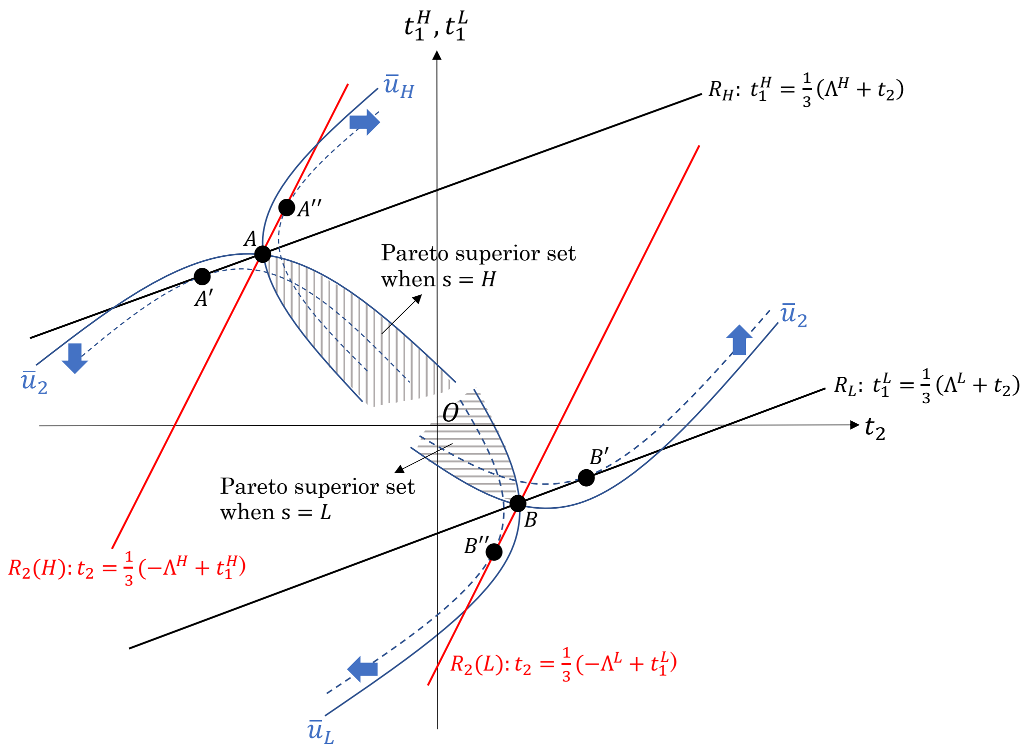

Figure 5: Reaction curves and Pareto superior sets in Proposition 1

3.5.1 Pure Action Strategies: non-equilibrium outcomes

Here, we focus on the pure action strategies under which both countries make timing choices with probability 1. Eight possible outcomes exist in our setting. First, we focus on the action profiles $(a_1^H, a_1^L, a_2) = (e, l, e)$, (l, e, e) and (e, e, e), and compare the results with the equilibrium under complete information. While countries 1 and 2 choose *early* under complete information, this is not commitment robust under our setting, and either country has an incentive to deviate from *early* to *late*.

Proposition 2. *There exists no PBE such that* $(a_1^H, a_1^L, a_2) = (e, l, e)$ *and* (l, e, e)*. Moreover, if* $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$ *holds with* $\tilde{\Lambda}^s = 3 \cdot \left(1 - \frac{2\sqrt{5}}{5}\right) \Lambda^s$ *for each* $s \in \{H, L\}$*, there exists no PBE such that* $(a_1^H, a_1^L, a_2) = (e, e, e)$*.*

Proof. See Appendix C. □

First, (e, l, e) is not attained in an equilibrium, because type H deviates from e to l. Figure 6 illustrates type H's incentive for deviation. Four reaction functions and point A, B and B' in Figure 6 are the same as in Figure 5. Suppose to the contrary that (e, l, e) is an equilibrium outcome. Since the type is fully revealed ($e \to$ type H and $l \to$ type L), type H and country 2's equilibrium payoffs are the same as those under complete information when $s = H$, which are shown by $\overline{u}_H$ and $\overline{u}_2$ in Figure 6. If type H deviates to *late*, it can pretend to be type L and make country 2 choose point B' as a Stackelberg leader. Then, type H can choose point C, which is beneficial for type H, since the capital price at point C is lower than that in point A because of the higher tax rate of country 2. That's why type H has an incentive to deviate to *late* (point $A \to C$). Similarly, (l, e, e) is not attained in an equilibrium because of the deviation by type L.

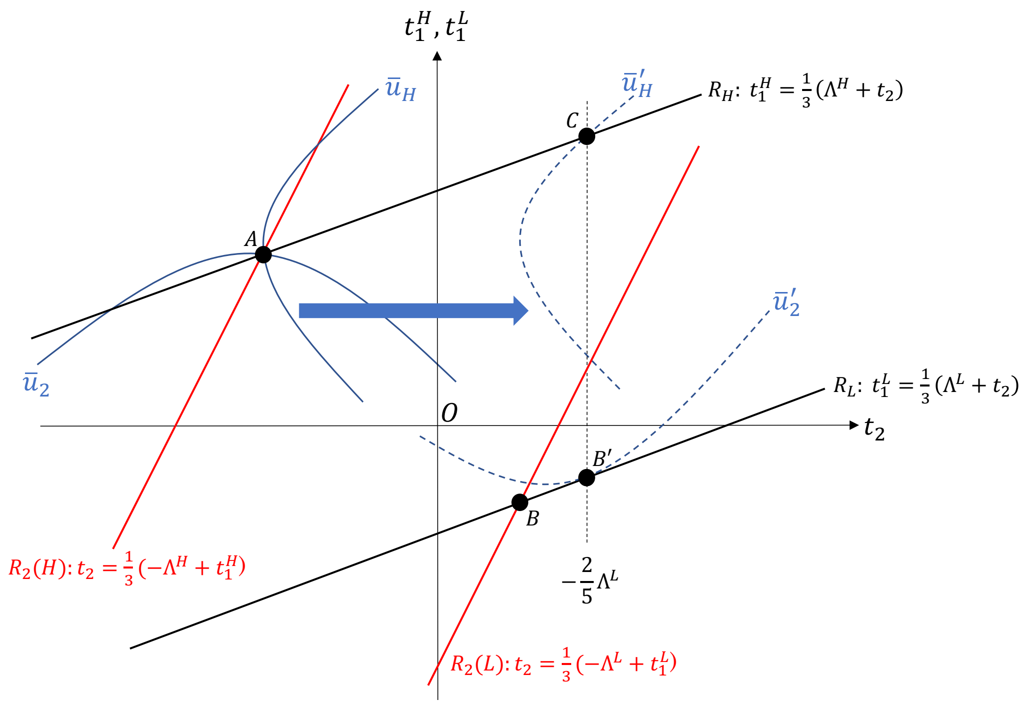

Figure 6: Type H's deviation from (e, l, e) in Proposition 2

Second, (e, e, e) is not attained in an equilibrium when $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$, because country 2 has an incentive to delay. If country 2 deviates to *late*, the signaling game on tax rates occurs. Then, the concealment incentive of both types will cause them to choose the same tax rate $t = 0$ (Lemma 1), which is a less aggressive tax rate and is beneficial for country 2.

Figure 7 illustrates country 2's incentive for deviation. Suppose to the contrary that (e, e, e) is an equilibrium outcome when $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$. Reaction functions R_H and R_L, and point A and B in Figure 7 are the same as in Figure 5. Since country 2 is not aware of country 1's true type, it chooses a tax rate based on an expected value of country 1's tax rate, denoted as $\hat{t}_1^e$.[18] The

[18] $\hat{t}_1^a$ is defined as $\hat{t}_1^a \equiv \sum_s \rho_s^a t_1^s$ in (12).

expected value $\hat{t}_1^e$ and country 2's reaction function are $\hat{t}_1^e = \frac{1}{3}(\overline{\Lambda} + t_2)$ and $t_2 = \frac{1}{3}(-\overline{\Lambda} + \hat{t}_1^e)$ by (11) and (12), respectively.[19] Then, the tax rate of country 2 in the simultaneous-move equilibrium is illustrated as the point D. Moreover, type H and type L's tax rates are illustrated as points E and F, and $\overline{u}_2(H)$ and $\overline{u}_2(L)$ show country 2's ex-post payoffs in the equilibrium when $s = H$ and L, respectively. If country 2 deviates to *late*, by Lemma 1, both types choose the same tax rate $t = 0$ and country 2 chooses point G as its best response.[20] By the deviation, type H (L) will decrease (increase) the tax rate from point E to G (from point F to G), which improves country 2's utility from $\overline{u}_2(H)$ to $\overline{u}_2'(H)$ (from $\overline{u}_2(L)$ to $\overline{u}_2'(L)$). Country 2 makes best use of concealment incentives to induce less aggressive tax rates from both types.

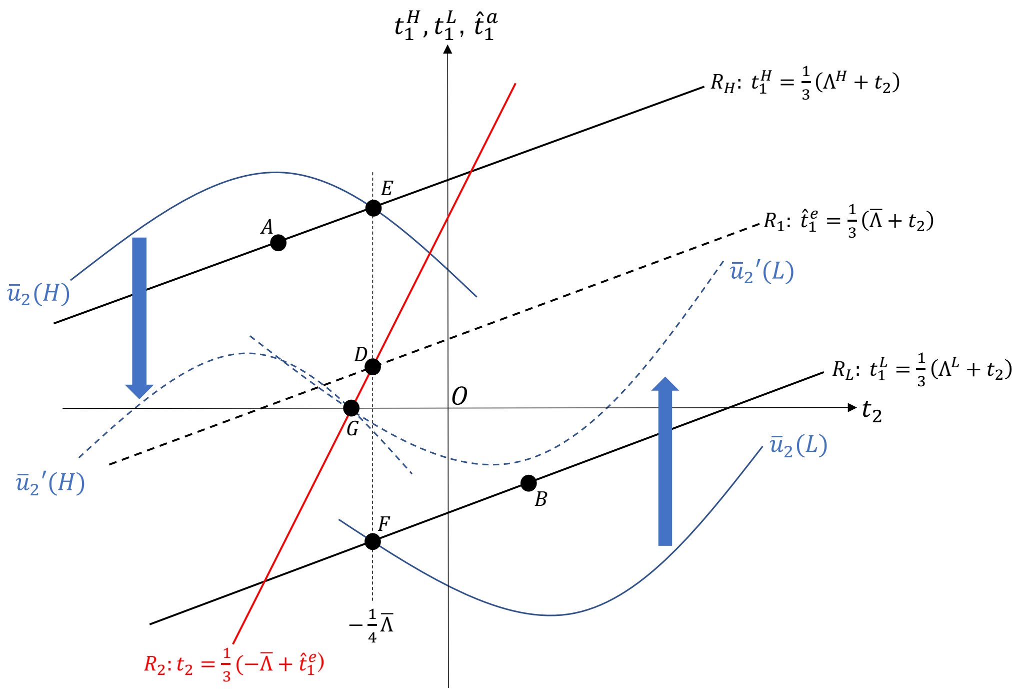

Figure 7: Country 2's deviation from (e, e, e) in Proposition 2

The others' non-equilibrium outcomes are as follows:

Proposition 3. *There exists no PBE such that* $(a_1^H, a_1^L, a_2) = (e, e, l)$, (l, e, l), (e, l, l), *and* (l, l, l).

Proof. See Appendix D. □

In the profile (e, e, l), the signaling situation on tax rates occurs and both types choose less aggressive tax rate $(t = 0)$ by the concealment incentive. Then, for any off-the-equilibrium belief

[19] Note that $\hat{\Lambda}^e = \overline{\Lambda}$ because of the pooling equilibrium.
[20] Note that $\hat{t}_1^e = 0$ when $t_1^H = t_1^L = 0$, and $\hat{\Lambda}^0 = \overline{\Lambda}$.

$\hat{\Lambda}^l$, either type has an incentive to deviate to *late* in order to choose a more aggressive tax rate. In the profile (l, e, l) and (e, l, l), since the type is fully revealed, either type has an incentive to deviate for pretending to be the other type, as in the case of (e, l, e). In the profile (l, l, l), country 2 wants to deviate to take a first-mover advantage as we will see in Proposition 4.

3.5.2 Pure Action Strategies: equilibrium outcome

We derive the equilibria of this timing game. Depending on prior belief $\overline{\Lambda}$, a Stackelberg outcome emerges in a PBE.

Proposition 4. *If $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$ holds, an outcome attained under PBE is $(a_1^H, a_1^L, a_2) = (l, l, e)$. Moreover, this survives D1.*

Proof. See Appendix E. □

In Ogawa (2013)'s setting, as we can see from Figure 5, both countries have a first-move incentive, since they seek an advantage by choosing the tax rate earlier to manipulate the capital price.[21] Even under our asymmetric information, country 2 chooses *early* for the first-mover benefit and has no incentive to deviate from (l, l, e), which is illustrated by Figure 8.

Four reaction functions and points A, B, D, E and F in Figure 8 are the same as in Figure 7. Under the equilibrium outcome (l, l, e), by (16), country 2 chooses point H based on the expected value of country 1's reaction $R_1 : \hat{t}_1^l = \frac{1}{3}(\overline{\Lambda} + t_2)$.[22] Type H and type L's tax rates are illustrated at points I and J, respectively. Then, country 2's expected payoff in the equilibrium is composed of $\overline{u}_2(H)$ and $\overline{u}_2(L)$, which show country 2's ex-post utility levels.

In this situation, even if country 2 deviates to *late*, it will not be better off in terms of expected utility for the following reason. If country 2 deviates to *late*, it chooses point D under the situation G_N, and country 2's expected payoff is composed of $\overline{u}'_2(H)$ and $\overline{u}'_2(L)$. Now, Figure 8 presents a case where $\overline{\Lambda} > 0$. In this situation, country 2 expects its role to be an exporter.[23] Thus, it makes a decision as an exporter to take a gain in the case of $s = H$, even though it will suffer a loss when country 1 is actually type L. Then, we can intuitively see that country 2's indifference curve is an inverted U-shape on $t_2 - \hat{t}_1^a$ plane. In Figure 8, two indifference curves $\overline{u}_2$ and $\overline{u}'_2$ are drawn to roughly convey country 2's expected utility level at points H and D, respectively.[24] We can see

[21] For example, when $s = H$, type H prefers point A'' to A and country 2 prefers point A' to A in Figure 5.

[22] Note that $\hat{\Lambda}^l = \overline{\Lambda}$ because of the pooling equilibrium.

[23] Country 2 expects its capital to be $\sum_s \rho_s^l k_2^s = \overline{k} - \frac{1}{6}\left(\overline{\Lambda} - \frac{2}{5}\overline{\Lambda}\right) < \overline{k}$ by (18).

[24] As mentioned later, it is not an accurate depiction.

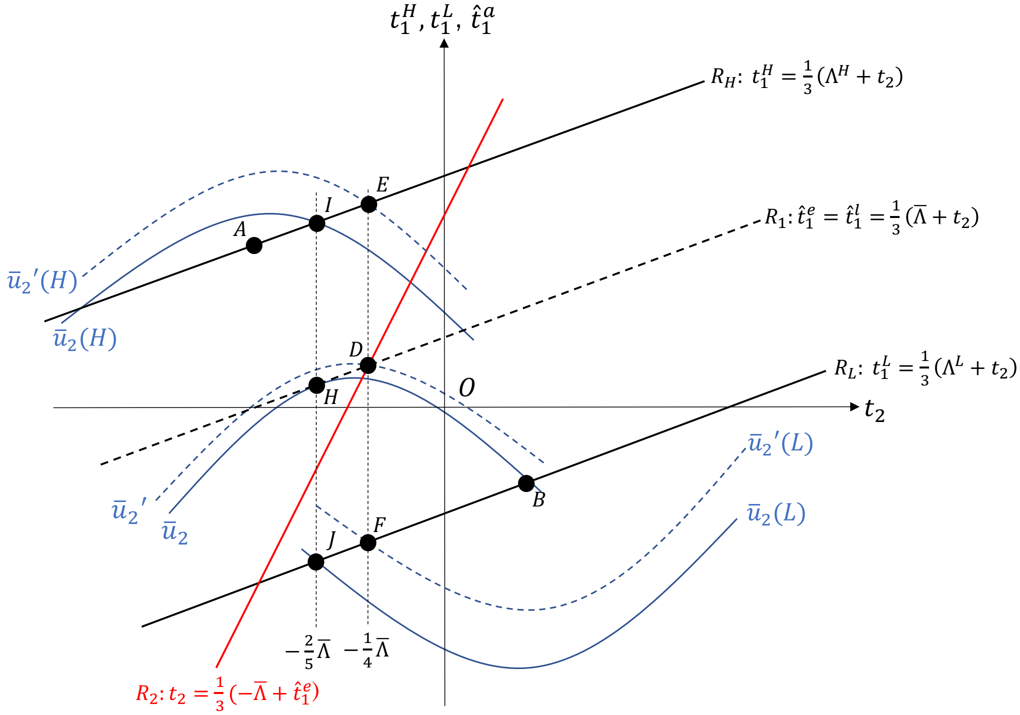

Figure 8: First-move incentive of country 2 in Proposition 4 in a case where $\overline{\Lambda} > 0$

that country 2 prefers point H to D and has no incentive to deviate to *late*. The same argument can be applied to the case where $\overline{\Lambda} < 0$.[25] In a precise sense, we cannot draw country 2's expected utility level as indifference curves on $t_2 - \hat{t}_1^a$ plane, because the expected utility cannot be expressed as a function of $\hat{t}_1^a$,[26] and this depiction is to understand country 2's incentive in an intuitive way.

Conversely, both types of country 1 may have a second-move incentive in "both" choosing *late*. When both types choose *late*, the type will not be revealed and $\hat{\Lambda}^l = \overline{\Lambda}$ holds, which is beneficial for both types.

Let us observe the mechanism in which each type does not deviate from choosing *late* in Figure 9. Reaction functions R_H, R_L, R_1 and R_2, and points A, B, D, I, H and J in Figure 9 are the same as in Figure 8. In the profile (l, l, e), country 2 chooses point H as a leader and its tax rate is $t_2 = -\frac{2}{5}\overline{\Lambda}$ by (16).[27] If either type deviates to *early*, since the situation G_N is realized, $t_2 = -\frac{1}{4}\hat{\Lambda}^e$ holds by (13). Here, we can find off-the-equilibrium belief $\hat{\Lambda}^e$ such that country 2's tax rate does not change before and after the deviation, that is $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda}$. If $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda}$, no type has an incentive

[25] When $\overline{\Lambda} = 0$, choosing *early* and *late* are indifferent for country 2.

[26] The expected utility includes $\mathbb{E}[(t_1^s)^2]$ and $\mathbb{E}[\Lambda^s t_1^s]$, which cannot be expressed by $\hat{t}_1^a$. $\overline{u}_2$ and $\overline{u}_2'$ in Figure 8 are indifference curves of the form $\mathbb{E}[u_2] - b_1 Var[t_1^s] - b_2 Cov[\Lambda^s, t_1^s]$ with b_1 and b_2 constant.

[27] Note that $\hat{\Lambda}^l = \overline{\Lambda}$ because of the pooling equilibrium.

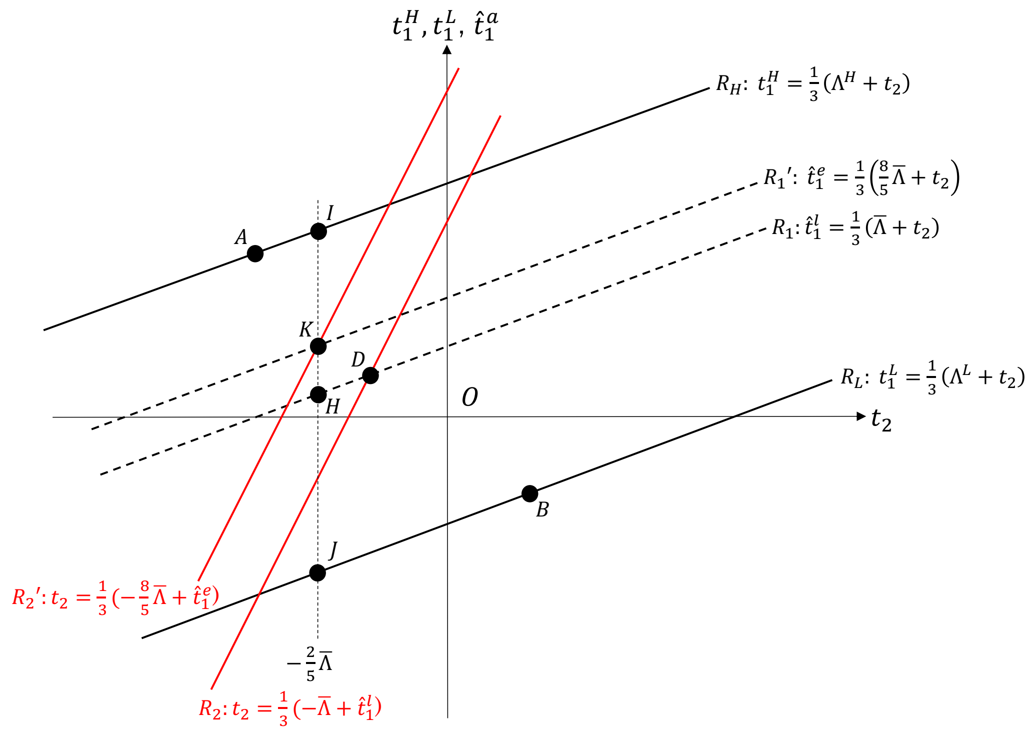

Figure 9: Second-move incentive of both types in Proposition 4

to deviate to *early*. In Figure 9, under off-the-equilibrium belief $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda}$, the expected value $\hat{t}_1^e$ and country 2's reaction function after the deviation are drawn as $R_1' : \hat{t}_1^e = \frac{1}{3}(\frac{8}{5}\overline{\Lambda} + t_2)$ and $R_2' : t_2 = \frac{1}{3}(-\frac{8}{5}\overline{\Lambda} + \hat{t}_1^e)$ by (11) and (12), respectively. Point K is the outcome after the deviation. We can see that country 2's tax rate at point K is the same as that at point H, so that each type's reaction also does not change after the deviation.

In order for the above situation to occur, $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$ must hold, which is the condition for $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda}$ to be a belief: $\hat{\Lambda}^e \in [\ \Lambda^L, \Lambda^H]$.[28] In the proof of Proposition 4, it is shown that $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda}$ is the unique off-the-equilibrium belief that sustains this pooling equilibrium: specifically, the conditions for no deviation by type H and L are $\frac{8}{5}\overline{\Lambda} \leq \hat{\Lambda}^e$ and $\frac{8}{5}\overline{\Lambda} \geq \hat{\Lambda}^e$, respectively. Thus, $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$ is the necessary and sufficient condition for the equilibrium outcome (l, l, e). The belief $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda}$ is reasonable in terms of $D1$ criterion by Cho and Kreps (1987).

We summarize the results of pure action strategies in the following table. The cases $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$ and $\overline{\Lambda} \in (\Lambda^L, \frac{5}{8}\Lambda^L)$ are analyzed in the next subsection.

[28] When $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$, $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda} \in [\frac{8}{5} \cdot \frac{5}{8}\Lambda^L, \frac{8}{5} \cdot \frac{5}{8}\Lambda^H] = [\ \Lambda^L, \Lambda^H]$ holds.

Table 1: Equilibrium (◯) and non-equilibrium (×) outcomes: pure action strategies

(e,e,e)	(e,e,l)	(e,l,e)	(l,e,e)	(e,l,l)	(l,e,l)	(l,l,e)	(l,l,l)
$\times(\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H])$	$\times$	$\times$	$\times$	$\times$	$\times$	◯ $(\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H])$	$\times$

The equilibrium outcome in this game is not always unique. Actually, the outcome (e,e,e) can be the equilibrium when uncertainty is small, that is, when $\overline{\Lambda}$ is close enough to Λ^L or Λ^H.[29] However, our equilibrium is supported by a reasonable belief in terms of the Cho and Kreps (1987) criterion as mentioned above, and the outcome (e,e,e) does not indicate its clear dominance in payoffs.

So far, we have analyzed the first- or second-move incentives and endogenous timing under asymmetric information by using reaction functions and the properties of payoffs toward other's actions as in the classical studies on endogenous timing (Gal-Or, 1985; Dowrick, 1986; Hamilton and Slutsky, 1990; Amir, 1995). The critical difference between analyses under complete information (classical studies) and asymmetric information (this study) is that reaction functions vary depending on the belief of the uninformed country. This makes it difficult to provide graphical characterizations of equilibria such as Hamilton and Slutsky (1990).

3.5.3 Mixed Action Strategies

Finally, we derive the equilibrium under which either country resorts to mixed action strategies. In this section, we denote a timing profile as $(p_1^H(e), p_1^L(e), p_2(e))$, which is a profile of probabilities of choosing *early* determined by type H, type L and country 2, respectively.

Proposition 5. *If $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$ holds, an outcome attained under PBE is $(p_1^H(e), p_1^L(e), p_2(e)) = (p_H^*, 0, 1)$, where $p_H^* = 1 - \frac{\rho'}{1-\rho'} \cdot \frac{1-\overline{\rho}}{\overline{\rho}}$ with $\rho' = \frac{\frac{5}{8}\Lambda^H - \Lambda^L}{\Lambda^H - \Lambda^L}$. If $\overline{\Lambda} \in (\Lambda^L, \frac{5}{8}\Lambda^L)$ holds, an outcome attained under PBE is $(p_1^H(e), p_1^L(e), p_2(e)) = (0, p_L^*, 1)$, where $p_L^* = 1 - \frac{1-\rho}{\rho} \cdot \frac{\overline{\rho}}{1-\overline{\rho}}$ with $\rho = \frac{-\frac{3}{8}\Lambda^L}{\Lambda^H - \Lambda^L}$. Moreover, each PBE survives D1.*

Proof. See Appendix E. □

The driving mechanism behind this result is similar to that of Proposition 4: country 2 chooses *early* ($p_2(e) = 1$) because of a first-move incentive, and each type s chooses *late* ($p_1^s(e) = 0$) because

[29] We can show that there exist $\Lambda^- \in (\Lambda^L, \tilde{\Lambda}^L)$ and $\Lambda^+ \in (\tilde{\Lambda}^H, \Lambda^H)$ such that (e,e,e) is an equilibrium outcome when $\overline{\Lambda} \in (\Lambda^L, \Lambda^-]$ or $\overline{\Lambda} \in [\Lambda^+, \Lambda^H)$.

of the concealment incentives. However, an outcome (l, l, e) cannot be sustained in this situation because it does not hold $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$, which is the necessary condition in Proposition 4 for off-the-equilibrium belief $\hat{\Lambda}^e$ to be a belief.

In this situation, either type adopts a mixed action strategy between *early* and *late*. The mixed action strategy is constructed so that country 2's tax rate does not change regardless of country 1's choice of timing. The idea behind this is the same as in Proposition 4. In the case of $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$, type H resorts to a mixed action strategy. In this case, country 2's tax rate after observing *early* is $-\frac{1}{4}\Lambda^H$ by (13), and that after observing *late* is $-\frac{2}{5}\hat{\Lambda}^l$ by (16). Note that both *early* and *late* are on-the-equilibrium paths, and $\hat{\Lambda}^l$ is calculated using Bayes' rule and equilibrium action strategies. We can find type H's mixed action strategy p_H^* such that $-\frac{1}{4}\Lambda^H = -\frac{2}{5}\hat{\Lambda}^l$ when $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$.[30] At that time, each type has no incentive for deviation, since country 2's tax rate does not vary. The same argument can be applied to the case of $\overline{\Lambda} \in (\Lambda^L, \frac{5}{8}\Lambda^L)$.

As $\overline{\Lambda}$ gets closer to Λ^H or Λ^L, the type using mixed action strategy increases the frequency of choosing *early*,[31] which implies that the equilibrium outcome approaches to that under complete information. Let us summarize the results from Proposition 4 and 5 in Figure 10.

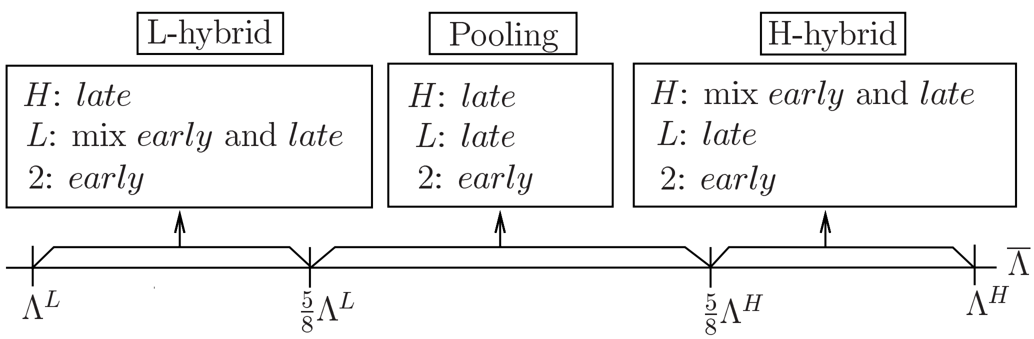

Figure 10: Equilibrium outcomes in Proposition 4 and 5

[30] By $(p_1^H(e), p_1^L(e), p_2(e)) = (p_H^*, 0, 1)$ and Bayes' rule, $\hat{\Lambda}^l = \frac{(1-p_H^*)\overline{\rho}}{(1-p_H^*)\overline{\rho}+(1-\overline{\rho})}\Lambda^H + \frac{(1-\overline{\rho})}{(1-p_H^*)\overline{\rho}+(1-\overline{\rho})}\Lambda^L$. When $\frac{5}{8}\Lambda^H < \overline{\Lambda}$, there exists p_H^* such that $\hat{\Lambda}^l = \frac{5}{8}\Lambda^H$, since the right-hand side is $\overline{\Lambda}$ when $p_H^* = 0$ and Λ^L when $p_H^* = 1$, and continuous function of p_H^*.

[31] We can check $\lim_{\overline{\rho}\to 1} p_H^* = 1$ and $\lim_{\overline{\rho}\to 0} p_L^* = 1$.

4 Welfare

4.1 Comparison of Equilibrium Welfare between Complete and Asymmetric information

Let $W^{com}(s)$ be the sum of utilities of both countries under complete information when country 1's type is $s \in \{H, L\}$, which is driven by the equilibrium in Proposition 1 (Ogawa, 2013). Let $W^{asy}(s)$ be the sum of the ex-post utilities of both countries under asymmetric information when country 1's type is s, which is driven by the equilibrium of Proposition 4 and 5. Now, we compare $W^{com}(s)$ with $W^{asy}(s)$ to disclose the effect of the information asymmetry on social welfare.

Proposition 6. *Consider the equilibria in Proposition 1, 4 and 5. Under the assumption that $-\frac{1}{5}\Lambda^L < \Lambda^H < -5\Lambda^L$, the following holds:*

(i) If $\overline{\Lambda} \in [\frac{5}{8}\Lambda^H, \Lambda^H)$, then $W^{asy}(H) = W^{com}(H)$ and $W^{asy}(L) > W^{com}(L)$;

(ii) If $\overline{\Lambda} \in (\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H)$, then $W^{asy}(H) > W^{com}(H)$ and $W^{asy}(L) > W^{com}(L)$;

(iii) If $\overline{\Lambda} \in (\Lambda^L, \frac{5}{8}\Lambda^L]$, then $W^{asy}(H) > W^{com}(H)$ and $W^{asy}(L) = W^{com}(L)$.

Proof. See Appendix F. □

These results indicate that social welfare under asymmetric information can be larger than that under complete information, where social welfare is defined as the sum of ex-post utilities.

In this model, social welfare is represented as the function of the gap between two tax rates $t_1^s - t_2$.[32] Social welfare increases when the gap is reduced, and it is maximized when $t_1^s = t_2$ holds.[33] Especially, when $t_1^s = t_2$ holds, we can see from (3) that the marginal productivity levels of both countries are balanced and the first-order condition for welfare maximization is satisfied. Under complete information, the gap between the tax rates is $t_1^s - t_2 = \frac{1}{2}\Lambda^s$ by (13),[34] so that inefficiency arises. Their tax rates differ under complete information because of the heterogeneous productivity ($\Lambda^s \neq 0$). In addition, in our environment, each country has a different incentive to manipulate the capital price, which causes the countries to set their tax rates in different directions.

However, in our model with asymmetric information, country 2 chooses a less aggressive tax rate toward each type, because of the uncertainty on the actual capital flow. For example, in the

[32]Social welfare in this model is $f_1(k_1^s) + f_2(k_2^s)$. The return and payment of capital are canceled out, since capital is fully owned by both countries. Social welfare is represented as the function of $t_1^s - t_2$, since k_1^s and k_2^s are characterized by $t_1^s - t_2$ from (6) and (7).

[33]Let $T \equiv t_1^s - t_2$ and $F(T) \equiv f_1(k_1^s) + f_2(k_2^s)$. Then, by (6) and (7), we have $F'(T) = \frac{\partial k_1^s}{\partial T} f_1'(k_1^s) + \frac{\partial k_2^s}{\partial T} f_2'(k_2^s) = -\frac{1}{4}T$. Thus, we have $F'(T) > 0$ when $T < 0$, $F'(T) < 0$ when $T > 0$ and $F'(T) = 0$ when $T = 0$.

[34]By (13), $t_1^s - t_2 = \frac{1}{3}\left(\Lambda^s - \frac{1}{4}\hat{\Lambda}^a\right) - \left(-\frac{1}{4}\hat{\Lambda}^a\right) = \frac{1}{2}\Lambda^s$. Note that $\hat{\Lambda}^a = \Lambda^s$ under complete information.

case of (ii) : $\overline{\Lambda} \in (\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H)$, country 2's tax rate under asymmetric information is $t_2 = -\frac{2}{5}\overline{\Lambda}$ by Proposition 4 and (16). This is a less aggressive tax rate than that under complete information ($t_2 = -\frac{1}{4}\Lambda^s$), because the tax rate under asymmetric information is higher toward type H and lower toward type L, that is $-\frac{1}{4}\Lambda^H < -\frac{2}{5}\overline{\Lambda}$ and $-\frac{1}{4}\Lambda^L > -\frac{2}{5}\overline{\Lambda}$, respectively.[35] This reduces the gap of the tax rates and improves welfare in both cases $s = H$ and $s = L$.[36] In the cases of (i) with $s = L$ and (iii) with $s = H$, we can apply a similar argument.

Alternatively, in the cases of (i) with $s = H$ and (iii) with $s = L$, welfare does not improve, because the same outcome is realized as that under complete information in a hybrid equilibrium. For example, in the case of (i) with $s = H$, by Proposition 5, a hybrid equilibrium emerges in which type H resorts to a mixed action strategy. In the hybrid equilibrium, type H's payoff is always the same as that under complete information, because a simultaneous-move game under complete information is realized when type H chooses *early*,[37] and type H's payoffs from *early* and *late* must be the same in the mixed action strategy.[38] Moreover, as mentioned in subsection 3.5.3, the hybrid equilibrium is constructed so that country 2's tax rate does not vary depending on country 1's choice of timing, which is fixed to the tax rate under complete information. As a result, $W^{asy}(H) = W^{com}(H)$ holds. We can apply the same argument to the case of (iii) with $s = L$.

Finally, note that the information asymmetry does not lead to Pareto improvement at all in terms of ex-post welfare. To verify this, let $u_2^{com}(s)$ be the utility of country 2 under complete information, and $u_2^{asy}(s)$ be the ex-post utility of country 2 under asymmetric information when country 1's type is s. These utilities are driven by Proposition 1, 4 and 5. We have the following.

Proposition 7. *Consider the equilibria in Proposition 1, 4 and 5. The following holds:*

(i) If $\overline{\Lambda} \in [\frac{5}{8}\Lambda^H, \Lambda^H)$, *then* $u_2^{com}(H) = u_2^{asy}(H)$ *and* $u_2^{com}(L) > u_2^{asy}(L)$;

(ii) If $\overline{\Lambda} \in (\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H)$, *then* $u_2^{com}(H) > u_2^{asy}(H)$ *and* $u_2^{com}(L) > u_2^{asy}(L)$;

(iii) If $\overline{\Lambda} \in (\Lambda^L, \frac{5}{8}\Lambda^L]$, *then* $u_2^{com}(H) > u_2^{asy}(H)$ *and* $u_2^{com}(L) = u_2^{asy}(L)$.

Proof. See Appendix F. □

[35] $-\frac{1}{4}\Lambda^H < -\frac{2}{5}\overline{\Lambda}$ and $-\frac{1}{4}\Lambda^L > -\frac{2}{5}\overline{\Lambda}$ hold by $\overline{\Lambda} \in (\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H)$.

[36] In the case of (ii) : $\overline{\Lambda} \in (\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H)$, the gap of tax rates is $t_1^s - t_2 = \frac{1}{3}\left(\Lambda^s - \frac{2}{5}\overline{\Lambda}\right) - \left(-\frac{2}{5}\overline{\Lambda}\right) = \frac{1}{3}\Lambda^s + \frac{4}{15}\overline{\Lambda}$ by (16). When $-\frac{1}{5}\Lambda^L < \Lambda^H < -5\Lambda^L$ (assumption of Proposition 6) holds, we have $|\frac{1}{3}\Lambda^s + \frac{4}{15}\overline{\Lambda}| < |\frac{1}{2}\Lambda^s|$, which means the gap of tax rates is reduced by the information asymmetry.

[37] When type H chooses *early*, the type is fully revealed in this hybrid equilibrium.

[38] Otherwise, type H seeks the probability that gives higher utility.

These results state that country 2's utility under complete information is larger than or equal to that under asymmetric information. By the definition of welfare, Propositions 6 and 7, we can see that there is no Pareto improvement by the information asymmetry in any cases.

Country 2's payoff does not improve, since it chooses the less aggressive tax rate owing to the misperception of the actual capital flow. For example in the case of (ii), as shown above, country 2's tax rate under asymmetric information is higher (lower) toward type H (L) than that under complete information. Then, type H (L)'s tax rate is also higher (lower) under asymmetric information because of $\frac{\partial t_1^s}{\partial t_2} > 0$ (strategic complementarity). These imply that the capital price under asymmetric information is higher (lower) in the case of $s = H$ (L) than that under complete information because of $\frac{\partial r}{\partial t_i} < 0$. This is harmful for country 2 regardless of country 1's type. A similar argument can be applied in the cases of (i) with $s = L$ and (iii) with $s = H$. Conversely, in the cases of (i) with $s = H$ and (iii) with $s = L$, the same outcome is realized as that under complete information, as mentioned above. Thus, country 2's utility does not change in those cases.

4.2 Further Analysis on Welfare: Excluding Timing Effect

The welfare results in section 4.1 are based on the comparison between simultaneous-move outcome with complete information and sequential-move outcome with asymmetric information. To isolate the positive effect of the information asymmetry on welfare, this section presents welfare analysis in a fixed timing. Let $W_T^I(s)$ be the sum of equilibrium utilities of both countries in a given timing structure $T \in \{G_N, G_1, G_2\}$ with information structure $I \in \{com(plete), asy(mmetric)\}$ when country 1's type is $s \in \{H, L\}$.[39] One scenario behind this analysis is that a social planner is able to organize the information set and timing structure in the economy, and both countries make decisions in the environment given by the planner. We have the following results.

Proposition 8. *Under the assumption that* $-\frac{1}{5}\Lambda^L < \Lambda^H < -5\Lambda^L$*, the following holds for any* $s \in \{H, L\}$*:*

(i) $W_{G_N}^{asy}(s) > W_{G_N}^{com}(s)$,

(ii) $W_{G_2}^{asy}(s) > W_{G_2}^{com}(s)$,

(iii) $W_{G_1}^{asy}(s) \geq W_{G_1}^{com}(s)$*, and the equality holds when* $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ *and type* L *chooses* $t_1^L = \frac{2}{5}\Lambda^L$ *or when* $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$ *and type* H *chooses* $t_1^H = \frac{2}{5}\Lambda^H$.

[39] Again, in G_N, G_1 and G_2, two countries move simultaneously, country 1 moves first, and country 2 moves first, respectively.

Proof. See Appendix F. □

In almost all cases, social welfare under asymmetric information is larger than that under complete information in a given timing. This can also be interpreted that a benevolent social planner would create an information asymmetry in many cases.

The mechanism behind welfare improvement is the same as explained in Proposition 6: country 2 becomes less aggressive because of the uncertainty regarding the actual capital flow, and hence the gap between two tax rates $t_1^s - t_2$ is reduced. For example, in the case of (i), the gaps under complete and asymmetric information are $t_1^s - t_2 = \frac{1}{2}\Lambda^s$ and $t_1^s - t_2 = \frac{1}{3}\Lambda^s + \frac{1}{6}\overline{\Lambda}$ by (13), respectively. We can see $|\frac{1}{3}\Lambda^s + \frac{1}{6}\overline{\Lambda}| < |\frac{1}{2}\Lambda^s|$ under the assumption that $-\frac{1}{5}\Lambda^L < \Lambda^H < -5\Lambda^L$.

The result in the case of (ii) shows the positive effect of the information asymmetry more clearly. In this case, we can show that $W_{G_2}^{asy}(s) \geq W_{G_N}^{com}(s) > W_{G_2}^{com}(s)$ holds. The first inequality holds by Proposition 6, that is our main welfare result. The second inequality means that, in the game with complete information, social welfare under a simultaneous-move situation is larger than that under a sequential-move situation. In our environment, a Stackelberg sequence induces a leader to choose a more aggressive tax rate to commit a desirable capital price,[40] because each country has a different incentive for manipulation of capital price, which has a negative effect on welfare. From two inequalities, we can see that the information asymmetry brings a large improvement in welfare.

In the case of (iii), while the mechanism for welfare improvement is the same as above, the equality $W_{G_1}^{asy}(s) = W_{G_1}^{com}(s)$ also holds in some situations where country 1 chooses the same tax rate as that under complete information in its mixed action strategy, as we see in Lemma 1.

5 Discussion

5.1 Types of Uncertainty

In our model, we incorporate the type of uncertainty in which the uninformed country does not know whether its productivity is higher or lower than the opponent's productivity. However, we can consider another type of uncertainty in which the uninformed country knows whether its productivity is higher than the opponent's or not but does not know the degree of the gap between their productivity levels. In our model, this can be referred to as $A_2 \in (0, A^L)$ or $A_2 \in (A^H, \infty)$.

[40] Suppose $s = H$, which means country 1 is an importer and country 2 is an exporter. Under complete information, country 2 tax rates in the situation G_N and G_2 are $t_2 = -\frac{1}{4}\Lambda^H$ and $t_2 = -\frac{2}{5}\Lambda^H$, respectively. The tax rate in the situation G_2 is more aggressive toward country 1, an importer, because of $-\frac{2}{5}\Lambda^H < -\frac{1}{4}\Lambda^H$.

For example, we consider the case $A_2 \in (0, A^L)$, where country 2 knows that it is a capital exporter owing to $A_2 < A_1$. In this case, we can show that type L has an incentive to differentiate itself from type H. Since country 2 is a capital exporter, it wants to increase the capital price. Then, type H wants to pretend to be type L, because, when country 2 thinks that its opponent is type L, it acts less aggressively and the capital price becomes lower than if it thinks that the opponent is type H. However, type L does not have an incentive to mimic type H, because it would make country 2 more aggressive, and thereby contribute to an increase in the capital price. Thus, this case can lead to a separating equilibrium. Contrary to our results, social welfare may not improve in this case, because the separating equilibrium causes inefficiency owing to over-investment, as in the standard signaling models stemming from Spence (1973). The case of $A_2 \in (A^H, \infty)$ is similar.

While our assumption on uncertainty may seem specific, this model implies that each country may have an incentive to conceal or distort its true productivity to improve its welfare. Our model provides theoretical and empirical motivations on countries' strategic interactions through information disclosure.

5.2 Capital Ownership and Share

In the case where the capital is fully owned by the outsiders of the economy (Kempf and Rota-Graziosi, 2010), all the countries in the economy act as capital importers, and a separating equilibrium emerges for the similar mechanism as explained in 5.1. Thus, our results cannot be applied to an absentee capital ownership environment. However, our model may be extended to the situation where the capital is partially owned by the outsiders of the economy. Hindriks and Nishimura (2017) analyze cases between full capital ownership (Ogawa, 2013) and no-capital ownership (Kempf and Rota-Graziosi, 2010). They conclude that a simultaneous-move outcome is realized as in Ogawa (2013), if the degree of capital ownership is more than a certain level. This implies that our result is not limited to the polar case of full capital ownership.

The equilibrium outcomes are also affected by the share of the capital between the two countries. Assuming strategic complementarity, Kempf and Rota-Graziosi (2015) develop the model with flexible capital share and characterize the equilibrium by *plain complementarity* and *substitutability* of tax rates.[41] Particularly, they conclude that the SPE yields two sequential move outcomes when both countries display either plain complementarity or substitutability (Kempf and Rota-Graziosi,

[41] A country displays *plain complementarity (substitutability)* if an increase in the tax rate of that country improves (reduces) welfare of another country.

2010), and a unique simultaneous move outcome when one country displays plain complementarity and the other displays plain substitutability (Ogawa, 2013). Based on their work, this study's findings may hold in the flexible capital share economy where one country displays plain complementarity and the other, plain substitutability.

6 Conclusion

An analysis of the leadership in a simple tax competition environment under asymmetric information leads to a new strategic effect. Under Ogawa (2013)'s setup, this study introduces information asymmetry regarding a country's productivity level of private goods, where one country is informed about its own and opponent's productivity, while the other country knows only its own productivity. The study shows that each type of the informed country will have an incentive to pretend to be the other type. This concealment incentives will lead to pooling equilibria and, consequently, a Stackelberg outcome emerges endogenously, while the simultaneous move is the unique outcome under complete information. Our results show that the simultaneous move in the tax competition game may not be commitment robust even in Ogawa (2013)'s non-absentee capital ownership environment.

Moreover, the study shows that ex-post social welfare can become larger under asymmetric information than under complete information. A positive effect of the information asymmetry on welfare is also shown in the analysis with a fixed timing structure. Especially, taking advantage of the uncertainty, the informed country can improve its utility by misleading the uninformed country about the actual capital flow in the economy. This implies that each country has an incentive to manipulate information or create uncertainty and that incomplete information environments are realized endogenously. Future studies can extend this model to a situation where each country can manipulate its information, and can also conduct empirical analyses to uncover countries' disincentives for information disclosure.

Bibliography

Amir, Rabah, "Endogenous Timing in Two-Player Games: A Counterexample," *Games and Economic Behavior*, 1995, *9* (2), 234–237.

Bacchetta, Philippe and Maria Paz Espinosa, "Information Sharing and Tax Competition among Governments," *Journal of International Economics*, 1995, *39* (1-2), 103–121.

_ **and María Paz Espinosa**, "Exchange-of-Information Clauses in International Tax Treaties," *International Tax and Public Finance*, 2000, *7* (3), 275–293.

Chen, ChiaHui, Junichiro Ishida, and Wing C. Suen, "Signaling under Double-Crossing Preferences," *Working Paper*, 2021.

Cho, In-Koo and David M. Kreps, "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 1987, *102* (2), 179–221.

Daley, Brendan and Brett Green, "Market Signaling with Grades," *Journal of Economic Theory*, 2014, *151*, 114–145.

de GaridelThoron, Thomas, "Welfare-Improving Asymmetric Information in Dynamic Insurance Markets," *Journal of Political Economy*, 2005, *113* (1), 121–150.

Dowrick, Steve, "von Stackelberg and Cournot Duopoly: Choosing Roles," *The RAND Journal of Economics*, 1986, *17* (2), 251–260.

Gal-Or, Esther, "First Mover and Second Mover Advantages," *International Economic Review*, 1985, *26* (3), 649–653.

Goodspeed, Timothy J., "Decentralization and Natural Disasters," *Working Paper*, 2013.

_ **and Andrew F. Haughwout**, "On the Optimal Design of Disaster Insurance in a Federation," *Economics of Governance*, 2012, *13* (1), 1–27.

Hamilton, Jonathan H and Steven M Slutsky, "Endogenous Timing in Duopoly Games: Stackelberg or Cournot Equilibria," *Games and Economic Behavior*, 1990, *2* (1), 29–46.

Hindriks, Jean and Yukihiro Nishimura, "Equilibrium Leadership in Tax Competition Models with Capital Ownership: A Rejoinder," *International Tax and Public Finance*, 2017, *24* (2), 338–349.

Huizinga, Harry and Søren Bo Nielsen, "Withholding Taxes or Information Exchange: The Taxation of International Interest Flows," *Journal of Public Economics*, 2003, *87* (1), 39–72.

Kempf, Hubert and Grégoire Rota-Graziosi, "Endogenizing Leadership in Tax Competition," *Journal of Public Economics*, 2010, *94*, 768–776.

_ **and** _ , "Further Analysis on Leadership in Tax Competition: The Role of Capital Ownership—A Comment," *International Tax and Public Finance*, 2015, *22*, 1028–1039.

Kolev, Dobrin R. and Thomas J. Prusa, "Tariff Policy for a Monopolist in a Signaling Game," *Journal of International Economics*, 1999, *49*, 51–76.

_ **and** _ , "Dumping and Double Crossing: The (In)effectiveness of Cost-Based Trade Policy under Incomplete Information," *International Economic Review*, 2002, *43* (3), 895–918.

Koufopoulos, Kostas and Roman Kozhan, "Welfare-Improving Ambiguity in Insurance Markets with Asymmetric Information," *Journal of Economic Theory*, 2014, *151*, 551–560.

Mailath, George J., "Endogenous Sequencing of Firm Decisions," *Journal of Economic Theory*, 1993, *59* (1), 169–182.

Morris, Stephen and Hyun Song Shin, "Social Value of Public Information," *American Economic Review*, 2002, *92* (5), 1521–1534.

Normann, Hans-Theo, "Endogenous Stackelberg Equilibria with Incomplete Information," *Journal of Economics*, 1997, *66* (2), 177–187.

_ , "Endogenous Timing with Incomplete Information and with Observable Delay," *Games and Economic Behavior*, 2002, *39* (2), 282–291.

Ogawa, Hikaru, "Further Analysis on Leadership in Tax Competition: The Role of Capital Ownership," *International Tax and Public Finance*, 2013, *20* (3), 474–484.

Spence, Michael, "Job Market Signaling," *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

Wildasin, David E., "Disaster Avoidance, Disaster Relief, and Policy Coordination in a Federation," *Working Paper*, 2011.

Appendices

A Preliminary

The reminder of basic notations: $A_1 \in \{A^H, A^L\}$ $(A^H > A^L)$ is the productivity of country 1, $A_2 \in (A^L, A^H)$ is the productivity of country 2, $\Lambda^s \equiv A^s - A_2$ where $s \in \{H, L\}$, and $\overline{\Lambda} \equiv \overline{\rho}_H \Lambda^H + \overline{\rho}_L \Lambda^L$ is a prior belief. By the assumption of $A_2 \in (A^L, A^H)$, it holds that $\Lambda^H > 0 > \Lambda^L$ and $\overline{\Lambda} \in [\Lambda^L, \Lambda^H]$. $\hat{\Lambda}^a$ and $\hat{\Lambda}^t$ are country 2's posterior beliefs after observing country 1's timing choice a and tax rate t, respectively.

First, we prove the following lemma, which is useful for the later analysis. Let $u_i(a_1^H, a_1^L, a_2|s)$ be country i's utility when country 1's type is s.

Lemma 2. *Let* $v_i(a_1^H, a_1^L, a_2|s) \equiv 36u_i(a_1^H, a_1^L, a_2|s) - \varepsilon_{is}$ *for each* $i \in \{1, 2\}$*, where* $\varepsilon_{1s} \equiv 36(A^s - \overline{k})\overline{k}$ *and* $\varepsilon_{2s} \equiv 36(A_2 - \overline{k})\overline{k}$ *for each* s*. We define* $\alpha^s \equiv \Lambda^s - \frac{1}{4}\hat{\Lambda}^a$*,* $\beta^s \equiv \Lambda^s - \frac{2}{5}\hat{\Lambda}^a$*, and* $\gamma^s \equiv \frac{3}{2}\Lambda^s - t_1 - \frac{1}{2}\hat{\Lambda}^{t_1}$*. Then, the following holds:*

(i) : The situation G_N (Simultaneous move)[42]

[Type s] $v_1(a_1^H, a_1^L, a_2|s) = 3(\alpha^s)^2$

[Country 2] $v_2(a_1^H, a_1^L, a_2|s) = -5(\alpha^s)^2 + 6\Lambda^s\alpha^s$

(ii) : The situation G_2 (Country 2 moves first)[43]

[Type s] $v_1(a_1^H, a_1^L, e|s) = 3(\beta^s)^2$

[Country 2] $v_2(a_1^H, a_1^L, e|s) = -5(\beta^s)^2 + 6\Lambda^s\beta^s$

(iii) : The situation G_1 (Country 1 moves first)

[Type s] $v_1(a_1^H, a_1^L, l|s) = (\gamma^s)^2 + 6t_1\gamma^s$

[Country 2] $v_2(a_1^H, a_1^L, l|s) = -3(\gamma^s)^2 + 6(\Lambda^s - t_1)\gamma^s$

Proof. (i) [Type s]: When a game is a simultaneous move, by (14)(15)

$$k_1^s = \overline{k} + \frac{1}{6}\alpha^s, \tag{A.1}$$

$$r^s = A^s - 2\overline{k} - \frac{2}{3}\alpha^s. \tag{A.2}$$

(5), ($A.1$) and ($A.2$) imply that type s's utility is $u_1(a_1^H, a_1^L, a_2|s) = \frac{1}{36}[3(\alpha^s)^2 + 36(A^s - \overline{k})\overline{k}]$. Then, we have $v_1(a_1^H, a_1^L, a_2|s) = 3(\alpha^s)^2$. (i) [Country 2], (ii) and (iii) are similar to (i) [Type s]. □

[42]For example, if country 1 is type H and both countries choose *early*, then type H's utility is $v_1(e, a_1^L, e|H) = 3(\alpha^H)^2$. The timing profile (a_1^H, a_1^L, a_2) is not specified in the Lemma, since there are some cases which describe the situation G_N.

[43]For example, if country 1 is type L, then its utility is $v_1(a_1^H, l, e|L) = 3(\beta^L)^2$.

α^s, β^s and γ^s correspond to the situation G_N, G_2 and G_1, respectively. Note that the utilities in case (iii) of Lemma 2 depend on t_1. This is because, as we see in the analysis of situation G_1 (3.4 Sequential Move (G_1)), the structure of tax competition changes depending on both types' timing choices. If $a_1^H \neq a_1^L$ in an equilibrium, since the type is fully revealed in the separating equilibrium, then the tax rate of type s is $t_1 = \frac{2}{5}\Lambda^s$ (by (22)) and country 2's belief after observing t_1 is $\hat{\Lambda}^{t_1} = \hat{\Lambda}^e = \Lambda^s$ (and then $\gamma^s = \frac{3}{5}\Lambda^s$). [44] If $a_1^H = a_1^L (= e)$, then t_1 and $\hat{\Lambda}^{t_1}$ are determined by the signaling game on the tax choice (Lemma 1). For example, in the case of $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$, $t_1 = 0$ for both types, and $\hat{\Lambda}^0 = \overline{\Lambda}$ hold (and then $\gamma^s = \frac{3}{2}\Lambda^s - \frac{1}{2}\overline{\Lambda}$).[45]

In the following proofs, we consider v_i as country i's utility because the terms $\frac{1}{36}$, ε_{1s}, and ε_{2s} have no effect on countries' decisions.

B Proof of Lemma 1

Here, we will prove Lemma 1. Lemma 1 analyzes the case where the situation G_1 and $a_1^H = a_1^L = e$ are realized in an equilibrium. Thus, by Lemma 2, we focus on the utility $v_1(e, e, l|H)$ and $v_1(e, e, l|L)$.

To analyze the signaling game in Lemma 1, it is useful to consider both types' utilities as functions of $(\Lambda^s, \hat{\Lambda}^t, t)$. Here, using Lemma 2 (iii), we define $w_1(\Lambda^s, \hat{\Lambda}^t, t) \equiv v_1(e, e, l|s) = (\gamma^s)^2 + 6t\gamma^s$ with $\gamma^s = \frac{3}{2}\Lambda^s - t - \frac{1}{2}\hat{\Lambda}^t$. We can consider Λ^s as the type. For example, $w_1(\Lambda^H, \overline{\Lambda}, 0)$ represents type H's utility when its tax rate is 0 and country 2's belief is $\hat{\Lambda}^0 = \overline{\Lambda}$.

To prove Lemma 1, we start with the following lemma.

Lemma 3. *We have the following:*

(i) $$w_1(\Lambda^s, \hat{\Lambda}^t, t) = -5t^2 + 6\left(\Lambda^s - \tfrac{1}{3}\hat{\Lambda}^t\right)t + \tfrac{9}{4}\left(\Lambda^s - \tfrac{1}{3}\hat{\Lambda}^t\right)^2 = \tfrac{1}{4}\left(-2t + 3(\Lambda^s - \tfrac{1}{3}\hat{\Lambda}^t)\right)\left(10t + 3(\Lambda^s - \tfrac{1}{3}\hat{\Lambda}^t)\right),$$

(ii) $w_1(\Lambda^s, \hat{\Lambda}^t, t)$ *is maximized at* $t = \frac{3}{5}\left(\Lambda^s - \frac{1}{3}\hat{\Lambda}^t\right)$, *and the maximum value is* $\frac{81}{20}(\Lambda^s - \frac{1}{3}\hat{\Lambda}^t)^2$,

(iii) There exists a unique value $\tilde{\Lambda}^s \in (\Lambda^L, \Lambda^H)$ *such that* $w_1(\Lambda^s, \tilde{\Lambda}^s, 0) = \max_t w_1(\Lambda^s, \Lambda^s, t)\ \forall s$, *and the value is* $\tilde{\Lambda}^s = 3\left(1 - \frac{2\sqrt{5}}{5}\right)\Lambda^s$ $(\approx 0.317\Lambda^s)$. *Note that* $\tilde{\Lambda}^H > 0$ *and* $\tilde{\Lambda}^L < 0$.

[44] Since the type is fully revealed by the timing choices, country 2's belief is not updated by the tax rate t_1. Thus, $\hat{\Lambda}^{t_1} = \hat{\Lambda}^e$ holds.

[45] Both types choose $t_1 = 0$; then, the belief is not updated.

Proof. Proof of (i): By $w_1(\Lambda^s, \hat{\Lambda}^t, t) \equiv v_1(e, e, l|s) = (\gamma^s)^2 + 6t\gamma^s$ with $\gamma^s = \frac{3}{2}\Lambda^s - t - \frac{1}{2}\hat{\Lambda}^t$, we can obtain (i). Proof of (ii): Use Lemma 3(i). Proof of (iii): Use Lemma 3$(i)(ii)$. □

Let us summarize the results of Lemma 3 as Figure B.1. We can see that there exist two-way concealment incentives around $t = 0$,[46] which leads to the pooling equilibrium.

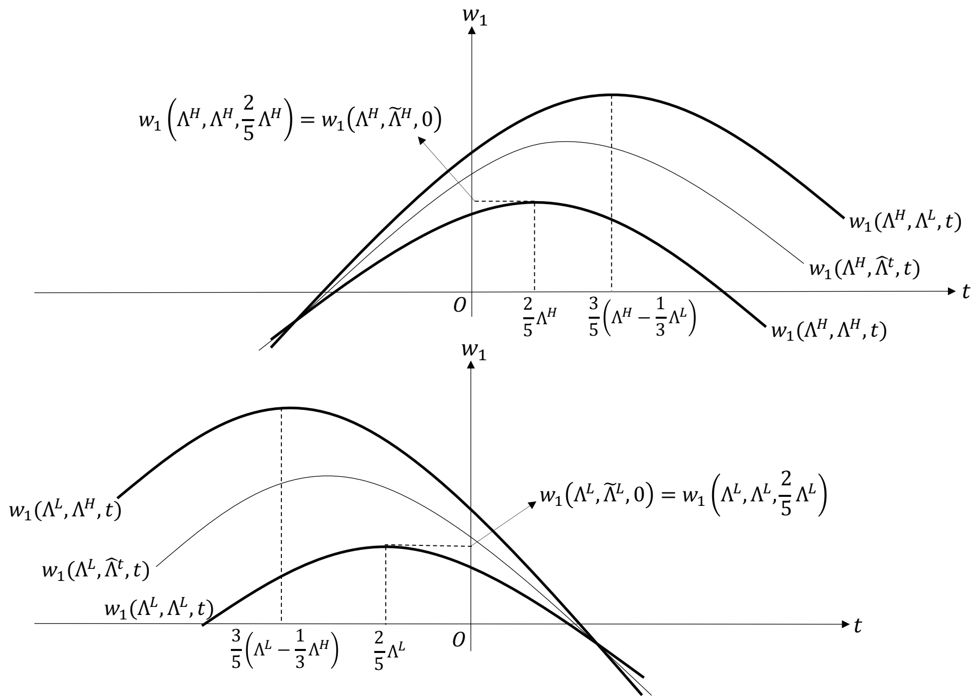

Figure B.1: The summary of Lemma 3: $\hat{\Lambda}^t \in (\Lambda^L, \Lambda^H)$

To understand each proof clearly, we will often use both types' indifference curves on the $t - \hat{\Lambda}^t$ plane. By Lemma 3(i), type s's indifference curve with utility level $\overline{w}$ satisfies $-5t^2 + 6\left(\Lambda^s - \frac{1}{3}\hat{\Lambda}^t\right)t + \frac{9}{4}\left(\Lambda^s - \frac{1}{3}\hat{\Lambda}^t\right)^2 = \overline{w}$. Each quadratic form has two curves on the t-$\hat{\Lambda}^t$ plane, but one curve violates $\hat{\Lambda}^t \in [\Lambda^L, \Lambda^H]$.[47] As a result, type H's indifference curve is $\hat{\Lambda}^t = 3\Lambda^H + 4t - 4(\frac{9}{4}t^2 + \frac{1}{4}\overline{w})^{\frac{1}{2}}$, and L's one is $\hat{\Lambda}^t = 3\Lambda^L + 4t + 4(\frac{9}{4}t^2 + \frac{1}{4}\overline{w})^{\frac{1}{2}}$. It can be also shown that type H's curve is concave, and type L's curve is convex. Moreover, their indifference curves are tangent to each other at $t = 0$. The slope of the tangent line of type H's indifference curve is $\frac{d\hat{\Lambda}^t}{dt} = 4 - 18t(\frac{9}{4}t^2 + \frac{1}{4}\overline{w})^{-\frac{1}{2}}$, and type L's one is $\frac{d\hat{\Lambda}^t}{dt} = 4 + 18t(\frac{9}{4}t^2 + \frac{1}{4}\overline{w})^{-\frac{1}{2}}$. We can see that the slopes are the same when $t = 0$.

[46]In Figure B.1, around $t = 0$, we can see that type H's utility increases as the belief $\hat{\Lambda}^t$ gets close to Λ^L, while type L's utility increases as $\hat{\Lambda}^t$ gets close to Λ^H.

[47]The indifference curves can be rewritten as $\hat{\Lambda}^t = 3\Lambda^s + 4t \pm 4(t^2 + \frac{1}{4}(5t^2 + \overline{w}))^{\frac{1}{2}}$ for each $s \in \{H, L\}$. Then, for type H, it holds that $\hat{\Lambda}^t = 3\Lambda^H + 4t + 4(t^2 + \frac{1}{4}(5t^2 + \overline{w}))^{\frac{1}{2}} > \Lambda^H$ for any t. Similarly, for type L, it holds that $\hat{\Lambda}^t = 3\Lambda^L + 4t - 4(t^2 + \frac{1}{4}(5t^2 + \overline{w}))^{\frac{1}{2}} < \Lambda^L$ for any t.

Let us summarize these results as Figure B.2.

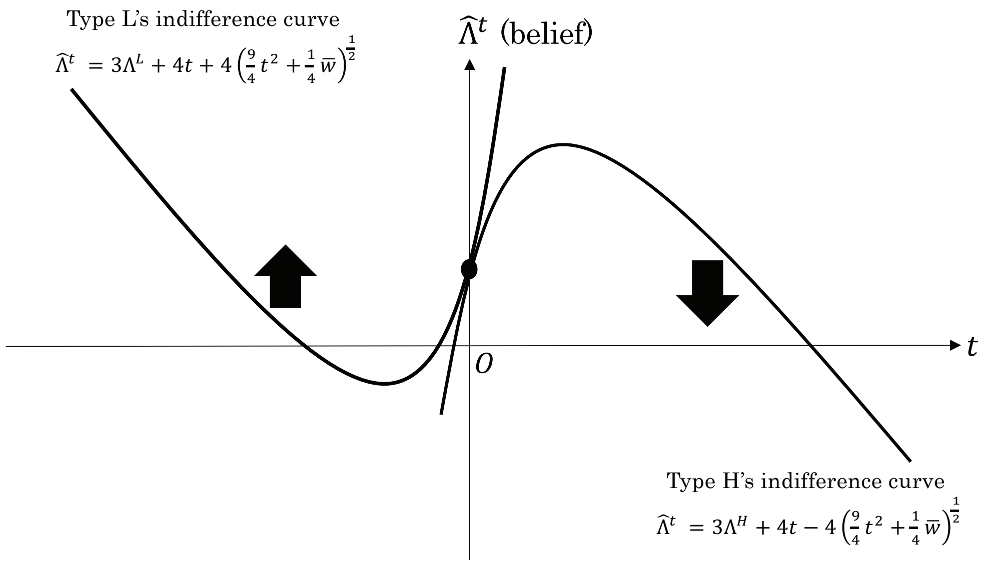

Figure B.2: Indifference curves on $t - \hat{\Lambda}^t$ plane

Lemma 4. *Suppose the situation G_1 is realized and $a_1^H = a_1^L = e$ holds in an equilibrium. Then, in the signaling game of tax choices, there is no separating equilibrium.*

Proof. Suppose to the contrary that there exists a separating equilibrium, and type H and L choose $t_1^H \neq 0$ and $t_1^L \neq 0$ ($t_1^H \neq t_1^L$) in the equilibrium, respectively. Then, the equilibrium payoffs of type H and L are $w_1(\Lambda^H, \Lambda^H, t_1^H)$ and $w_1(\Lambda^L, \Lambda^L, t_1^L)$, respectively. Now, we focus on the tax rate $t = 0$, which is off-the-equilibrium path. If $\hat{\Lambda}^0 \in (\tilde{\Lambda}^L, \Lambda^H]$ holds, where $\hat{\Lambda}^0$ is the belief after observing $t = 0$, by Lemma 3(i)(ii)(iii), it holds that $w_1(\Lambda^L, \hat{\Lambda}^0, 0) = \frac{9}{4}(\Lambda^L - \frac{1}{3}\hat{\Lambda}^0)^2 > \frac{9}{4}(\Lambda^L - \frac{1}{3}\tilde{\Lambda}^L)^2 = w_1(\Lambda^L, \tilde{\Lambda}^L, 0) = \max_t w_1(\Lambda^L, \Lambda^L, t) \geq w_1(\Lambda^L, \Lambda^L, t_1^L)$.[48] This implies that $\hat{\Lambda}^0 \in [\Lambda^L, \tilde{\Lambda}^L]$ must hold in the separating equilibrium, since type L has an incentive to deviate to $t_1^L = 0$ under $\hat{\Lambda}^0 \in (\tilde{\Lambda}^L, \Lambda^H]$. Similarly, if $\hat{\Lambda}^0 \in [\Lambda^L, \tilde{\Lambda}^H)$ holds, by Lemma 3(i)(ii)(iii), it holds that $w_1(\Lambda^H, \hat{\Lambda}^0, 0) = \frac{9}{4}(\Lambda^H - \frac{1}{3}\hat{\Lambda}^0)^2 > \frac{9}{4}(\Lambda^H - \frac{1}{3}\tilde{\Lambda}^H)^2 = w_1(\Lambda^H, \tilde{\Lambda}^H, 0) = \max_t w_1(\Lambda^H, \Lambda^H, t) \geq w_1(\Lambda^H, \Lambda^H, t_1^H)$. This implies that $\hat{\Lambda}^0 \in [\tilde{\Lambda}^H, \Lambda^H]$ must hold, since type H has an incentive to deviate to $t_1^H = 0$ under $\hat{\Lambda}^0 \in [\Lambda^L, \tilde{\Lambda}^H)$. However, this contradicts $\hat{\Lambda}^0 \in [\Lambda^L, \tilde{\Lambda}^L]$, since $\tilde{\Lambda}^H > \tilde{\Lambda}^L$.

From the above, it must hold $t_1^H = 0$ or $t_1^L = 0$ in the equilibrium. Suppose $t_1^L = 0$. Then, for type H, it holds that $w_1(\Lambda^H, \Lambda^L, 0) = \frac{9}{4}(\Lambda^H - \frac{1}{3}\Lambda^L)^2 > \frac{81}{20}(\Lambda^H - \frac{1}{3}\Lambda^H)^2 = \max_t w_1(\Lambda^H, \Lambda^H, t) \geq$

[48]Note that $\frac{9}{4}(\Lambda^L - \frac{1}{3}\hat{\Lambda}^0)^2 > \frac{9}{4}(\Lambda^L - \frac{1}{3}\tilde{\Lambda}^L)^2$ holds, since $\frac{\partial w_1(\Lambda^L, \Lambda, 0)}{\partial \Lambda} > 0$.

$w_1(\Lambda^H, \Lambda^H, t_1^H)$.[49] This implies that $t_1^L \neq 0$ must hold in the separating equilibrium because of the deviation of type H. Thus, it must hold $t_1^H = 0$. However, for type L, we can see $w_1(\Lambda^L, \Lambda^H, 0) = \frac{9}{4}(\Lambda^L - \frac{1}{3}\Lambda^H)^2 > \frac{81}{20}(\Lambda^L - \frac{1}{3}\Lambda^L)^2 \geq w_1(\Lambda^L, \Lambda^L, t_1^L)$. Therefore, the separating equilibrium cannot be supported by any tax rates, and this is a contradiction. □

In the first part of the proof, we show that there exists no separating equilibrium such that type H and L choose $t_1^H \neq 0$ and $t_1^L \neq 0$, respectively. We show that there always exists a type who wants to deviate to $t = 0$ for any beliefs $\hat{\Lambda}^0$. In figure B.3, the indifference curves are drawn, which describe their maximum utility levels under separating equilibria,[50] as well as the sets of beliefs under which they want to deviate to $t = 0$. We can see that the sets of beliefs overlap each other in figure B.3, which implies that there is no belief $\hat{\Lambda}^0$ such that both types never deviate. This result can be attributed to the fact that their indifference curves are tangent to each other at $t = 0$.

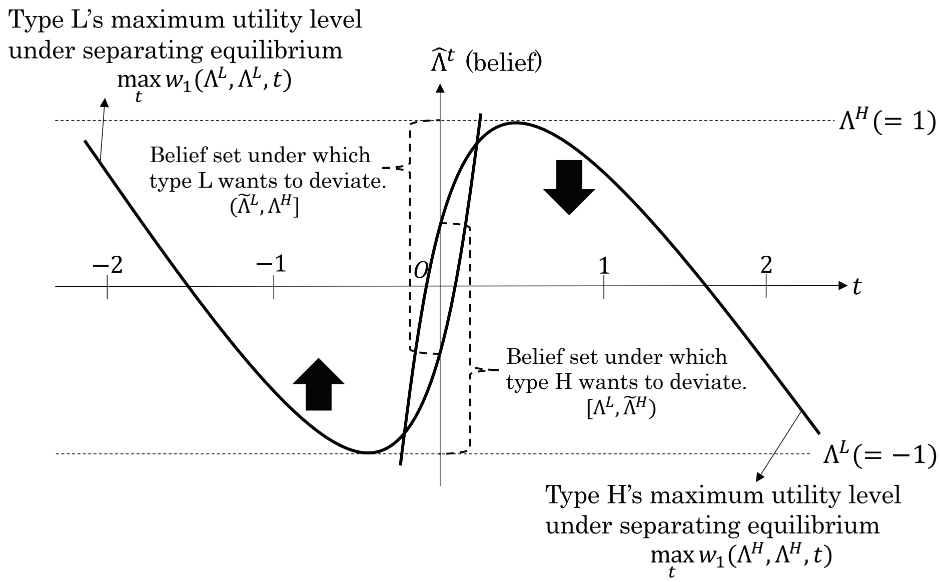

Figure B.3: No separating equilibrium such that type H and L choose $t_1^H \neq 0$ and $t_1^L \neq 0$ (First part of the proof for Lemma 4): An example of $\Lambda^H = 1$, $\Lambda^L = -1$, and $\overline{\Lambda} = 0$

Next, in the second part of the proof, we show that there exists no separating equilibrium such that $t_1^H = 0$ or $t_1^L = 0$. We show that the type that does not choose $t = 0$ wants to deviate to $t = 0$. Figure B.4 illustrates an example of a separating equilibrium with $t_1^H \neq 0$ and $t_1^L = 0$, where

[49]Note that $\frac{81}{20}(\Lambda^H - \frac{1}{3}\Lambda^H)^2 = \frac{9}{5}(\Lambda^H)^2$ and $\Lambda^L < 0$.

[50]The belief must be of the real type in a separating equilibrium, and indifference curves of type H and L must pass through the horizontal lines Λ^H and Λ^L at the chosen tax rates, respectively. Thus, we can see that each maximum utility level under a separating equilibrium is attained when each indifference curve is tangent to the horizontal line of the real type.

type H's indifference curve describes the maximum utility level under separating equilibria. In this example, $(t, \hat{\Lambda}^t) = (0, \Lambda^L)$ is available for type H. We can see that type H has an incentive to deviate to $t = 0$. A similar argument can be applied to the case where $t_1^H = 0$ and $t_1^L \neq 0$.

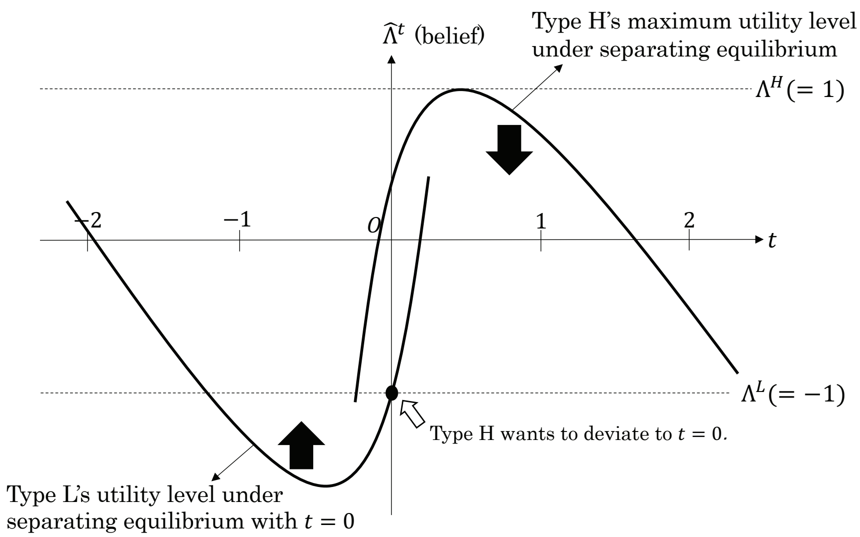

Figure B.4: No separating equilibrium with $t_1^L = 0$ (Second part of the proof for Lemma 4): An example of $\Lambda^H = 1$, $\Lambda^L = -1$, and $\overline{\Lambda} = 0$

Lemma 5. *Suppose the situation G_1 is realized and $a_1^H = a_1^L = e$ holds in an equilibrium. Let t_p be a tax rate attained under a pooling equilibrium. Then, $t_p = 0$ must hold.*

Proof. We consider two cases: $t_p \in (-\infty, 0)$ and $t_p \in (0, \infty)$, and only show that $t_p \in (-\infty, 0)$ is not chosen under any pooling equilibria. The proof in the case of $t_p \in (0, \infty)$ is similar.

First, suppose $t_p \in (-\infty, -\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda})]$. Then, it holds $w_1(\Lambda^H, \hat{\Lambda}^0, 0) = \frac{9}{4}(\Lambda^H - \frac{1}{3}\hat{\Lambda}^0)^2 > 0 \geq \frac{1}{4}(-2t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))(10t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda})) = w_1(\Lambda^H, \overline{\Lambda}, t_p)$ for any $\hat{\Lambda}^0$. This implies that $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$ must hold in a pooling equilibrium, since type H has an incentive to deviate to $t = 0$. Next, suppose $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$. Then, for a pooling equilibrium to exist, $w_1(\Lambda^H, \overline{\Lambda}, t_p) \geq w_1(\Lambda^H, \hat{\Lambda}^0, 0)$ must hold for some $\hat{\Lambda}^0$.[51] Now, let Λ_- satisfy $w_1(\Lambda^H, \overline{\Lambda}, t_p) = w_1(\Lambda^H, \Lambda_-, 0)$.[52] We have $w_1(\Lambda^H, \overline{\Lambda}, t_p) \geq w_1(\Lambda^H, \hat{\Lambda}^0, 0)$ if and only if $\hat{\Lambda}^0 \in [\Lambda_-, \Lambda^H]$.[53] Then, for

[51] That is, $-5(t_p)^2 + 6(\Lambda^H - \frac{1}{3}\overline{\Lambda})t_p + \frac{9}{4}(\Lambda^H - \frac{1}{3}\overline{\Lambda})^2 \geq \frac{9}{4}(\Lambda^H - \frac{1}{3}\hat{\Lambda}^0)^2$.

[52] The value Λ_- exists. If t_p is a pooling-equilibrium tax rate, we can find $\hat{\Lambda}^0$ such that $w_1(\Lambda^H, \overline{\Lambda}, t_p) \geq w_1(\Lambda^H, \hat{\Lambda}^0, 0)$. Since $w_1(\Lambda^H, \Lambda^L, 0) > w_1(\Lambda^H, \overline{\Lambda}, t_p)$ holds for any $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$, the continuity of $w_1(\Lambda^H, \hat{\Lambda}^0, 0)$ implies the existence of Λ_-.

[53] Note that $w_1(\Lambda^H, \hat{\Lambda}^0, 0)$ decreases in $\hat{\Lambda}^0$.

any $\hat{\Lambda}^0 \in [\Lambda_-, \Lambda^H]$, we can obtain $w_1(\Lambda^L, \hat{\Lambda}^0, 0) - w_1(\Lambda^L, \overline{\Lambda}, t_p) \geq \frac{3}{2}(\Lambda^H - \Lambda^L)(\Lambda_- - \overline{\Lambda} + 4t_p)$.[54] Here, let $F(t_p) \equiv \Lambda_- - \overline{\Lambda} + 4t_p$. Then, it holds $F(0) = 0$ and $F'(t_p) < 0$ for any $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$.[55] This implies $F(t_p) > 0$ for any $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$, so that $w_1(\Lambda^L, \hat{\Lambda}^0, 0) - w_1(\Lambda^L, \overline{\Lambda}, t_p) > 0$ holds for any $\hat{\Lambda}^0 \in [\Lambda_-, \Lambda^H]$ and $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$. However, this contradicts $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$ being a tax rate in the pooling equilibrium. □

In the second part of the proof, we show that there is no belief $\hat{\Lambda}^0$ under which both type H and L do not deviate from $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$ to $t = 0$. Especially, we drive the set of beliefs under which type H never deviates to $t = 0$ (that is, $[\Lambda_-, \Lambda^H]$), and show that type L always deviates to $t = 0$ for all $\hat{\Lambda}^0 \in [\Lambda_-, \Lambda^H]$. This is summarized as Figure B.5.

Similarly, we can show that $t_p \in (0, -\frac{3}{10}(\Lambda^L - \frac{1}{3}\overline{\Lambda}))$ is not chosen in a pooling equilibrium. As illustrated in Figure B.6, we can show that type H always deviates to $t = 0$ for any belief under which type L never deviates to $t = 0$ (that is, $[\Lambda^L, \Lambda_+]$).

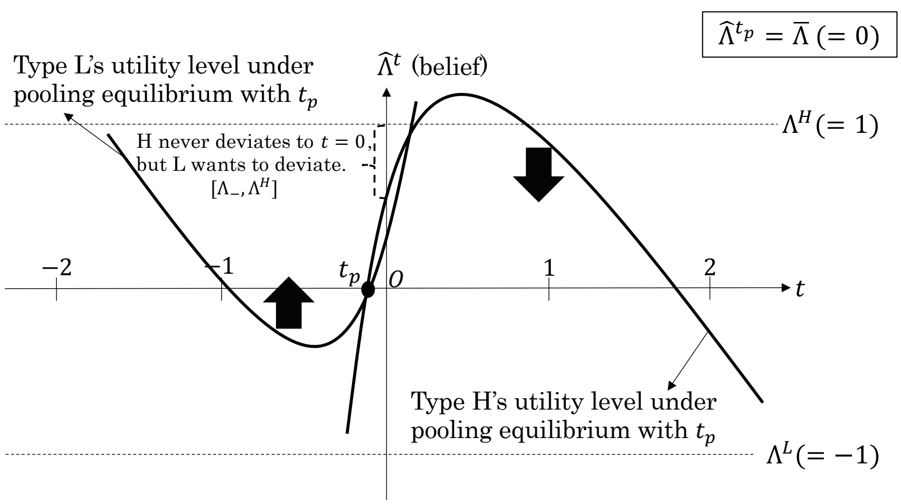

Figure B.5: No pooling equilibrium with $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$ (Second part of the proof for Lemma 5): An example of $\Lambda^H = 1$, $\Lambda^L = -1$, and $\overline{\Lambda} = 0$

[54] Since $\hat{\Lambda}^0 \geq \Lambda_-$, $w_1(\Lambda^L, \hat{\Lambda}^0, 0) - w_1(\Lambda^L, \overline{\Lambda}, t_p) = \frac{9}{4}(\Lambda^L - \frac{1}{3}\hat{\Lambda}^0)^2 + 5(t_p)^2 - 6(\Lambda^L - \frac{1}{3}\overline{\Lambda})t_p - \frac{9}{4}(\Lambda^L - \frac{1}{3}\overline{\Lambda})^2 \geq \frac{9}{4}(\Lambda^L - \frac{1}{3}\Lambda_-)^2 + 5(t_p)^2 - 6(\Lambda^L - \frac{1}{3}\overline{\Lambda})t_p - \frac{9}{4}(\Lambda^L - \frac{1}{3}\overline{\Lambda})^2$. By the definition of Λ_-, $5(t_p)^2 + 2\overline{\Lambda}t_p = 6\Lambda^H t_p + \frac{9}{4}(\Lambda^H - \frac{1}{3}\overline{\Lambda})^2 - \frac{9}{4}(\Lambda^H - \frac{1}{3}\Lambda_-)^2$. Using this, $w_1(\Lambda^L, \hat{\Lambda}^0, 0) - w_1(\Lambda^L, \overline{\Lambda}, t_p) \geq \frac{9}{4}(\Lambda^L - \frac{1}{3}\Lambda_-)^2 - \frac{9}{4}(\Lambda^L - \frac{1}{3}\overline{\Lambda})^2 + \frac{9}{4}(\Lambda^H - \frac{1}{3}\overline{\Lambda})^2 - \frac{9}{4}(\Lambda^H - \frac{1}{3}\Lambda_-)^2 + 6\Lambda^H t_p - 6\Lambda^L t_p = \frac{3}{2}(\Lambda^H - \Lambda^L)(\Lambda_- - \overline{\Lambda} + 4t_p)$.

[55] We can see $\Lambda_- = 3\Lambda^H - \{(-2t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))(10t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))\}^{\frac{1}{2}}$. Thus, $F'(t_p) = \frac{\partial \Lambda_-}{\partial t_p} + 4 = 4 - 4(-5t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda})\{(-2t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))(10t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))\}^{-\frac{1}{2}} = 4 - 4\{(-5t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))^2\}^{\frac{1}{2}}\{(-2t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))(10t_p + 3(\Lambda^H - \frac{1}{3}\overline{\Lambda}))\}^{-\frac{1}{2}} < 0$ for any $t_p \in (-\frac{3}{10}(\Lambda^H - \frac{1}{3}\overline{\Lambda}), 0)$.

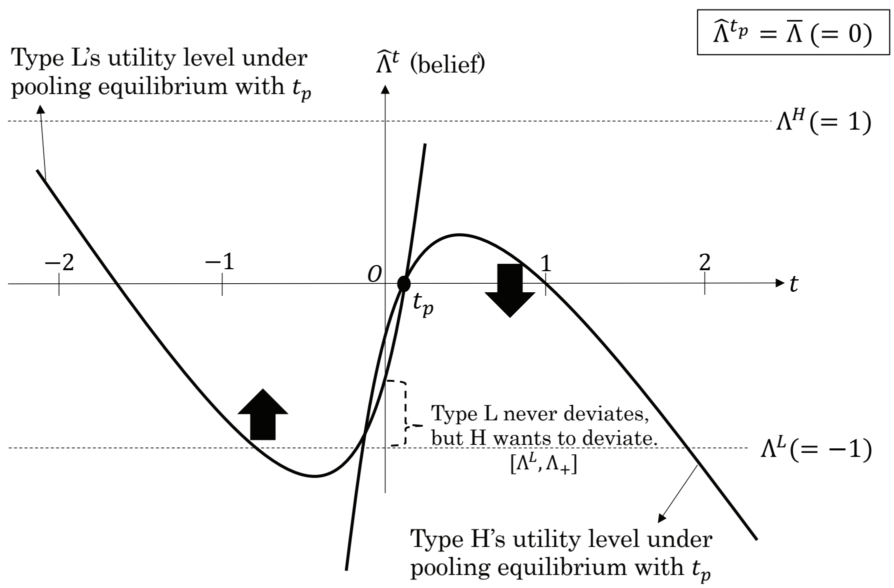

Figure B.6: No pooling equilibrium with $t_p \in (0, -\frac{3}{10}(\Lambda^L - \frac{1}{3}\overline{\Lambda}))$: An example of $\Lambda^H = 1$, $\Lambda^L = -1$, and $\overline{\Lambda} = 0$

Lemma 6. *Suppose the situation G_1 is realized and $a_1^H = a_1^L = e$ holds in an equilibrium. Then, if a hybrid equilibrium exists, it must hold that (i) type H mixes 0 and $\frac{2}{5}\Lambda^H$ and type L takes 0 surely or (ii) type H takes 0 surely and type L mixes 0 and $\frac{2}{5}\Lambda^L$.*

Proof. Consider the case that type H mixes t_p and t_h, and type L takes t_p with probability 1. Suppose to the contrary that type H takes $t_h \neq \frac{2}{5}\Lambda^H$ in the hybrid equilibrium. If type H chooses t_h, its type is revealed, and type H's utility is $w_1(\Lambda^H, \Lambda^H, t_h)$. However, we can see $w_1(\Lambda^H, \hat{\Lambda}, \frac{2}{5}\Lambda^H) \geq w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H) > w_1(\Lambda^H, \Lambda^H, t_h)$ for any belief $\hat{\Lambda}$, which is a contradiction.[56] Thus, it must hold $t_h = \frac{2}{5}\Lambda^H$ in the hybrid equilibrium.

Now, suppose type H takes t_p with probability q and $\frac{2}{5}\Lambda^H$ with probability $1-q$, and type L takes t_p in the equilibrium. By the definition of PBE (Definition 1), it must hold $w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H) = w_1(\Lambda^H, \hat{\Lambda}^{t_p}, t_p)$.[57] Thus, we can see that type H's expected utility in the hybrid equilibrium is equal to $w_1(\Lambda^H, \hat{\Lambda}^{t_p}, t_p)$. Moreover, type L's utility is $w_1(\Lambda^L, \hat{\Lambda}^{t_p}, t_p)$. Thus, we can apply the same argument as Lemma 5, so that $t_p = 0$ must hold in the equilibrium.[58]

[56]For the first equality, we can check $\frac{\partial w_1(\Lambda^H, \hat{\Lambda}, \frac{2}{5}\Lambda^H)}{\partial \hat{\Lambda}} < 0$ for any $\hat{\Lambda} \in (\Lambda^L, \Lambda^H)$. For the second inequality, note that $w_1(\Lambda^H, \Lambda^H, t)$ is uniquely maximized at $t = \frac{2}{5}\Lambda^H$.

[57]If $w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H) \neq w_1(\Lambda^H, \hat{\Lambda}^{t_p}, t_p)$ holds, type H wants to deviate from the mixed action strategy q. Note that $\hat{\Lambda}^{t_p}$ is calculated using Bayes' rule as $\hat{\Lambda}^{t_p} = \frac{q\overline{\rho}_H}{q\overline{\rho}_H + \overline{\rho}_L}\Lambda^H + \frac{\overline{\rho}_L}{q\overline{\rho}_H + \overline{\rho}_L}\Lambda^L$.

[58]When we consider $\hat{\Lambda}^{t_p}$ as $\overline{\Lambda}$ in Lemma 5, the same proof can be applied.

In the case that type H takes t_p with prob 1 and type L mixes t_l and t_p, we can apply the same proof as above. □

Proof of Lemma 1: We show the following: (i) if $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$ holds, there exists a pooling equilibrium where both types take 0; (ii) if $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$ holds, there exists a hybrid equilibrium where type H mixes 0 and $\frac{2}{5}\Lambda^H$, and type L takes 0; (iii) if $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ holds, there exists a hybrid equilibrium where type H takes 0, and type L mixes 0 and $\frac{2}{5}\Lambda^L$; and (iv) the equilibrium outcome is unique for each $\overline{\Lambda}$.

(i: Derivation of pooling equilibrium)

Assume $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$. We will check that both types never deviate from $t_1^H = t_1^L = 0$ under the following belief:

$$\hat{\Lambda}^t \equiv \begin{cases} \Lambda^L & if \quad t \in (-\infty, \frac{1}{4}(\Lambda^L - \overline{\Lambda})] \\ 4t + \overline{\Lambda} & if \quad t \in (\frac{1}{4}(\Lambda^L - \overline{\Lambda}), \frac{1}{4}(\Lambda^H - \overline{\Lambda})) \\ \Lambda^H & if \quad t \in [\frac{1}{4}(\Lambda^H - \overline{\Lambda}), \infty). \end{cases} \tag{B.3}$$

Note that this is a belief, that is, $\hat{\Lambda}^t \in [\Lambda^L, \Lambda^H]$ for all t.

First, we consider the interval $(\frac{1}{4}(\Lambda^L - \overline{\Lambda}), \frac{1}{4}(\Lambda^H - \overline{\Lambda}))$. Then, $\hat{\Lambda}^t = 4t + \overline{\Lambda}$ holds. We show that the tax rate $t_1^H = t_1^L = 0$ maximize each utility on $(\frac{1}{4}(\Lambda^L - \overline{\Lambda}), \frac{1}{4}(\Lambda^H - \overline{\Lambda}))$. The first-order conditions for utility maximization are

$$\frac{\partial w_1(\Lambda^s, \hat{\Lambda}^t, t)}{\partial t} = -10t + 6\left(\Lambda^s - \frac{1}{3}\hat{\Lambda}^t\right) - 2\frac{d\hat{\Lambda}^t}{dt}t - \frac{3}{2}\frac{d\hat{\Lambda}^t}{dt}\left(\Lambda^s - \frac{1}{3}\hat{\Lambda}^t\right) = 0 \quad \forall s,$$

where $\hat{\Lambda}^t = 4t + \overline{\Lambda}$ and $\frac{d\hat{\Lambda}^t}{dt} = 4$. Then, $t = 0$ satisfies the first-order condition for both types. We can check that the second-order conditions are also satisfied. Thus, both types have no incentive to deviate from $t = 0$ to $t \in (\frac{1}{4}(\Lambda^L - \overline{\Lambda}), \frac{1}{4}(\Lambda^H - \overline{\Lambda}))$.

Second, we consider the interval $(-\infty, \frac{1}{4}(\Lambda^L - \overline{\Lambda})]$. Then, $\hat{\Lambda}^t = \Lambda^L$. For type H, the tax rate maximizing its utility on $(-\infty, \frac{1}{4}(\Lambda^L - \overline{\Lambda})]$ is $\frac{1}{4}(\Lambda^L - \overline{\Lambda})$, since $w_1(\Lambda^H, \Lambda^L, t)$ is increasing on $(-\infty, \frac{1}{4}(\Lambda^L - \overline{\Lambda})]$. We can see $w_1(\Lambda^H, \overline{\Lambda}, 0) = \frac{9}{4}\left(\Lambda^H - \frac{1}{3}\overline{\Lambda}\right)^2 \geq \frac{9}{4}((\Lambda^H - \frac{1}{3}\overline{\Lambda}) - \frac{1}{2}(\overline{\Lambda} - \Lambda^L))((\Lambda^H - \frac{1}{3}\overline{\Lambda}) + \frac{1}{2}(\overline{\Lambda} - \Lambda^L)) = w_1(\Lambda^H, \Lambda^L, \frac{1}{4}(\Lambda^L - \overline{\Lambda}))$ for any $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$.[59] For type L, its utility is $w_1(\Lambda^L, \Lambda^L, t)$, and by Lemma 3$(ii)$ the maximum utility on $(-\infty, \frac{1}{4}(\Lambda^L - \overline{\Lambda})]$ is at most $\frac{9}{5}(\Lambda^L)^2$.

[59] $w_1(\Lambda^H, \Lambda^L, \frac{1}{4}(\Lambda^L - \overline{\Lambda})) = \frac{1}{4}(-2 \cdot \frac{1}{4}(\Lambda^L - \overline{\Lambda}) + 3(\Lambda^H - \frac{1}{3}\Lambda^L))(10 \cdot \frac{1}{4}(\Lambda^L - \overline{\Lambda}) + 3(\Lambda^H - \frac{1}{3}\Lambda^L)) = \frac{9}{4}(\Lambda^H - \frac{1}{2}\Lambda^L + \frac{1}{6}\overline{\Lambda})(\Lambda^H + \frac{1}{2}\Lambda^L - \frac{5}{6}\overline{\Lambda}) = \frac{9}{4}((\Lambda^H - \frac{1}{3}\overline{\Lambda}) - \frac{1}{2}(\overline{\Lambda} - \Lambda^L))((\Lambda^H - \frac{1}{3}\overline{\Lambda}) + \frac{1}{2}(\overline{\Lambda} - \Lambda^L))$.

We can see $w_1(\Lambda^L, \overline{\Lambda}, 0) = \frac{9}{4}\left(\Lambda^L - \frac{1}{3}\overline{\Lambda}\right)^2 \geq \frac{9}{5}(\Lambda^L)^2 \geq w_1(\Lambda^L, \Lambda^L, t)$ for any $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$.[60] Thus, both types have no incentive to deviate from $t = 0$ to $t \in (-\infty, \frac{1}{4}(\Lambda^L - \overline{\Lambda})]$.

Finally, we consider the interval $[\frac{1}{4}(\Lambda^H - \overline{\Lambda}), \infty)$. Then, $\hat{\Lambda}^t = \Lambda^H$. For type H, since its utility is $w_1(\Lambda^H, \Lambda^H, t)$, the maximum utility on $[\frac{1}{4}(\Lambda^H - \overline{\Lambda}), \infty)$ is at most $\frac{9}{5}(\Lambda^H)^2$. We can see $w_1(\Lambda^H, \overline{\Lambda}, 0) = \frac{9}{4}\left(\Lambda^H - \frac{1}{3}\overline{\Lambda}\right)^2 \geq \frac{9}{5}(\Lambda^H)^2 \geq w_1(\Lambda^H, \Lambda^H, t)$ for any $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$. For type L, the tax rate maximizing its utility on $[\frac{1}{4}(\Lambda^H - \overline{\Lambda}), \infty)$ is $\frac{1}{4}(\Lambda^H - \overline{\Lambda})$, since $w_1(\Lambda^L, \Lambda^H, t)$ is decreasing on $[\frac{1}{4}(\Lambda^H - \overline{\Lambda}), \infty)$. We can see $w_1(\Lambda^L, \overline{\Lambda}, 0) = \frac{9}{4}\left(\Lambda^L - \frac{1}{3}\overline{\Lambda}\right)^2 \geq \frac{9}{4}((\Lambda^L - \frac{1}{3}\overline{\Lambda}) - \frac{1}{2}(\overline{\Lambda} - \Lambda^H))((\Lambda^L - \frac{1}{3}\overline{\Lambda}) + \frac{1}{2}(\overline{\Lambda} - \Lambda^H)) = w_1(\Lambda^L, \Lambda^H, \frac{1}{4}(\Lambda^H - \overline{\Lambda}))$ for any $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$. Thus, both types have no incentive to deviate from $t = 0$ to $t \in [\frac{1}{4}(\Lambda^H - \overline{\Lambda}), \infty)$.

(ii: Derivation of H-hybrid equilibrium)

Assume $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$. We consider the action strategies that type L chooses 0, and type H chooses 0 with probability $\tilde{q}$ and $\frac{2}{5}\Lambda^H$ with probability $1 - \tilde{q}$. We consider $\tilde{q}$ such that $\frac{\tilde{q}\overline{\rho}_H}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^H + \frac{\overline{\rho}_L}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^L = \tilde{\Lambda}^H$ is satisfied, which means $\hat{\Lambda}^0 = \tilde{\Lambda}^H$.[61] We define the consistent belief with the above equilibrium strategy profile, as follows:

$$\hat{\Lambda}^t \equiv \begin{cases} \Lambda^L & if \quad t \in (-\infty, \frac{1}{4}(\Lambda^L - \tilde{\Lambda}^H)] \\ 4t + \tilde{\Lambda}^H & if \quad t \in (\frac{1}{4}(\Lambda^L - \tilde{\Lambda}^H), \frac{1}{4}(\Lambda^H - \tilde{\Lambda}^H)) \\ \Lambda^H & if \quad t \in [\frac{1}{4}(\Lambda^H - \tilde{\Lambda}^H), \infty). \end{cases} \tag{B.4}$$

Let us check whether both types deviate from the strategy profile, given the belief ($B.4$). For type H, in the equilibrium, the utilities from $t = 0$ and $t = \frac{2}{5}\Lambda^H$ are $w_1(\Lambda^H, \tilde{\Lambda}^H, 0)$ and $w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H)$, respectively. $\tilde{q}$ is an optimal mixed strategy between 0 and $\frac{2}{5}\Lambda^H$, since it holds $w_1(\Lambda^H, \tilde{\Lambda}^H, 0) = w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H)$ by Lemma 3.[62] Since the expected utility of type H in the equilibrium is $w_1(\Lambda^H, \tilde{\Lambda}^H, 0)$, and the utility of type L in the equilibrium is $w_1(\Lambda^L, \tilde{\Lambda}^H, 0)$, we can apply the same argument as **(i: Derivation of pooling equilibrium)** to show that both types have no incentive to deviate from the strategy profile.[63]

Finally, we should show that there exists $\tilde{q}$ such that $\frac{\tilde{q}\overline{\rho}_H}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^H + \frac{\overline{\rho}_L}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^L = \tilde{\Lambda}^H$ when $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$. Let $H(q) \equiv \frac{q\overline{\rho}_H}{q\overline{\rho}_H + \overline{\rho}_L}\Lambda^H + \frac{\overline{\rho}_L}{q\overline{\rho}_H + \overline{\rho}_L}\Lambda^L - \tilde{\Lambda}^H$. Then, we can see $H(0) = \Lambda^L - \tilde{\Lambda}^H < 0$,

[60] For any $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$, $\frac{9}{4}\left(\Lambda^L - \frac{1}{3}\overline{\Lambda}\right)^2 \geq \frac{9}{4}\left(\Lambda^L - \frac{1}{3}\tilde{\Lambda}^L\right)^2 = \frac{9}{4}\left(\Lambda^L - \frac{1}{3} \cdot 3(1 - \frac{2\sqrt{5}}{5})\Lambda^L\right)^2 = \frac{9}{5}(\Lambda^L)^2$.

[61] The belief $\hat{\Lambda}^0$ must be calculated using Bayes' rule. In this case, $\hat{\Lambda}^0 = \frac{\tilde{q}\overline{\rho}_H}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^H + \frac{\overline{\rho}_L}{\tilde{q}\overline{\rho}_H + \overline{\rho}_L}\Lambda^L$.

[62] If it holds, type H does not have an incentive to choose other probabilities.

[63] The only difference from the case (i) is that $t = \frac{2}{5}\Lambda^H$ is on-the-equilibrium path. However, it does not matter since $\hat{\Lambda}^{\frac{2}{5}\Lambda^H} = \Lambda^H$ in both cases.

and $H(1) = \overline{\Lambda} - \tilde{\Lambda}^H > 0$. Since $H(q)$ is continuous in q and $H'(q) > 0$ holds, there exists $\tilde{q}$ such that $H(\tilde{q}) = 0$ uniquely.

(iii: Derivation of L-hybrid equilibrium)

Suppose $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$. We consider the action strategies that type H chooses 0, and type L chooses 0 with probability $\tilde{p}$ and $\frac{2}{5}\Lambda^L$ with probability $1-\tilde{p}$. We consider $\tilde{p}$ such that $\frac{\overline{\rho}_H}{\overline{\rho}_H + \tilde{p}\overline{\rho}_L}\Lambda^H + \frac{\tilde{p}\overline{\rho}_L}{\overline{\rho}_H + \tilde{p}\overline{\rho}_L}\Lambda^L = \tilde{\Lambda}^L$ is satisfied, which implies $\hat{\Lambda}^0 = \tilde{\Lambda}^L$. We define the consistent belief with the above equilibrium strategy profile, as follows:

$$\hat{\Lambda}^t \equiv \begin{cases} \Lambda^L & if \quad t \in (-\infty, \frac{1}{4}(\Lambda^L - \tilde{\Lambda}^L)] \\ 4t + \tilde{\Lambda}^L & if \quad t \in (\frac{1}{4}(\Lambda^L - \tilde{\Lambda}^L), \frac{1}{4}(\Lambda^H - \tilde{\Lambda}^L)) \\ \Lambda^H & if \quad t \in [\frac{1}{4}(\Lambda^H - \tilde{\Lambda}^L), \infty). \end{cases} \tag{B.5}$$

Then, the same argument as the H-hybrid equilibrium is applied to show that no types deviate from the strategy profile.

(iv: Uniqueness)

First, we consider the case of $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$. In this case, by Lemmas 4 and 5, it suffices to show that there is no hybrid equilibrium. Suppose to the contrary that there exists a hybrid equilibrium. Then, by Lemma 6, either type H or L mixes 0 and $\frac{2}{5}\Lambda^s$. Here, suppose type H mixes 0 and $\frac{2}{5}\Lambda^H$ in the hybrid equilibrium. Then, it must hold that $w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H) = w_1(\Lambda^H, \hat{\Lambda}^0, 0)$ and thus, by Lemma 3(iii), $\hat{\Lambda}^0 = \tilde{\Lambda}^H$. Moreover, since the posterior belief $\hat{\Lambda}^0$ must satisfy $\overline{\Lambda} > \hat{\Lambda}^0$,[64] we have $\overline{\Lambda} > \tilde{\Lambda}^H$. However, this is a contradiction to $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$. We can apply the same argument to type L.

Second, we consider the case of $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$. In this case, by Lemmas 4, 5, and 6, it suffices to show that there is no pooling equilibrium in which both types take 0, and no L-hybrid equilibrium in which type H takes 0 and type L mixes 0 and $\frac{2}{5}\Lambda^L$. Suppose the pooling equilibrium exists. By $\frac{\partial w_1(\Lambda^H, \hat{\Lambda}^t, \frac{2}{5}\Lambda^H)}{\partial \hat{\Lambda}^t} < 0$, Lemma 3($iii$) and $\overline{\Lambda} > \tilde{\Lambda}^H$, it holds that $w_1(\Lambda^H, \hat{\Lambda}^t, \frac{2}{5}\Lambda^H) \geq w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H) = w_1(\Lambda^H, \tilde{\Lambda}^H, 0) > w_1(\Lambda^H, \overline{\Lambda}, 0)$ for any $\hat{\Lambda}^t$. Thus, type H has an incentive to deviate from $t = 0$, and this is a contradiction. Suppose the L-hybrid equilibrium exists. It must hold that $w_1(\Lambda^L, \Lambda^L, \frac{2}{5}\Lambda^L) = w_1(\Lambda^L, \hat{\Lambda}^0, 0)$ in the equilibrium, and thus $\hat{\Lambda}^0 = \tilde{\Lambda}^L$ holds. Moreover, the posterior belief $\hat{\Lambda}^0$ must satisfy $\overline{\Lambda} < \hat{\Lambda}^0$, so that we have $\overline{\Lambda} < \tilde{\Lambda}^L$. However, this is a contradiction

[64] When type H chooses 0 with prob q and $\frac{2}{5}\Lambda^H$ with prob $1-q$, it holds $\hat{\Lambda}^0 = \frac{q\overline{\rho}_H}{q\overline{\rho}_H + \overline{\rho}_L}\Lambda^H + \frac{\overline{\rho}_L}{q\overline{\rho}_H + \overline{\rho}_L}\Lambda^L$. We can see $\overline{\Lambda} > \hat{\Lambda}^0$ for all $q \in (0, 1)$, since $\hat{\Lambda}^0$ is increasing in q and $\hat{\Lambda}^0 = \overline{\Lambda}$ when $q = 1$.

to $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$.

The proof in the case of $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ is similar to the second one. □

In the derivation of the pooling equilibrium, we consider a specific belief (the belief ($B.3$)), and show that each type has no incentive to deviate from the equilibrium under the belief. We can illustrate an example of the pooling equilibrium as Figure B.7.

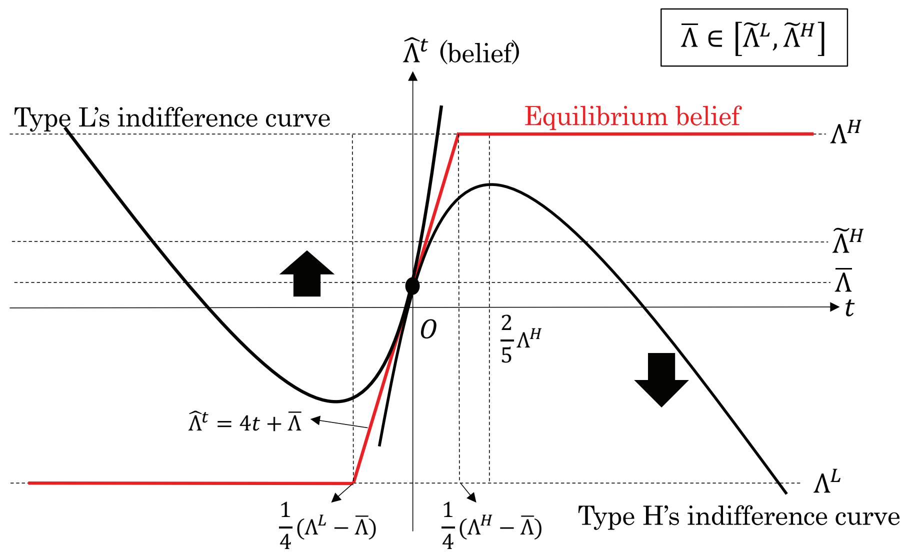

Figure B.7: An example of pooling equilibrium with $t = 0$

When $\overline{\Lambda}$ is sufficiently large or small, the pooling equilibrium does not exist. For example, let us consider the case that $\overline{\Lambda}$ is sufficiently large, that is, the case $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$. An example is drawn as Figure B.8.

In this case, there is no belief under which both types do not deviate from $t = 0$. For example, in Figure B.8, type H has an incentive to deviate to $t = \frac{2}{5}\Lambda^H$ for any belief, and this situation holds whenever $\overline{\Lambda} > \tilde{\Lambda}^H$. We can see that, for the pooling actions at $t = 0$ to be sustained, the belief $\hat{\Lambda}^0$ must be less than $\tilde{\Lambda}^H$.

Here, if type H increases the probability of choosing $t = \frac{2}{5}\Lambda^H$, it would increase the probability that the type choosing $t = 0$ is type L, so that $\hat{\Lambda}^0$ decreases (to $\tilde{\Lambda}^H$), and H-hybrid equilibrium can emerge as illustrated in Figure B.9. The same logic can be applied to L-hybrid equilibrium when $\overline{\Lambda}$ is sufficiently small, that is, $\overline{\Lambda} < \tilde{\Lambda}^L$.

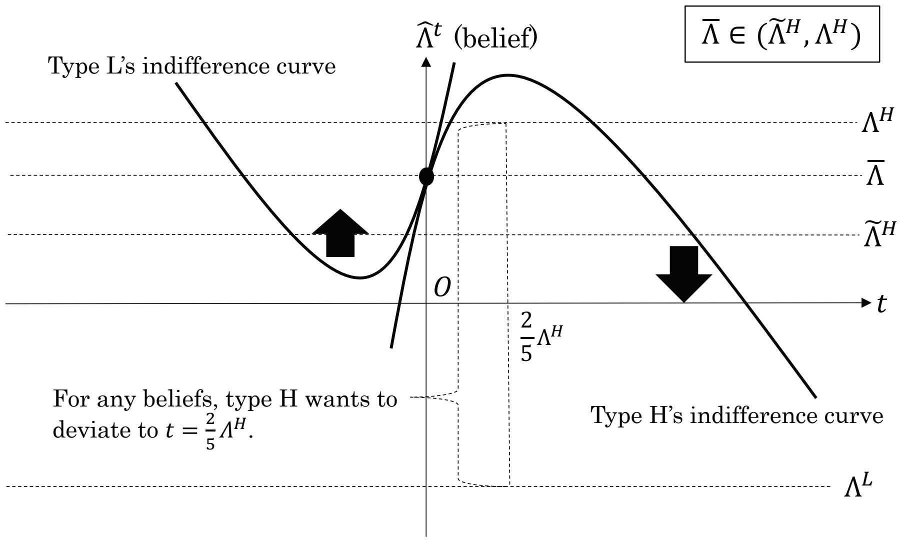

Figure B.8: An example of no pooling equilibrium with $t = 0$

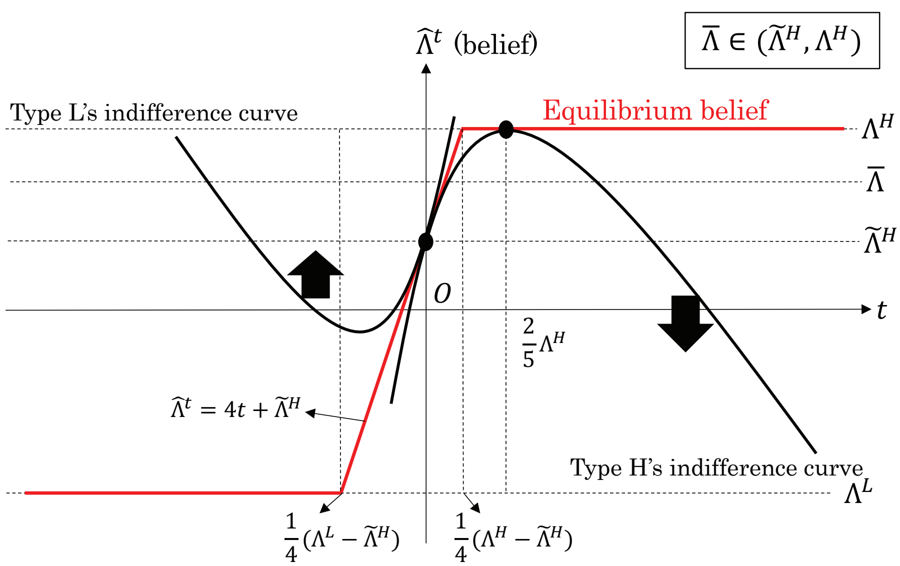

Figure B.9: An example of H-hybrid equilibrium

C Proof of Proposition 2

Proof. First, suppose to the contrary that the profile (e, l, e) is an equilibrium outcome. Then, since the type is fully revealed ($e \to H$ and $l \to L$), by Lemma 2, type H's utility is $v_1(e, l, e|H) = 3\left(\Lambda^H - \frac{1}{4}\Lambda^H\right)^2$. However, if type H deviates to l, then its utility is $v_1(l, l, e|H) = 3\left(\Lambda^H - \frac{2}{5}\Lambda^L\right)^2$. Since $v_1(l, l, e|H) > v_1(e, l, e|H)$ holds, type H has an incentive to deviate to l, and this is a contradiction.

Second, suppose the profile (l, e, e) is an equilibrium outcome. Then, since the type is fully revealed ($e \to L$ and $l \to H$), type L's utility is $v_1(l, e, e|L) = 3\left(\Lambda^L - \frac{1}{4}\Lambda^L\right)^2$. However, if type L deviates to l, then its utility is $v_1(l, l, e|L) = 3\left(\Lambda^L - \frac{2}{5}\Lambda^H\right)^2$. Since $v_1(l, l, e|L) > v_1(l, e, e|L)$ holds, type L has an incentive to deviate to l, and this is a contradiction.

Finally, suppose the profile (e, e, e) is an equilibrium outcome when $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$. Since the belief is not updated by the pooling behaviors,[65] by Lemma 2, country 2's expected utility is $\mathbb{E}v_2(e, e, e|s) = \sum_s \overline{\rho}_s \left[-5\left(\Lambda^s - \frac{1}{4}\overline{\Lambda}\right)^2 + 6\left(\Lambda^s - \frac{1}{4}\overline{\Lambda}\right)\Lambda^s\right] = \overline{(\Lambda^s)^2} + \frac{11}{16}(\overline{\Lambda})^2$, where $\overline{(\Lambda^s)^2} = \sum_s \overline{\rho}_s (\Lambda^s)^2$. If country 2 deviates to l, by Lemma 1, $t_1^H = t_1^L = 0$ holds.[66] Since country 2 never updates its belief, country 2's expected utility is $\mathbb{E}v_2(e, e, l|s) = \sum_s \overline{\rho}_s \left[-3\left(\frac{3}{2}\Lambda^s - \frac{1}{2}\overline{\Lambda}\right)^2 + 6\left(\frac{3}{2}\Lambda^s - \frac{1}{2}\overline{\Lambda}\right)\Lambda^s\right] = \frac{9}{4}\overline{(\Lambda^s)^2} + \frac{3}{4}(\overline{\Lambda})^2$. Then, we can see $\mathbb{E}v_2(e, e, l|s) > \mathbb{E}v_2(e, e, e|s)$ for any $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$, and this is a contradiction. □

D Proof of Proposition 3

Proof. Suppose to the contrary that the profile (e, e, l) is an equilibrium outcome. When $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$ holds, by Lemmas 1 and 2, it holds that $v_1(l, e, l|H) = 3(\Lambda^H - \frac{1}{4}\hat{\Lambda}^l)^2 > (\frac{3}{2}\Lambda^H - \frac{1}{2}\overline{\Lambda})^2 = v_1(e, e, l|H)$ if $\hat{\Lambda}^l < (4 - 2\sqrt{3})\Lambda^H + \frac{2\sqrt{3}}{3}\overline{\Lambda}$. Similarly, it holds that $v_1(e, l, l|L) = 3(\Lambda^L - \frac{1}{4}\hat{\Lambda}^l)^2 > (\frac{3}{2}\Lambda^L - \frac{1}{2}\overline{\Lambda})^2 = v_1(e, e, l|L)$ if $\hat{\Lambda}^l > (4 - 2\sqrt{3})\Lambda^L + \frac{2\sqrt{3}}{3}\overline{\Lambda}$. We can see that, for any off-the-equilibrium belief $\hat{\Lambda}^l$, either type has an incentive to deviate, and this is a contradiction. Next, when $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$, by Lemmas 1 and 2, it holds $v_1(e, l, l|L) = 3(\Lambda^L - \frac{1}{4}\hat{\Lambda}^l)^2 > (\frac{3}{2}\Lambda^L - \frac{1}{2}\tilde{\Lambda}^H)^2 = v_1(e, e, l|L)$ if $\hat{\Lambda}^l > (4 - 2\sqrt{3})\Lambda^L + \frac{2\sqrt{3}}{3}\tilde{\Lambda}^H$.[67] Now, we define $G(\hat{\Lambda}^l) \equiv v_1(e, e, l|H) - v_1(l, e, l|H) = \frac{9}{5}(\Lambda^H)^2 -$

[65] In other words, $\hat{\Lambda}^e = \overline{\Lambda}$.

[66] Note that $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$ holds.

[67] Note that $\tilde{\Lambda}^H$ is country 2's equilibrium belief after observing $t = 0$ when $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$. For the detail, see the proof of Lemma 1.

$3(\Lambda^H - \frac{1}{4}\hat{\Lambda}^l)^2$.[68] Since $G'(\hat{\Lambda}^l) > 0$ and $G((4-2\sqrt{3})\Lambda^L + \frac{2\sqrt{3}}{3}\tilde{\Lambda}^H) < 0$,[69] $G(\hat{\Lambda}^l) < 0$ holds for any $\hat{\Lambda}^l \leq (4-2\sqrt{3})\Lambda^L + \frac{2\sqrt{3}}{3}\tilde{\Lambda}^H$. Thus, for any off-the-equilibrium belief $\hat{\Lambda}^l$, either type has an incentive to deviate, and this is a contradiction. The proof of case $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ is similar to the latter one.

Suppose the profile (l, e, l) is an equilibrium outcome. Since the type is revealed ($e \to L$ and $l \to H$), we can see $v_1(l, e, l|L) = \frac{9}{5}(\Lambda^L)^2 < 3\left(\Lambda^L - \frac{1}{4}\Lambda^H\right)^2 = v_1(l, l, l|L)$. Thus, type L wants to deviate to l, and this is a contradiction.

Suppose the profile (e, l, l) is an equilibrium outcome. Since the type is revealed ($e \to H$ and $l \to L$), we can see $v_1(e, l, l|H) = \frac{9}{5}(\Lambda^H)^2 < 3(\Lambda^H - \frac{1}{4}\Lambda^L)^2 = v_1(l, l, l|H)$. Thus, type H wants to deviate to l, and this is a contradiction.

Suppose the profile (l, l, l) is an equilibrium outcome. Since the belief is not updated, it holds $\mathbb{E}v_2(l, l, l|s) = \sum_s \overline{\rho}_s \left(-5(\alpha^s)^2 + 6\Lambda^s\alpha^s\right) = \sum_s \overline{\rho}_s(\Lambda^s)^2 + \frac{11}{16}(\overline{\Lambda})^2 < \sum_s \overline{\rho}_s(\Lambda^s)^2 + \frac{4}{5}(\overline{\Lambda})^2 = \sum_s \overline{\rho}_s \left(-5(\beta^s)^2 + 6\Lambda^s\beta^s\right) = \mathbb{E}v_2(l, l, e|s)$, where $\alpha^s = \Lambda^s - \frac{1}{4}\overline{\Lambda}$ and $\beta^s = \Lambda^s - \frac{2}{5}\overline{\Lambda}$. Thus, country 2 wants to deviate to e, and this is a contradiction. □

E Proof of Proposition 4, 5

Before the proof, we introduce the D1 criterion developed by Cho and Kreps (1987). Fix a PBE, and suppose $a' \in \{e, l\}$ is off-the-equilibrium action. Let $v^*(s)$ be type s's utility in the PBE and $v'(s)$ be type s's utility from deviating to a' under the belief $\hat{\Lambda}^{a'}$. Moreover, let $D_s \equiv \{\hat{\Lambda}^{a'} | v'(s) > v^*(s)\}$ and $D_s^0 \equiv \{\hat{\Lambda}^{a'} | v'(s) = v^*(s)\}$. Then, *a belief $\hat{\Lambda}^{a'}$ imposed D1* is restricted to $\hat{\Lambda}^{a'} = \Lambda^{-s}$ if there exist s and $-s$ such that $D_s \cup D_s^0 \subset D_{-s}$.

Proof of proposition 4: We show that no type has an incentive to deviate from the profile (l, l, e) when $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$. First, we consider country 2. Given both types' action strategies, it holds that $\mathbb{E}v_2(l, l, e|s) = \sum_s \overline{\rho}_s \left(-5(\beta^s)^2 + 6\Lambda^s\beta^s\right) = \sum_s \overline{\rho}_s(\Lambda^s)^2 + \frac{4}{5}(\overline{\Lambda})^2 > \sum_s \overline{\rho}_s(\Lambda^s)^2 + \frac{11}{16}(\overline{\Lambda})^2 = \sum_s \overline{\rho}_s \left(-5(\alpha^s)^2 + 6\Lambda^s\alpha^s\right) = \mathbb{E}v_2(l, l, l|s)$, where $\alpha^s = \Lambda^s - \frac{1}{4}\overline{\Lambda}$ and $\beta^s = \Lambda^s - \frac{2}{5}\overline{\Lambda}$. Next, for type H, it holds that $v_1(l, l, e|H) = 3(\Lambda^H - \frac{2}{5}\overline{\Lambda})^2 \geq 3(\Lambda^H - \frac{1}{4}\hat{\Lambda}^e)^2 = v_1(e, l, e|H)$ when $\frac{8}{5}\overline{\Lambda} \leq \hat{\Lambda}^e$. For type L, $v_1(l, l, e|L) = 3(\Lambda^L - \frac{2}{5}\overline{\Lambda})^2 \geq 3(\Lambda^L - \frac{1}{4}\hat{\Lambda}^e)^2 = v_1(l, e, e|L)$ when $\frac{8}{5}\overline{\Lambda} \geq \hat{\Lambda}^e$. Therefore, if

[68] $v_1(e, e, l|H) = \frac{9}{5}(\Lambda^H)^2$ holds, since it holds $v_1(e, e, l|H) = w_1(\Lambda^H, \tilde{\Lambda}^H, 0) = w_1(\Lambda^H, \Lambda^H, \frac{2}{5}\Lambda^H) = \frac{9}{5}(\Lambda^H)^2$ in the H-hybrid equilibrium. For the detail, see the proof of Lemma 1.

[69] It holds $G(\hat{\Lambda}^l) = \frac{3}{2}\Lambda^H(\hat{\Lambda}^l - \frac{4}{5}\Lambda^H) - \frac{3}{16}(\hat{\Lambda}^l)^2$. Thus, to show $G((4-2\sqrt{3})\Lambda^L + \frac{2\sqrt{3}}{3}\tilde{\Lambda}^H) < 0$, it suffices to show $\hat{\Lambda}^l - \frac{4}{5}\Lambda^H < 0$ when $\hat{\Lambda}^l = (4-2\sqrt{3})\Lambda^L + \frac{2\sqrt{3}}{3}\tilde{\Lambda}^H$. By $\tilde{\Lambda}^H = 3(1 - \frac{2\sqrt{5}}{5})\Lambda^H$, we can see $\hat{\Lambda}^l - \frac{4}{5}\Lambda^H = (4-2\sqrt{3})\Lambda^L + \frac{10\sqrt{3} - 4\sqrt{15} - 4}{5}\Lambda^H < 0$.

off-the-equilibrium belief is $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda}$, both types have no incentive to deviate. Note that it holds $\hat{\Lambda}^e = \frac{8}{5}\overline{\Lambda} \in [\Lambda^L, \Lambda^H]$ when $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$.

Finally, we show that the equilibrium survives D1 criterion. Since $D_H = \{\hat{\Lambda}^e | \frac{8}{5}\overline{\Lambda} > \hat{\Lambda}^e\}$, $D_H^0 = \{\frac{8}{5}\overline{\Lambda}\}$, $D_L = \{\hat{\Lambda}^e | \hat{\Lambda}^e > \frac{8}{5}\overline{\Lambda}\}$ and $D_L^0 = \{\frac{8}{5}\overline{\Lambda}\}$, $D_s \cup D_s^0 \nsubseteq D_{-s}$ holds for any s. Thus, $D1$ does not put any restrictions on off-the-equilibrium belief $\hat{\Lambda}^e$. □

Proof of proposition 5: First, we consider the case of $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$ and show that no type has an incentive to deviate from H-hybrid equilibrium in which $(p_1^H(e), p_1^L(e), p_2(e)) = (p_H^*, 0, 1)$, where $p_H^* = 1 - \frac{\rho'}{1-\rho'} \cdot \frac{1-\overline{\rho}}{\overline{\rho}}$ with $\rho' = \frac{\frac{5}{8}\Lambda^H - \Lambda^L}{\Lambda^H - \Lambda^L}$. By Bayes' rule and equilibrium strategies, we can see $\hat{\Lambda}^e = \Lambda^H$ and $\hat{\Lambda}^l = \frac{(1-p_H^*)\overline{\rho}}{(1-p_H^*)\overline{\rho}+(1-\overline{\rho})}\Lambda^H + \frac{(1-\overline{\rho})}{(1-p_H^*)\overline{\rho}+(1-\overline{\rho})}\Lambda^L = \frac{5}{8}\Lambda^H$.[70]

For country 2, given both types' equilibrium strategies, it holds that
$\mathbb{E}v_2(a_1^H, l, e|s) = \overline{\rho}p_H^* \left(-5(\alpha^H)^2 + 6\Lambda^H\alpha^H\right) + \overline{\rho}(1-p_H^*)\left(-5(\beta^H)^2 + 6\Lambda^H\beta^H\right) + (1-\overline{\rho})\left(-5(\beta^L)^2 + 6\Lambda^L\beta^L\right)$, where $\alpha^H = \Lambda^H - \frac{1}{4}\Lambda^H$, $\beta^H = \Lambda^H - \frac{2}{5}\hat{\Lambda}^l$ and $\beta^L = \Lambda^L - \frac{2}{5}\hat{\Lambda}^l$. If country 2 deviates to l, then $\mathbb{E}v_2(a_1^H, l, l|s) = \overline{\rho}p_H^* \left(-3(\gamma^H)^2 + 6(\Lambda^H - \frac{2}{5}\Lambda^H)\gamma^H\right) + \overline{\rho}(1-p_H^*)\left(-5(\alpha^H)^2 + 6\Lambda^H\alpha^H\right) + (1 - \overline{\rho})\left(-5(\alpha^L)^2 + 6\Lambda^L\alpha^L\right)$, where $\gamma^H = \frac{3}{2}\Lambda^H - \frac{2}{5}\Lambda^H - \frac{1}{2}\Lambda^H$, $\alpha^H = \Lambda^H - \frac{1}{4}\hat{\Lambda}^l$ and $\alpha^L = \Lambda^L - \frac{1}{4}\hat{\Lambda}^l$. Then, we have $\mathbb{E}v_2(a_1^H, l, e|s) - \mathbb{E}v_2(a_1^H, l, l|s) = \overline{\rho}p_H^*\left[\frac{27}{16}(\Lambda^H)^2 - \frac{27}{25}(\Lambda^H)^2\right] + (\overline{\rho}(1-p_H^*) + (1 - \overline{\rho}))\left[\frac{3}{5}(\hat{\Lambda}^l)^2 - \frac{39}{80}(\hat{\Lambda}^l)^2\right] > 0$. Thus, we can see country 2 never deviates. For type H, in the mixed action strategy, $v_1(e, l, e|H) = v_1(l, l, e|H)$ holds, that is, $3(\Lambda^H - \frac{1}{4}\Lambda^H)^2 = 3(\Lambda^H - \frac{2}{5}\hat{\Lambda}^l)^2$.[71] Given beliefs $\hat{\Lambda}^e$ and $\hat{\Lambda}^l$, type H has no incentive to deviate from the mixed action strategy p_H^*. For type L, it holds $v_1(p_H^*, e, e|L) = 3(\Lambda^L - \frac{1}{4}\Lambda^H)^2 = 3(\Lambda^L - \frac{2}{5}\hat{\Lambda}^l)^2 = v_1(p_H^*, l, e|L)$ since $\frac{1}{4}\Lambda^H = \frac{2}{5}\hat{\Lambda}^l$.[72] Thus, type L also has no incentive to deviate. We do not have to check whether this equilibrium survives $D1$, since there is no off-the-equilibrium path.

Finally, let us check $p_H^* \in (0, 1)$ when $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$. Now, define $\rho(x) = \frac{x\Lambda^H - \Lambda^L}{\Lambda^H - \Lambda^L}$ with $x \in [0, 1]$. Then, we have $\rho(x) \in (0, 1]$ and $\rho(x)\Lambda^H + (1 - \rho(x))\Lambda^L = x\Lambda^H$. By $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$ and the definition of ρ', it holds $\overline{\rho} > \rho(\frac{5}{8}) = \rho'$. Thus, we have $1 > \frac{\rho'}{1-\rho'} \cdot \frac{1-\overline{\rho}}{\overline{\rho}} > 0$, and $p_H^* \in (0, 1)$.

The proof of the case $\overline{\Lambda} \in (\Lambda^L, \frac{5}{8}\Lambda^L)$ is similar to the above. □

F Proof of proposition 6, 7 and 8

To prove Proposition 6 and 8, we use the following lemma.

[70]By $p_H^* = 1 - \frac{\rho'}{1-\rho'} \cdot \frac{1-\overline{\rho}}{\overline{\rho}}$ with $\rho' = \frac{\frac{5}{8}\Lambda^H - \Lambda^L}{\Lambda^H - \Lambda^L}$, we can see $\hat{\Lambda}^l = \rho'\Lambda^H + (1-\rho')\Lambda^L = \frac{5}{8}\Lambda^H$.

[71]p_H^* is constructed so that $v_1(e, l, e|H) = v_1(l, l, e|H)$ holds, that is $\frac{1}{4}\Lambda^H = \frac{2}{5}\hat{\Lambda}^l$.

[72]While v_1 is a function of (a_1^H, a_1^L, a_2), we use this notation to express a utility under mixed action strategies for convenience. That is, we use the notation $v_1(p_1^H(e), p_1^L(e), p_2(e))$.

Lemma 7. *Let $T \equiv t_1^s - t_2$ and $U \equiv u_1^s + u_2$. Then, U is represented as a function of T and satisfies the following: $U'(T) > 0$ when $T < 0$, $U'(T) < 0$ when $T > 0$ and $U'(T) = 0$ when $T = 0$.*

Proof. Since $r(\overline{k} - k_1^s) + r(\overline{k} - k_2^s) = 0$ holds by (4), we have $U = f_1(k_1^s) + f_2(k_2^s)$ by (5). By (6) and (7), we can see that U is a function of T, and $U'(T) = \frac{\partial k_1^s}{\partial T} f_1'(k_1^s) + \frac{\partial k_2^s}{\partial T} f_2'(k_2^s) = -\frac{1}{4}(A_1^s - 2k_1^s) + \frac{1}{4}(A_2 - 2k_2^s) = -\frac{1}{4}T$. This implies $U'(T) > 0$ when $T < 0$, $U'(T) < 0$ when $T > 0$ and $U'(T) = 0$ when $T = 0$. □

$U(T)$ is social welfare as the sum of both countries' utilities. Lemma 7 means that social welfare increases when the gap $t_1^s - t_2$ is reduced, and it is maximized when $t_1^s = t_2$ holds.

Proof of proposition 6: We use Lemma 7. First of all, the gap of tax rates $t_1^s - t_2$ under complete information is $\frac{1}{2}\Lambda^s$ by Proposition 1 and (13).

[Proof of (*ii*)]: When $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$, by Proposition 4 and (16), the gap of tax rates $t_1^s - t_2$ is $\frac{1}{3}\left(\Lambda^s - \frac{2}{5}\overline{\Lambda}\right) - \left(-\frac{2}{5}\overline{\Lambda}\right) = \frac{1}{3}\Lambda^s + \frac{4}{15}\overline{\Lambda}$. Under the assumption of $-\frac{1}{5}\Lambda^L < \Lambda^H < -5\Lambda^L$ and $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$, we have $|\frac{1}{3}\Lambda^s + \frac{4}{15}\overline{\Lambda}| < |\frac{1}{2}\Lambda^s|$, which implies $W^{asy}(s) > W^{com}(s)$ for any $s \in \{H, L\}$ by Lemma 7.

[Proof of (*i*)]: First, we consider the case of $s = L$. When $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$, by Proposition 5 and (16), the gap of tax rates $t_1^L - t_2$ is $\frac{1}{3}\Lambda^L + \frac{4}{15}\hat{\Lambda}^l$, where $\hat{\Lambda}^l = \frac{5}{8}\Lambda^H$ in the hybrid equilibrium.[73] Under the assumption of $-\frac{1}{5}\Lambda^L < \Lambda^H < -5\Lambda^L$, we have $|\frac{1}{3}\Lambda^L + \frac{4}{15}\hat{\Lambda}^l| < |\frac{1}{2}\Lambda^L|$, which implies $W^{asy}(L) > W^{com}(L)$ by Lemma 7. Next, we consider the case of $s = H$. When $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$, by Proposition 5, type H resorts to a mixed action strategy between e and l. In the situation G_N, since the type is fully revealed, the same outcome is realized as in complete information. In the situation G_2, by (16) and $\hat{\Lambda}^l = \frac{5}{8}\Lambda^H$, the gap of tax rates $t_1^L - t_2$ is $\frac{1}{3}\Lambda^H + \frac{4}{15}\hat{\Lambda}^l = \frac{1}{2}\Lambda^H$, which is the same as the gap under complete information. These implies that $W^{asy}(H) = W^{com}(H)$ holds.

[Proof of (*iii*)]: The proof is similar to the proof of (*i*). □

Proof of proposition 7: By Proposition 1 and Lemma 2, $u_2^{comp}(s) = v_2(e, e, e|s) = -5(\alpha^s)^2 + 6\Lambda^s\alpha^s = \frac{27}{16}(\Lambda^s)^2$ with $\alpha^s = \Lambda^s - \frac{1}{4}\Lambda^s$, regardless of $\overline{\Lambda}$.

When $\overline{\Lambda} \in [\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H]$, by Proposition 4 and Lemma 2, $u_2^{asy}(s) = -5(\beta^s)^2 + 6\Lambda^s\beta^s = (\Lambda^s)^2 + \frac{8}{5}\overline{\Lambda}\Lambda^s - \frac{4}{5}(\overline{\Lambda})^2$ with $\beta^s = \Lambda^s - \frac{2}{5}\overline{\Lambda}$. Then, for any s and $\overline{\Lambda} \in (\frac{5}{8}\Lambda^L, \frac{5}{8}\Lambda^H)$, we have $u_2^{comp}(s) - u_2^{asy}(s) = \frac{11}{16}(\Lambda^s)^2 - \frac{8}{5}\overline{\Lambda}\Lambda^s + \frac{4}{5}(\overline{\Lambda})^2 > \frac{11}{16}(\Lambda^s)^2 - \frac{8}{5}\Lambda^s \cdot \frac{5}{8}\Lambda^s + \frac{4}{5}(\frac{5}{8}\Lambda^s)^2 = 0$. Moreover, $u_2^{comp}(H) - u_2^{asy}(H) = 0$ when $\overline{\Lambda} = \frac{5}{8}\Lambda^H$, and $u_2^{comp}(L) - u_2^{asy}(L) = 0$ when $\overline{\Lambda} = \frac{5}{8}\Lambda^L$.

[73] See the proof of Proposition 5 for the detail of $\hat{\Lambda}^l$.

When $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$ and $s = H$, type H resorts to a mixed action strategy. Since $\hat{\Lambda}^l = \frac{5}{8}\Lambda^H$ and $\hat{\Lambda}^e = \Lambda^H$ hold in the hybrid equilibrium, by (13) and (16), the equilibrium outcome is the same as in complete information regardless of type H's choice of timing.[74] Thus, $u_2^{comp}(H) = u_2^{asy}(H)$ holds. When $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$ and $s = L$, the situation G_2 is realized. Then, by Lemma 2, $u_2^{asy}(L) = -5(\beta^L)^2 + 6\Lambda^L\beta^L = (\Lambda^L)^2 + \Lambda^H\Lambda^L - \frac{5}{16}(\Lambda^H)^2$ with $\beta^L = \Lambda^L - \frac{2}{5}\hat{\Lambda}^l$, and hence we have $u_2^{comp}(L) - u_2^{asy}(L) = \frac{1}{16}(11\Lambda^L - 5\Lambda^H)(\Lambda^L - \Lambda^H) > 0$.

When $\overline{\Lambda} \in (\Lambda^L, \frac{5}{8}\Lambda^L)$, $\hat{\Lambda}^l = \frac{5}{8}\Lambda^L$ holds in the hybrid equilibrium. Using this condition, we can show $u_2^{comp}(L) = u_2^{asy}(L)$ holds regardless of type L's timing choice, and $u_2^{comp}(H) - u_2^{asy}(H) = \frac{1}{16}(11\Lambda^H - 5\Lambda^L)(\Lambda^H - \Lambda^L) > 0$, as in the proof of the case of $\overline{\Lambda} \in (\frac{5}{8}\Lambda^H, \Lambda^H)$ above. □

Proof of proposition 8: (i: $W_{G_N}^{com}$ vs $W_{G_N}^{asy}$) In the situation G_N, by (13), the gaps of tax rates $t_1^s - t_2$ under complete and asymmetric information are $\frac{1}{2}\Lambda^s$ and $\frac{1}{3}\Lambda^s + \frac{1}{6}\overline{\Lambda}$, respectively. Under the assumption of $-\frac{1}{5}\Lambda^L < \Lambda^H < -5\Lambda^L$, we have $|\frac{1}{3}\Lambda^s + \frac{1}{6}\overline{\Lambda}| < |\frac{1}{2}\Lambda^s|$, which implies $W_{G_N}^{asy}(s) > W_{G_N}^{com}(s)$ for any $s \in \{H, L\}$ by Lemma 7.

(ii: $W_{G_2}^{com}$ vs $W_{G_2}^{asy}$) First, we compare $W_{G_N}^{com}(s)$ and $W_{G_2}^{com}(s)$. By (13) and (16), the gaps of tax rates $t_1^s - t_2$ in the situations G_N and G_2 under complete information are $\frac{1}{2}\Lambda^s$ and $\frac{3}{5}\Lambda^s$, respectively. Then, we have $|\frac{1}{2}\Lambda^s| < |\frac{3}{5}\Lambda^s|$, which implies $W_{G_N}^{com}(s) > W_{G_2}^{com}(s)$ holds by Lemma 7. Since $W_{G_2}^{asy}(s) \geq W_{G_N}^{com}(s)$ by Proposition 6, we can see $W_{G_2}^{asy}(s) \geq W_{G_N}^{com}(s) > W_{G_2}^{com}(s)$ for any s.

(iii: $W_{G_1}^{com}$ vs $W_{G_1}^{asy}$) First, by (22), the gap of tax rates $t_1^s - t_2$ in the situation G_1 under complete information is $\frac{3}{5}\Lambda^s$. When $\overline{\Lambda} \in [\tilde{\Lambda}^L, \tilde{\Lambda}^H]$, by Lemma 1 and (19), the gap of tax rates $t_1^s - t_2$ under asymmetric information is $\frac{1}{3}\overline{\Lambda}$. Then, we have $|\frac{1}{3}\overline{\Lambda}| < |\frac{3}{5}\Lambda^s|$, which implies $W_{G_1}^{asy}(s) > W_{G_1}^{com}(s)$. Similarly, when $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ and $t_1^L = 0$, or when $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$ and $t_1^H = 0$, we can see $W_{G_1}^{asy}(s) > W_{G_1}^{com}(s)$ for the same reason. When $\overline{\Lambda} \in (\Lambda^L, \tilde{\Lambda}^L)$ and $t_1^L = \frac{2}{5}\Lambda^L$, or when $\overline{\Lambda} \in (\tilde{\Lambda}^H, \Lambda^H)$ and $t_1^H = \frac{2}{5}\Lambda^H$, since the type is fully revealed in the hybrid equilibrium and each type's tax rate is the same as in complete information, country 2's reaction is also the same as in complete information. Therefore, $W_{G_1}^{asy}(s) = W_{G_1}^{com}(s)$ in those cases. □

[74] When type H chooses *early*, the situation G_N is realized. Then, by (13) and $\hat{\Lambda}^e = \Lambda^H$, $t_1^H = \frac{1}{3}\left(\Lambda^H - \frac{1}{4}\Lambda^H\right) = \frac{1}{4}\Lambda^H$ and $t_2 = -\frac{1}{4}\Lambda^H$. When type H chooses *late*, the situation G_2 is realized. Then, by (16) and $\hat{\Lambda}^l = \frac{5}{8}\Lambda^H$, $t_1^H = \frac{1}{3}\left(\Lambda^H - \frac{2}{5} \cdot \frac{5}{8}\Lambda^H\right) = \frac{1}{4}\Lambda^H$ and $t_2 = -\frac{2}{5} \cdot \frac{5}{8}\Lambda^H = -\frac{1}{4}\Lambda^H$.