// NO.26-057 | 11/2026

DISCUSSION PAPER

// NIKOLAJ MORETTI

Minimum Tax Rates in Markov Perfect Equilibrium

Leibniz Association

ZEW

Minimum Tax Rates in Markov Perfect Equilibrium

Nikolaj Moretti*†

November 2026

Abstract

Static models of tax competition suggest that a minimum tax can make even low-tax jurisdictions better off, so that the global minimum tax of 15% on multinational firms' profits would be in tax havens' own interest. This article asks whether this holds when governments are forward-looking. It analyzes Markov perfect equilibria of a dynamic game of tax competition between a high-tax and a low-tax player with asynchronous, stochastic moves, in which tax rates are strategic complements. In this setting, each player sets its rate above its static best response to pull the opponent's future rate up. A binding minimum tax can make the low-tax player unresponsive, thus removing the high-tax player's wedge. If the minimum tax is below a threshold, the high-tax player lowers its rate despite strategic complementarity. While the high-tax player gains in steady-state returns, the low-tax player loses for minimum taxes just above the rate it sets without one. Calibrated to business tax competition between German municipalities around the introduction of a minimum tax in 2004, the model produces a cut in the high-tax player's rate close to a recent causal estimate, the low-tax player loses for every minimum tax that binds it alone, and aggregate tax revenue increases unless the minimum tax lies just above the low-tax player's rate without one. If these results carry over to the international setting, havens are less likely to gain from a minimum tax than static models suggest, which makes their participation in the global minimum tax harder to secure.

Keywords: tax competition, minimum taxation, global minimum tax, dynamic game, strategic complements, Markov perfect equilibrium
JEL codes: C73, F68, H25, H71, R30

*ZEW – Leibniz Centre for European Economic Research, L 7, 1, 68161 Mannheim, Germany, and Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Telegrafenberg A31, 14473 Potsdam, Germany. E-mail: nikolaj.moretti@zew.de

†I am deeply grateful to Lennart Stern for many discussions that shaped this article. I also thank Beatriz Gaitan, Eckhard Janeba, Matthias Kalkuhl, and the participants of the 2026 Public Finance Conference (ZEW, Mannheim), of the 27th Global Conference on Environmental Taxation (Berlin), and of the *Föderales Forum 2026* (German Federal Ministry of Finance) for helpful comments. I gratefully acknowledge a doctoral scholarship from the German Federal Environmental Foundation (DBU) and funding from the German Federal Ministry of Research, Technology and Space (BMFTR) as part of the Kopernikus project Ariadne (grant reference 03SFK5J0).

1 Introduction

In October 2021, 136 countries agreed on a global minimum tax of 15% on multinational firms' profits.[1] The agreement is aimed at reducing race-to-the-bottom dynamics in international tax competition and the profit shifting that comes with them, which causes sizable losses in tax revenues for many governments.[2] Observers consider the agreement "unique when it comes to international tax coordination" (Janeba and Schjelderup 2023, p. 1) and "a pathbreaking innovation in global tax policies" (Hebous and Keen 2023, p. 1). An agreement without a global enforcing authority, however, lasts only if countries are better off inside it than outside, that is, if their participation constraints are met. That high-tax countries gain from a minimum tax is straightforward, because it raises the tax rates of havens and thereby narrows the tax differentials that drive profit shifting. For a haven, the condition is more demanding; the minimum tax constrains its own tax rate, so it can gain only through the other countries' reactions. Absent sanctions against non-members, a haven that stays outside is at least as well off as without the agreement, so a haven that loses from the minimum tax would not join.[3] In the European Union, where there are calls for a more ambitious minimum tax of 21% or 25% (Greens/EFA Group in the European Parliament 2021; European Trade Union Confederation 2025), the same condition follows from the unanimity requirement for tax legislation: every member state must be at least as well off with such a minimum tax as without it.[4] The case that havens might gain from such a floor, too, rests on a simple argument. If tax rates are strategic complements, a minimum tax that raises the haven's rate makes the high-tax country raise its rate as well, and the haven benefits from this response. Hebous and Keen (2023) formalize the argument in a static, simultaneous-move game, in which a floor marginally above the haven's Nash rate is a Pareto improvement, so that havens gain as well. In an illustrative calibration, the minimum tax most preferred by the haven is 16.7%, exceeding the agreed 15%.

The cleanest available test of this argument points the other way. Empirical studies of the global minimum tax are to be expected only a few years from now, but Buettner and Poehnlein

[1] The agreement was reached within the OECD and G20's Inclusive Framework on Base Erosion and Profit Shifting. It includes extra-territorial top-up taxation of profits in non-member states to reach this threshold.

[2] Using country-by-country reporting data, Garcia-Bernardo and Janský (2024, p. 2) put the profits that multinational corporations moved to tax havens in 2017 at more than 850 billion USD, and the tax revenue that other countries lost as a result at 200 to 300 billion USD. Chiari (2024) arrives at higher figures: a loss in global tax revenue of approximately 480 billion USD in 2019 based on statutory corporate tax rates, and of 600 billion USD based on effective tax rates.

[3] In the presence of sanctions, like the extra-territorial top-up taxation that is part of the global minimum tax agreement, a haven that loses from the minimum tax may be better off inside the agreement than outside. However, it can join the agreement in form without joining in substance. When the G7 proposed the minimum tax in June 2021, four months before the agreement, Switzerland prepared to sign up to it while drawing up plans for subsidies and other incentives to offset the higher tax rates (Jones 2021). Switzerland subsequently joined the agreement; to date, several Swiss cantons have implemented such offsetting measures or are in the process of doing so (KPMG 2026). Sanctions can thus secure a haven's formal participation but not its interest in the minimum tax. The question of whether a haven gains absent sanctions, as analyzed in this article, thus remains relevant for whether the agreement holds in substance.

[4] That Ireland, Hungary, and Estonia did not sign up to the agreement on a 15% minimum tax in July 2021 (Greens/EFA Group in the European Parliament 2021) suggests that unanimity for a higher rate cannot be taken for granted. To date, no minimum tax above 15% has been enacted in the EU.

(2024) analyze the introduction of a minimum business tax rate for German municipalities in 2004, which they consider "[a] more promising case" for empirical evaluation than the international experience with minimum taxes (Buettner and Poehnlein 2024, p. 1). The authors find that high-tax municipalities in geographical proximity of a tax haven *lower* their tax rates after the introduction of the floor, compared to other high-tax municipalities. As the authors point out, this response is hard to reconcile with the common view that the tax rates of competing jurisdictions are strategic complements. They further document a contemporaneous decline in havens' tax revenue, which suggests that havens lose from the minimum tax. Hebous and Keen (2023, p. 3, fn. 19) interpret the drop in high-tax municipalities' tax rates, reported in an earlier version of the study, as evidence of strategic substitutability at the subnational level. Yet estimated reaction functions point toward strategic complementarity, in the international sphere (Leibrecht and Hochgatterer 2012; Heimberger 2021) as well as in business tax competition between German municipalities (Thunecke et al. 2026).

This article presents a model in which tax rates are strategic complements and the high-tax player can nevertheless lower its tax rate in response to a minimum tax. The model is an infinite-horizon sequential two-player game that combines the model of dynamic duopolistic competition from Maskin and Tirole (1987) with the model of cross-border shopping from Kanbur and Keen (1993). The latter has been reinterpreted to represent tax competition over profits shifted between jurisdictions of revenue-maximizing governments by Keen and Konrad (2013) and Agrawal and Wildasin (2020). The two players are jurisdictions, which can be read as municipalities or as countries. They compete for a mobile tax base through their tax rates, maximizing intertemporal tax revenues and non-tax benefits. Moves are stochastic: in each period, one of the two players or none has the opportunity to adjust their tax rates. In this game, I characterize steady states of Markov perfect equilibria (MPE) in general, study the affine MPE in depth, and analyze the consequences of introducing a minimum tax. The mechanism that reconciles the observation by Buettner and Poehnlein (2024) with strategic complementarity also works against the optimistic conclusion by Hebous and Keen (2023) that havens gain from a minimum tax.

At the heart of the model is a *dynamic incentive*: a player that sets its tax rate knows that its opponent will respond to this rate once it has the opportunity to move. Because tax rates are strategic complements, the higher the player's own rate, the higher the opponent's response, so the player sets its rate above its static best response to pull the opponent's future rate up. I show that, under general regularity conditions, a steady-state tax rate induced by an MPE (if one exists) equals the player's static best response plus an *incentive wedge*, which is larger the sooner and the more strongly the opponent responds. When restricting the admissible Markovian strategies to the class of affine-linear functions, there is a unique MPE with slopes between zero and one, its tax profile converges to a steady state almost surely, and strategies and steady-state tax rates are available in closed form. Both steady-state tax rates feature a positive *dynamic premium* over their static Nash equilibrium counterparts. The dynamic incentive is akin to the leader's *commitment* in Stackelberg games.

A minimum tax at or above the low-tax player's tax rate can switch this incentive off for the high-tax player. Intuitively, the floor now keeps the low-tax player's rate from falling, as the high-tax player's own high rate did before. Because the high-tax player has no future rate

to pull up, its incentive wedge vanishes and it chooses its static best response to the minimum tax. I construct a *local* MPE with this steady state, that is, an MPE of the game in which both players' tax rates are bounded from above, in which the low-tax player stays at the minimum tax. The bounds can be chosen wide enough to include the tax profile at which the game stands when the minimum tax is introduced, and the steady state is reached from there almost surely; the existence of an MPE without these bounds is not shown. The high-tax player's rate change then consists of two parts: a positive *complementarity effect*, the reaction to the minimum tax's mechanical increase of the low-tax player's rate, and a negative *wedge effect*, the loss of the incentive wedge. I derive a threshold for the minimum tax below which the wedge effect dominates, so that the high-tax player lowers its tax rate *despite* strategic complementarity. In terms of steady-state returns, the high-tax player gains from every minimum tax that binds only the low-tax player, while the low-tax player loses from minimum taxes just above its rate in the affine steady state, because the high-tax player's lower rate takes tax base away from it. Aggregate tax revenue rises strictly with the minimum tax. Depending on the parameters, it can nevertheless fall when a minimum tax is introduced, namely for minimum taxes below a second threshold. This threshold lies below the first one; the high-tax player lowering its rate is thus a necessary condition for aggregate tax revenue to fall.

The model is then calibrated to the context of business tax competition between German municipalities around 2004, the year in which a minimum tax rate was enacted, using data for the German federal state of Mecklenburg-Vorpommern, a subset of the data used for the empirical analysis by Buettner and Poehnlein (2024). At the calibration, the minimum tax lies below the first threshold: the high-tax player changes its tax rate by -0.38 percentage points (pp), against a difference-in-differences estimate of -0.51 pp in Buettner and Poehnlein (2024). This net change masks two larger effects, a complementarity effect of $+1.31$ pp and a wedge effect of -1.69 pp. Moreover, the haven's loss obtains over the whole range of admissible minimum taxes. At the minimum tax introduced in 2004, aggregate tax revenue rises by 24.3%; it falls only for minimum taxes less than 0.17 pp above the haven's rate without one, that is, below the second threshold. The model does not account for the havens' observed tax rate increases beyond the minimum tax; Section 6 discusses possible explanations.

Several policy implications follow from the model. At the national level, a central government that cares about aggregate tax revenue should only introduce minimum taxes that are sufficiently high to prevent revenue losses. If transferable to the international setting, the results have two implications for the global minimum tax. First, havens are less likely to gain from a minimum tax than static models suggest, which makes their participation constraints harder to meet. At the calibration, the haven loses even from minimum taxes to which the high-tax player responds with a higher tax rate. Second, while this message is pessimistic, understanding the underlying mechanism is crucial in designing the remedy. An instrument that bounds tax rates may remove the incentive wedge, while an instrument that shifts them does not. In Section 6, I discuss an instrument of the latter kind, the *single market levy* for the EU proposed by García Antón and Lejour (2023): a common rate levied on top of the member states' domestic tax rates to fund the EU budget, which raises the tax rates firms face and leaves the dynamic incentive of both players intact.

The model contributes to two strands of the literature on tax competition. First, it con-

tributes to the empirical literature on the question of whether games of tax competition are games of strategic complements or substitutes. Against the evidence in favor of strategic complementarity (Leibrecht and Hochgatterer 2012; Heimberger 2021), Parchet (2019) uses data for tax setting behavior by Swiss municipalities and documents quasi-experimental evidence based on an instrumental variable in favor of strategic substitutability at the subnational level. As noted above, the drop in high-tax municipalities' tax rates after the introduction of the German minimum tax has been read as pointing in the same direction by Hebous and Keen (2023). However, Thunecke et al. (2026) provide evidence for strategic complementarity in tax competition between German municipalities. The mechanism described in my model explains why using the introduction of minimum taxes as an exogenous source of variation might lead to wrong conclusions: in the model, the high-tax player can lower its tax rate *despite* strategic complementarity and upward sloping equilibrium strategies, because a minimum tax changes not only the competitor's tax rate but also how the competitor responds. The quantitative results further suggest that even if high-tax municipalities are observed to increase their tax rates in reaction to the minimum tax, estimates of their reaction functions based on this observation may be biased because the observations are the net of two counteracting effects.

Second, it adds to the theoretical literature that analyzes the effect of minimum tax rates and the global minimum tax specifically. Many theoretical contributions use models of static, simultaneous-move games (Janeba and Peters 1999; Johannesen 2022; Hebous and Keen 2023; Janeba and Schjelderup 2023). Kiss (2012) is one of two dynamic exceptions and analyzes trigger strategies in an infinitely repeated game. Paulus et al. (2026), the second exception, study the global minimum tax in a differential game between non-haven countries in which mobile capital reallocates gradually because of tax differentials, and compare Markovian and open-loop equilibria under two enforcement rules of the minimum tax. In their model, Markovian strategies imply real-time adjustment of tax rates by governments, whereas here a player's tax rate stays in place until its next opportunity to move. To the best of my knowledge, this article is the first to analyze a minimum tax in MPE of an infinite-horizon game between a haven and a non-haven player.

This model's results are similar to the results obtained in Stackelberg models of tax competition (Wang 1999; Konrad 2009), in that the high-tax player *decreases* its tax rate after the introduction of a minimum tax, despite strategic complementarity, and the low-tax player may be worse off. Buettner and Poehnlein (2024, p. 2) note that their finding is in accordance with such a leader–follower structure. However, the roles of leader and follower may be difficult or implausible to attribute in some real-world settings. For instance, Hebous and Keen (2023, p. 3) note that the low-tax player loses from a minimum tax just above its uncoordinated rate if one player moves first, but argue that "the Nash case seems at least as plausible and important to explore". A contribution of this article is therefore to offer a model framework in which those effects arise without the need to assign the roles of leader and follower: with asynchronous moves and an infinite horizon, each player is leader and follower in turn.

The rest of the article is structured as follows. Section 2 introduces the static setting, derives the static Nash benchmark, embeds it in a dynamic model, and introduces the MPE as the solution concept. Section 3 studies steady-state tax profiles for MPE — where they exist — under general regularity conditions, before narrowing down the analysis to the class

of affine-linear strategies. Section 4 introduces a binding minimum tax and derives its effect on the high-tax player's tax rate, on both players' steady-state returns, and on aggregate tax revenue. Section 5 contains the calibration as well as the results based on it. Section 6 discusses the limitations of the model and of the calibration, the lessons for the global minimum tax, and a tax club as an alternative, before Section 7 concludes.

2 Model

In this section, I introduce the model. It is a dynamic game in infinite discrete time with two players, moving asynchronously, where the timing and frequency of moves are stochastic. It features infinite state and action spaces as well as deterministic transitions, that is, deterministic mappings from the current action to the future state. Each player chooses its tax rate to maximize the infinite discounted sum of returns from tax revenues and non-tax benefits, given the opponent's strategy. I will start by introducing the static setting from which I obtain the period returns for the players in the model and analyze the resulting simultaneous-move game. Next, I generalize this static game to a dynamic setting with an infinite horizon.

2.1 Instantaneous Returns

The structure of the model in every period t relies on the model of tax competition with the possibility of cross-border shopping by consumers in Kanbur and Keen (1993). It is reformulated as a model of competition in profit taxes between two jurisdictions. The jurisdictions can be read as municipalities, as in the business tax competition between German municipalities around the introduction of a minimum business tax rate in 2004 (Section 5), or as countries: Keen and Konrad (2013) explain how the results from the model by Kanbur and Keen (1993) can be translated directly to the context of international corporate taxation where multinational firms can shift profits between countries. Consider a model where firms do business on the territory of two jurisdictions, henceforth Players H and L. The profit of a firm is taxed depending on where it declares its legal seat. Assume that there is a continuum of firms with measure 2 that consider whether to locate in jurisdiction H or L. For each firm, gross profits $\bar{\pi}$ are independent of its location. However, net profits of firm $k \in [0, 2]$ for location decision H and L are given by

$$\pi_{\mathrm{H}}(k; \tau_{\mathrm{H}}) = (1 - \tau_{\mathrm{H}})\bar{\pi}(k) - \frac{\theta}{2}\, k, \qquad \pi_{\mathrm{L}}(k; \tau_{\mathrm{L}}) = (1 - \tau_{\mathrm{L}})\bar{\pi}(k) - \frac{\theta}{2}(2 - k). \tag{1}$$

Net profits thus depend on the tax rate of the jurisdiction and on some additional, jurisdiction-specific cost which differs for every firm depending on their location on the interval $[0, 2]$. The additional cost represents each firm's preference for one of the two jurisdictions in the absence of a tax differential. When firms choose where to operate, it could capture differences in housing and rental prices, the availability of (skilled) workers, amenities that make it easier to attract workers, etc. When firms choose where to declare their profits, it captures the cost of declaring them in a jurisdiction other than the one in which the firm operates, which differs across firms.[5]

[5]Keen and Konrad (2013) obtain the same tax base from a different assumption on firms: all firms are identical and each shifts a part of its profit to the other jurisdiction at a quadratic cost. With one unit of

The parameter $\theta > 0$ determines the relative importance of the jurisdiction's tax compared to the other cost for net profits.

A firm maximizes the expected discounted sum of its net profits and chooses its location in every period, after observing the current tax profile of the two players $\tau_t \equiv (\tau_{\mathrm{H},t}, \tau_{\mathrm{L},t})$:

$$J_k(\ell_{t-1}, \tau_t) = \max_{\ell \in \{\mathrm{H},\mathrm{L}\}} \left\{ \pi_\ell(k; \tau_{\ell,t}) - C\, \mathbb{I}_{[\ell \neq \ell_{t-1}]} + \delta_F \mathbb{E}_t \left[J_k(\ell, \tau_{t+1}) \right] \right\}, \tag{2}$$

where ℓ_{t-1} is the firm's location in the previous period, C is a fixed cost of relocation, $\delta_F \in (0,1)$ is the firms' discount factor, and next period's tax rates are stochastic because Player i adjusts its tax rate with probability λ_i, see Section 2.3.

In this article, I assume $C = 0$, that is, relocation is costless. When tax competition is mainly about making firms declare their profits within one's jurisdiction — for instance holding companies or letterbox firms — costs of relocating the legal seat may be small and the assumption warranted. When tax competition is mainly about attracting firms' real investments, the assumption is stark. Taking this into account would, however, introduce an additional state variable into the players' problem and is beyond the scope of this paper. See Paulus et al. (2026) for a dynamic analysis of the global minimum tax with reallocation of mobile capital.

Since each firm has measure zero, its choice does not influence the tax profile τ_{t+1} chosen by the players. With $C = 0$, the previous location ℓ_{t-1} drops out of the right-hand side of Equation (2), so J_k does not depend on it. The continuation value $\delta_F\, \mathbb{E}_t \left[J_k(\ell, \tau_{t+1}) \right]$ is therefore the same for both locations ℓ and does not affect the firm's choice. Hence, the firms' maximization problem reduces to repeatedly choosing the jurisdiction that maximizes their short-term net profits $\pi_\ell(k; \tau_{\ell,t})$ from Equation (1).

Assume that gross profits are the same for all firms, $\bar{\pi}(k) = \bar{\pi}$, so that firms differ only in their location k, and normalize $\bar{\pi} = 1$. After this normalization, θ can be interpreted as a cost per unit of gross profit, making it comparable with tax rates. Then, for each pair $(\tau_\mathrm{H}, \tau_\mathrm{L}) \in [0,1] \times [0,1]$ with $|\tau_\mathrm{L} - \tau_\mathrm{H}| \leq \theta$, the firm at location $x_\mathrm{H} = 1 + (\tau_\mathrm{L} - \tau_\mathrm{H})/\theta \in [0,2]$ is indifferent between locating in the two jurisdictions. All firms $k < x_\mathrm{H}$ locate in jurisdiction H, all others in jurisdiction L. If firms are assumed to be uniformly distributed on the interval $[0,2]$, x_H corresponds to the number of firms that locate in jurisdiction H. By the same token, $x_\mathrm{L} \equiv 2 - x_\mathrm{H}$ is the number of firms that locate in jurisdiction L.[6]

Thus, from the perspective of player i, instantaneous returns to choosing tax rate τ_i are given by

$$r^i(\tau_\mathrm{H}, \tau_\mathrm{L}) = (\tau_i + a_i)\, x_i(\tau_\mathrm{H}, \tau_\mathrm{L}), \qquad x_i(\tau_\mathrm{H}, \tau_\mathrm{L}) = 1 + (\tau_{-i} - \tau_i)/\theta \tag{3}$$

where x_i is the tax base, that is, the number of firms with unit gross profit on the territory of player i, which depends on the tax rates of both players. The parameter a_i captures the benefit

firms operating in each jurisdiction and a cost of $\theta s^2/2$ for shifting the share s of profit, a firm in the high-tax jurisdiction shifts $s = |\tau_\mathrm{L} - \tau_\mathrm{H}|/\theta$, which gives the tax base x_i in Equation (3) below.

[6] For tax differentials exceeding θ in absolute value, the indifference point lies outside $[0,2]$ and all firms locate in the low-tax jurisdiction. I nevertheless use the linear expressions for x_H and x_L at all tax profiles, so that the returns in Equation (3) below are quadratic on the whole state space. In the calibration of Section 5.1, the tax differential at the steady states and at the static Nash equilibrium reported in Table 2 is at most 0.118, compared with $\theta = 0.227$.

of x_i that goes beyond the direct tax revenue, for instance, employment opportunities on the territory of player i. Without loss of generality, the players are labeled such that $a_{\mathrm{H}} \leq a_{\mathrm{L}}$. Player H is then the high-tax and Player L the low-tax player, in the static Nash equilibrium of Section 2.2 as well as in the steady state of the affine MPE of Section 3.2. Henceforth, subscripts denote partial derivatives, $r^i_j \equiv \partial r^i/\partial \tau_j$ and $r^i_{jj'} \equiv \partial^2 r^i/\partial\tau_j\partial\tau_{j'}$, e.g., r^i_i and r^i_{-i}. For future reference, the partial derivatives of instantaneous returns are given by

$$r^i_i(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = 1 + \frac{\tau_{-i} - \tau_i}{\theta} - \frac{a_i + \tau_i}{\theta} = 1 + \frac{\tau_{-i} - 2\tau_i - a_i}{\theta}, \tag{4}$$

$$r^i_{-i}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = \frac{\tau_i + a_i}{\theta}, \tag{5}$$

$$r^i_{ii}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = -\frac{2}{\theta}, \tag{6}$$

$$r^i_{i,-i}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = \frac{1}{\theta}. \tag{7}$$

Since $r^i_{i,-i}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = 1/\theta > 0$ (Equation (7)), the tax decisions of the two players are strategic complements. As discussed in the introduction, this is the empirically relevant case, in the international sphere as well as in business tax competition between German municipalities.

I now briefly discuss the Nash equilibrium from a static simultaneous-move game before describing the dynamic setup used in the rest of the article.

2.2 Static Nash Equilibrium

In a static, simultaneous-move, two-player game where Players' returns r^i are described by (3), Equation (4) can be used to obtain the following first-order condition for a maximum,

$$1 + \frac{\tau_{-i} - \tau_i}{\theta} - \frac{a_i + \tau_i}{\theta} = 0.$$

Because r^i is concave in the own tax rate (Equation (6)), it is also a sufficient condition. The first-order condition weighs two opposing forces: the first two terms, equal to the tax base $x_i = 1 + (\tau_{-i} - \tau_i)/\theta$, are the revenue gained on the existing base per unit increase in τ_i (the intensive margin), while the last term is the erosion of that base — the higher rate drives $1/\theta$ firms per unit to the competitor, each worth $\tau_i + a_i$ (the extensive margin). Optimality dictates balancing the two. The first-order condition implies the (static) best response function

$$\tau^{\mathrm{BR}}_i(\tau_{-i}) = \frac{\theta - a_i + \tau_{-i}}{2}, \tag{8}$$

for $i = \mathrm{H}, \mathrm{L}$, with slope $1/2$ in the opponent's tax rate. The (static) Nash equilibrium of the game is given by the strategy profile $(\tau^{\mathrm{NE}}_{\mathrm{H}}, \tau^{\mathrm{NE}}_{\mathrm{L}})$. Using these best response functions for both players and solving for the fix point of the resulting coupled system of equations gives the tax profile resulting in the Nash equilibrium

$$\left(\tau^{\mathrm{NE}}_{\mathrm{H}}, \tau^{\mathrm{NE}}_{\mathrm{L}}\right) = \left(\theta - \frac{2a_{\mathrm{H}} + a_{\mathrm{L}}}{3}, \theta - \frac{2a_{\mathrm{L}} + a_{\mathrm{H}}}{3}\right). \tag{9}$$

Since $a_{\mathrm{H}} \leq a_{\mathrm{L}}$, Player H sets the higher rate, as $\tau^{\mathrm{NE}}_{\mathrm{H}} - \tau^{\mathrm{NE}}_{\mathrm{L}} = (a_{\mathrm{L}} - a_{\mathrm{H}})/3 \geq 0$. The static Nash equilibrium will be used as a benchmark later on.

2.3 Dynamic Game

I now introduce the dynamic version of the game analyzed in the rest of the article. Players discount the future with the discount factor $\delta \in (0,1)$ per period. Player i's intertemporal payoff at time t is given by

$$U_t^i = \sum_{s=0}^{\infty} \delta^s r^i(\tau_{i,t+s}, \tau_{-i,t+s}). \tag{10}$$

U_t^i is player i's objective function and it is defined over the complete time horizon of the game, which is infinite and in discrete time steps. When moving in a period t, player i's action is to choose a tax rate $\tau_{i,t} \in [\tau_0, 1] \equiv \mathcal{A}(\tau_0)$, where $\tau_0 < 1$ is a lower bound on tax rates.[7] Consequently, the action space $\mathcal{A}(\tau_0)$ is infinite. I assume that players, when moving, base their decision on payoff-relevant variables only. That is, when choosing their tax rate $\tau_{i,t+1}$, the only information player i takes into account is the other player's tax rate $\tau_{-i,t}$. In other words, I restrict attention to Markovian strategies. I denote these strategies as $T_i : \mathcal{A}(\tau_0) \to \mathcal{A}(\tau_0)$, $\tau_{-i,t} \mapsto \tau_{i,t+1}$.

Therefore, the state of the game at time t is given by the current tax profile of the two players $(\tau_{\mathrm{H},t}, \tau_{\mathrm{L},t})$. Thus, the state space is given by $\mathcal{S} \equiv \mathcal{A}(\tau_0) \times \mathcal{A}(\tau_0) = [\tau_0, 1]^2$ and is infinite, too, though compact.

The game considered is one of deterministic transitions. If the state of the game at time t is given by $(\tau_{\mathrm{H},t}, \tau_{\mathrm{L},t})$ and Player H moves in period $t+1$, choosing τ'_{H}, the state of the game at time $t+1$ is given by $(\tau_{\mathrm{H},t+1}, \tau_{\mathrm{L},t+1}) = (\tau'_{\mathrm{H}}, \tau_{\mathrm{L},t})$ with certainty.

The timing of the moves deviates from the setting in Maskin and Tirole (1987), in which players move in strict alternation. Here, a random process governs whether Player H, Player L, or neither of them moves in a given period. This allows the probability of moving to be calibrated to the observed frequency of tax rate changes (Section 5.1). I assume that players never move contemporaneously, which makes the game considered one of asynchronous moves. Hence, in every period t, there are three possible cases that arise with a given probability: Player H moves with probability λ_{H}, Player L moves with probability λ_{L}, no player moves with the residual probability $\lambda_\emptyset \equiv 1 - \lambda_{\mathrm{H}} - \lambda_{\mathrm{L}}$. The assumption of asynchronous moves is crucial for the results, because the opponent moves while the player's tax rate is in place and replies to it, similar to the follower in the Stackelberg games of Section 1 (see Section 3.1).

The following *dwell factor* will prove useful in what follows:

$$D_i := \sum_{t \geq 0} [\delta(1-\lambda_i)]^t = \frac{1}{1-\delta(1-\lambda_i)}, \ i = \mathrm{H}, \mathrm{L}. \tag{11}$$

It is the expected discounted value of a constant flow of unit returns from the current period until player i next moves and can be interpreted as the effective period length over which the

[7] The action space can be expanded to negative values to account for subsidies. Although I abstract from modeling the players' budget explicitly, the latter will always impose an upper bound on the feasible subsidy and thus a lower bound on $\tau_{i,t}$. Tax rates at or close to 1 eliminate any incentive for economic activity in the first place and are never chosen in equilibrium; throughout, I therefore restrict attention to steady states with rates below one. The upper endpoint is nevertheless included, so that $\mathcal{A}(\tau_0)$ is a closed interval. The lower endpoint must be included, because a player bound by a minimum tax sets its rate exactly at τ_0 (Section 4).

instantaneous returns accrue. A second timing object is the composite

$$\omega_i := \sum_{s\geq 1} \delta^s \lambda_i (1-\lambda_i)^{s-1} = \delta \lambda_i D_i \in (0,1),\ i = \mathrm{H}, \mathrm{L}. \tag{12}$$

The sum identifies ω_i as the expected discounted value of a single unit return received at the (random) date at which player i next moves, that is, the expected discount factor applied to that date: player i's next move arrives s periods ahead with probability $\lambda_i(1-\lambda_i)^{s-1}$ and the unit return is then discounted by δ^s.[8] It can be interpreted as the effective discount factor.

In what follows, I look for pairs of Markovian strategies $(T_\mathrm{H}, T_\mathrm{L})$ that form Subgame Perfect Nash Equilibria (SPNE). The Markovian property of T_i makes any of these a Markov perfect equilibrium (MPE). In order to be a SPNE, each player's strategy T_i must maximize the player's intertemporal payoff U_t^i from Equation (10) for any time t, given the other player's strategy T_{-i} for any starting state of the game $(\tau_{\mathrm{H},t-1}, \tau_{\mathrm{L},t-1})$ — and this among all continuation strategies. To find a strategy profile that fulfills this, one can rely on the one-shot deviation principle. The latter states that it is sufficient to exclude profitable one-shot deviations for a strategy profile to be a SPNE. Combined with the Markov property of memorylessness of the strategies considered, one can further simplify the search for the strategy profile by excluding profitable one-shot deviations in the first period.

Fix a profile of Markovian strategies $(T_\mathrm{H}, T_\mathrm{L})$. Whichever of the three move events realizes in a period, all payoff-relevant information is contained in the tax profile that obtains once the period's move — if any — has taken place. Accordingly, let $V_i(\tau_\mathrm{H}, \tau_\mathrm{L})$ denote player i's expected intertemporal payoff given that the current period's *post-move* tax profile is $(\tau_\mathrm{H}, \tau_\mathrm{L})$. In a period in which no player moves, the post-move profile is the inherited profile itself, so V_i is in particular player i's value in a no-move period.

Hence, relying on recursive formulations from dynamic programming, $(T_\mathrm{H}, T_\mathrm{L})$ form a MPE if and only if there exist valuation functions $V_i : \mathcal{S} \to \mathbb{R}$, $i = \mathrm{H}, \mathrm{L}$ such that for any $(\tau_\mathrm{H}, \tau_\mathrm{L}) \in \mathcal{S}$ the following conditions hold:[9]

$$V_\mathrm{H}(\tau_\mathrm{H}, \tau_\mathrm{L}) = r^\mathrm{H}(\tau_\mathrm{H}, \tau_\mathrm{L}) + \delta \left[\lambda_\mathrm{H} \max_{\tau'_\mathrm{H} \in \mathcal{A}(\tau_0)} V_\mathrm{H}(\tau'_\mathrm{H}, \tau_\mathrm{L}) + \lambda_\mathrm{L} V_\mathrm{H}\left(\tau_\mathrm{H}, T_\mathrm{L}(\tau_\mathrm{H})\right) + \lambda_\emptyset V_\mathrm{H}(\tau_\mathrm{H}, \tau_\mathrm{L}) \right] \tag{13}$$

$$T_\mathrm{H}(\tau_\mathrm{L}) \in \operatorname*{argmax}_{\tau'_\mathrm{H} \in \mathcal{A}(\tau_0)} V_\mathrm{H}(\tau'_\mathrm{H}, \tau_\mathrm{L}) \tag{14}$$

and symmetrically for Player L, with the roles of λ_H and λ_L interchanged.

I refer to Equation (13) as the *value recursion* and to Equation (14) as the *optimality condition*. The value recursion mirrors the timing of the game. In the next period, one of the three move events occurs: Player H itself moves (probability λ_H), in which case it re-optimizes and obtains the value of the best available post-move profile, $\max_{\tau'_\mathrm{H}} V_\mathrm{H}(\tau'_\mathrm{H}, \tau_\mathrm{L})$; Player L moves

[8] That $\omega_i < 1$ follows from the inverse form of Equation (11), $1/D_i = 1 - \delta + \delta\lambda_i$: $\omega_i = \delta\lambda_i/(1-\delta+\delta\lambda_i) < 1$ if and only if $0 < 1 - \delta$.

[9] The dynamic-programming formulation is equivalent to Markov perfect equilibrium in a deterministic game of alternating moves (Maskin and Tirole 1987). Maskin and Tirole (1988) (§4) note that when commitment lengths are governed by a Poisson process, the equations characterizing Markov perfect equilibrium take the same form as under deterministic alternating moves.

(probability λ_{L}), in which case Player H's rate stays at τ_{H} while Player L's rate is replaced by $T_{\mathrm{L}}(\tau_{\mathrm{H}})$; or no player moves (probability $\lambda_{\emptyset}$), in which case the profile persists. The maximization thus applies only in the branch in which the player actually moves, and the object maximized is V_{H} itself. The value of being called to move against opponent rate τ_{L}, the object of a standard Bellman equation, is $\max_{\tau'_{\mathrm{H}}} V_{\mathrm{H}}(\tau'_{\mathrm{H}}, \tau_{\mathrm{L}}) = V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right)$. The maximum is attained at $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ by the optimality condition (14). In Appendix A.1, I show that the values of moving and of watching the opponent move obey simple recursions weighted by the dwell factors (11).

3 Markov Perfect Equilibrium

In what follows I assume that instantaneous returns $r^i(\tau_{\mathrm{H}}, \tau_{\mathrm{L}})$, $i = \mathrm{H}, \mathrm{L}$ are given by Equation (3) and therefore quadratic. The partial derivatives are collected in Equations (4)–(7). Moreover, assume that both players move with equal probability, $\lambda \equiv \lambda_{\mathrm{H}} = \lambda_{\mathrm{L}}$. It follows that both the dwell factors (11) and the expected discount factors applied to the date of a player's next move (12) are equal, $D_{\mathrm{H}} = D_{\mathrm{L}} \equiv D$ and $\omega_{\mathrm{H}} = \omega_{\mathrm{L}} \equiv \omega = \delta\lambda D$.

I start from the first-order condition that Equations (13) and (14) imply for any MPE. When called to move against τ_{L}, Player H chooses its rate τ_{H} to maximize $V_{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}})$ over $\mathcal{A}(\tau_0)$ (Equation (14)). Differentiating the value recursion (13) with respect to τ_{H}, holding τ_{L} fixed, gives

$$(1 - \delta\lambda_{\emptyset}) \frac{\partial V_{\mathrm{H}}}{\partial \tau_{\mathrm{H}}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) + \delta\lambda_{\mathrm{L}} W'_{\mathrm{H}}(\tau_{\mathrm{H}}), \tag{15}$$

where

$$W_{\mathrm{H}}(\tau_{\mathrm{H}}) := V_{\mathrm{H}}\big(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\big) \tag{16}$$

is the value of *watching* Player L move while Player H's own rate is held at τ_{H}: the post-move value at the state Player L's reply $T_{\mathrm{L}}(\tau_{\mathrm{H}})$ brings about. Of the three branches of the value recursion, only one links the own rate to a new state. If Player H itself moves next, it faces the same opponent rate and re-chooses $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ whatever its current rate, so this branch does not depend on τ_{H}. If nobody moves, the profile persists, and the branch reproduces $\partial V_{\mathrm{H}}/\partial\tau_{\mathrm{H}}$, collected on the left-hand side. If Player L moves, it replies to τ_{H}, and Player H obtains the watch value $W_{\mathrm{H}}(\tau_{\mathrm{H}})$.

At an interior optimum $\tau_{\mathrm{H}} = T_{\mathrm{H}}(\tau_{\mathrm{L}}) \in (\tau_0, 1)$ at which W_{H} is differentiable, the derivative $\partial V_{\mathrm{H}}/\partial\tau_{\mathrm{H}}$ in Equation (15) vanishes. Multiplied by D_{L}, the first-order condition of Player H reads

$$D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{H}}\big(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}\big) + \omega_{\mathrm{L}}\, W'_{\mathrm{H}}\big(\tau_{\mathrm{H}}\big) = 0. \tag{17}$$

The equation's two terms illustrate that a move sets a rate that influences Player H's payoff in two ways. First, the rate stays in force until Player L next moves — if Player H moves next, it chooses the same rate again — so its marginal return is weighted by the dwell factor D_{L} of Equation (11). Second, the rate determines Player L's reply, whose value enters through the marginal watch value, discounted by ω_{L} of Equation (12) to the date of Player L's reply. At the optimum, the two effects balance. The corresponding objects for Player L follow by relabeling $\mathrm{H} \leftrightarrow \mathrm{L}$ and $\lambda_{\mathrm{H}} \leftrightarrow \lambda_{\mathrm{L}}$.

Before turning to steady states, one property holds for the strategies of any MPE. Since tax rates are strategic complements ($r^i_{i,-i} > 0$, Equation (7)), the strategies are upward sloping; Proposition 2 uses this to sign the slope of the opponent's strategy.

Lemma 1. *Let $(T_{\mathrm{H}}, T_{\mathrm{L}})$ be an MPE. Then $T_i(\tau_{-i}) \geq T_i(\tau'_{-i})$ for all $\tau_{-i}, \tau'_{-i} \in \mathcal{A}(\tau_0)$ with $\tau_{-i} > \tau'_{-i}$, $i = \mathrm{H}, \mathrm{L}$.*

Proof. See Appendix A.2. □

Lemma 1 carries the monotonicity of the static best response, which follows from strategic complementarity, over to Markovian strategies: in any MPE, a higher opponent rate never lowers a player's reply.

3.1 Steady-State Tax Profiles

In what follows, I denote the steady-state tax rates of an arbitrary MPE by $(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}})$ and reserve $(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}})$ for the affine equilibrium of Proposition 3. The following proposition is one of the core results of the paper. It starts from a player's first-order condition evaluated at a steady state and derives a *wedge representation* of the player's steady-state tax rate, at any steady state of any MPE at which local regularity conditions hold for that player. This wedge representation writes each player's steady-state rate as its static best response to the opponent's rate plus the excess over it, the *incentive wedge* (henceforth wedge). The wedge measures a player's *dynamic incentive*: because the opponent's next rate responds to the player's own rate, the player gains from raising its rate above the static best response to pull the opponent's next rate up. Evaluated at the affine equilibrium of Proposition 3 below, the wedges generate the dynamic premium of Corollary 4. The proposition does not claim existence or stability of a steady state, local or global.

Proposition 2 (Wedge representation)**.** *Assume equal probability of moving for Players H and L, $\lambda_{\mathrm{H}} = \lambda_{\mathrm{L}} \equiv \lambda$. Let $(T_{\mathrm{H}}, T_{\mathrm{L}})$ be an MPE and let $(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}}) \in [\tau_0, 1]^2$ be a steady state of the induced dynamics, i.e., $T_{\mathrm{H}}(\hat{\tau}_{\mathrm{L}}) = \hat{\tau}_{\mathrm{H}}$ and $T_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}) = \hat{\tau}_{\mathrm{L}}$. Fix a player $i \in \{\mathrm{H}, \mathrm{L}\}$ and suppose that*

(a) $\tau_0 < \hat{\tau}_i < 1$ (player i interior),

(b) the value functions of player i and its own strategy T_i are differentiable in a neighborhood of the steady state (one-sided at the boundary of $[\tau_0, 1]^2$),

(c) the opponent's strategy T_{-i} is differentiable at $\hat{\tau}_i$ with $\hat{m}_{-i} \equiv T'_{-i}(\hat{\tau}_i) < 1$; recall that $\hat{m}_{-i} \geq 0$ is guaranteed by Lemma 1.

Then, the steady state satisfies the wedge representation

$$\hat{\tau}_i = \underbrace{\frac{\theta - a_i + \hat{\tau}_{-i}}{2}}_{\text{static best response}} + \underbrace{\frac{\omega \hat{m}_{-i}}{2}\left(\hat{\tau}_i + a_i\right)}_{\text{incentive wedge}}, \tag{18}$$

with closed form

$$\hat{\tau}_i = \frac{\theta - (1 - \omega \hat{m}_{-i})\, a_i + \hat{\tau}_{-i}}{2 - \omega \hat{m}_{-i}}, \qquad \omega \hat{m}_{-i} \in [0, 1). \tag{19}$$

Provided $\hat{\tau}_i + a_i > 0$, the wedge is zero if and only if $\hat{m}_{-i} = 0$, and $\hat{\tau}_i$ is strictly increasing in both $\hat{m}_{-i}$ and $\hat{\tau}_{-i}$.

Proof. See Appendix A.5. □

The proviso $\hat{\tau}_i + a_i > 0$ is technical. The sum $\hat{\tau}_i + a_i$ is player i's effective marginal benefit per unit of tax base. The proviso fails only if a player is harmed, on net, by hosting its tax base. While optimality alone does not exclude such counter-intuitive configurations, the calibration has $\hat{\tau}_i + a_i > 0$ for both players.

It is instructive to study each player's first-order condition (Equation (17) for Player H), from which the wedge representation (18) is derived. Evaluated at a steady state $\hat{\tau} \equiv (\hat{\tau}_\mathrm{H}, \hat{\tau}_\mathrm{L})$ at which conditions (a)–(c) hold for player i, and with the marginal watch value of Lemma 9 (Appendix A.3), the first-order condition reduces to

$$r_i^i(\hat{\tau}) = -\omega \hat{m}_{-i}\, r_{-i}^i(\hat{\tau}) = -\omega \hat{m}_{-i} \frac{\hat{\tau}_i + a_i}{\theta}, \tag{20}$$

where I plugged in the expression for $r_{-i}^i(\hat{\tau})$ from Equation (5). Because $r_i^i(\tau_\mathrm{H}, \tau_\mathrm{L}) = 0$ is the static first-order condition behind the best response function (8), Equation (20) sheds light on the trade-off between the static and the dynamic incentive behind the incentive wedge, the core mechanism of the model. Provided $\hat{\tau}_i + a_i > 0$ and $\hat{m}_{-i} > 0$, the player accepts a static marginal loss on its own rate, $r_i^i(\hat{\tau}) < 0$, equal to the discounted gain from the opponent's reply. The latter has three factors: how soon the reply comes (ω), how strong the reply is ($\hat{m}_{-i}$), and what the backflow of the tax base caused by the reply is worth ($(\hat{\tau}_i + a_i)/\theta$). The dynamic incentive is this discounted marginal gain from the opponent's reply, and each player raises its rate until its static marginal loss offsets it. It relies on the players' asynchronous moves, as the opponent replies to the player's tax rate while this rate is still in place. If both players moved simultaneously in every period, the previous tax profile would not be payoff relevant, players would not condition on it, and the MPE would repeat the static Nash equilibrium of Section 2.2.

For fixed $\hat{\tau}_i + a_i$, the wedge depends only on the timing of the opponent's reply, summarized in ω, and on the strength of the reply, characterized by $\hat{m}_{-i}$. It vanishes if the opponent does not respond ($\hat{m}_{-i} = 0$) or if its reply lies far in the future ($\omega \to 0$ as $\lambda \to 0$). Intuitively, in these cases the relative importance of the dynamic incentive shrinks. In other words, influencing the opponent's future rate is less valuable and the steady-state rates approach the static best responses. Interestingly, given the steady state and the opponent's slope, the wedge does not depend on θ, the sensitivity of the tax base to the tax differential from Section 2.1. The reason is that θ scales the cost of raising the tax rate above the static best response, the static marginal loss $-r_i^i(\hat{\tau}) = \frac{2}{\theta}\big(\hat{\tau}_i - \tau_i^{\mathrm{BR}}(\hat{\tau}_{-i})\big)$ (see Equations (4) and (8)), and its benefit, the backflow of the base caused by the opponent's reply, $r_{-i}^i = (\hat{\tau}_i + a_i)/\theta$, alike: both arise from firms relocating in response to the tax differential.

The incentive wedge is a player's deviation from its static best response, given the opponent's rate. Section 4 shows that a minimum tax can remove the high-tax player's wedge, which drives the central result of this paper.

3.2 Affine Strategies without a Binding Floor

In what follows, I narrow down the set of considered Markovian strategies to affine-linear strategies of the form

$$T_i(\tau_{-i}) = k_i + m_i \tau_{-i}, \tag{21}$$

with intercept k_i and slope m_i. I refer to this combination of quadratic returns and affine-linear strategies as the "linear–quadratic" setting. It is shown that the slopes coincide, $m_{\mathrm{H}} = m_{\mathrm{L}} \equiv m$. Given quadratic returns, I follow the procedure by Maskin and Tirole (1987) in a structurally similar duopoly model and guess that equilibrium strategies are affine. The proposition then verifies the guess: there is a unique pair of affine strategies with slopes in $(0, 1)$ that satisfies both players' first-order conditions at every state, and the mover's objective is concave, so that the strategies are optimal replies to each other. Section 6 discusses a property that may make this MPE worth studying beyond its analytical tractability.

The proposition requires the bounds of the action space $\mathcal{A}(\tau_0)$ never to bind, including off the equilibrium path; that is, $T_i(\tau_{-i}) = k_i + m\tau_{-i} \in [\tau_0, 1]$ must hold for all $\tau_{-i} \in [\tau_0, 1]$ (the case with a binding lower bound τ_0 is the subject of Section 4). This requirement is equivalent to the four conditions $k_i + m \leq 1$ and $\tau_0 \leq k_i/(1-m)$ for $i = \mathrm{H}, \mathrm{L}$. Because $a_{\mathrm{H}} \leq a_{\mathrm{L}}$ entails that Player L has the smaller and Player H the larger intercept, $k_{\mathrm{L}} \leq k_{\mathrm{H}}$ (Equation (24) with $\eta \geq 0$), the requirement is equivalent to the two conditions $k_{\mathrm{H}} + m \leq 1$ and $\tau_0 \leq k_{\mathrm{L}}/(1-m)$.

The next proposition shows that such an MPE exists and is unique among affine MPE with slopes in $(0, 1)$, and that its tax profile converges to a steady state from any initial profile. The steady state and the intercepts k_i are given in closed form in the primitives and the slope m.

Proposition 3 (Affine MPE)**.** *Assume equal probability of moving for Players H and L, $\lambda_{\mathrm{H}} = \lambda_{\mathrm{L}} \equiv \lambda$. Let m be the smaller root of*

$$1 - 2(1+\omega)m + 2\omega m^2 + \omega^2 m^4 = 0, \tag{22}$$

which lies in $(0, 1/2)$, and define

$$\sigma \equiv \frac{\theta}{1-\omega m} - \frac{a_{\mathrm{H}} + a_{\mathrm{L}}}{2}, \qquad \eta \equiv \frac{(1-\omega m)(a_{\mathrm{L}} - a_{\mathrm{H}})}{2(3-\omega m)}, \tag{23}$$

$$k_{\mathrm{H}} \equiv (1-m)\,\sigma + (1+m)\,\eta, \qquad k_{\mathrm{L}} \equiv (1-m)\,\sigma - (1+m)\,\eta. \tag{24}$$

Since $a_{\mathrm{H}} \leq a_{\mathrm{L}}$, $\eta \geq 0$ and therefore $k_{\mathrm{L}} \leq k_{\mathrm{H}}$. Assume that $\tau_0 \leq k_{\mathrm{L}}/(1-m)$ and $k_{\mathrm{H}} + m \leq 1$. Then $T_i(\tau_{-i}) = k_i + m\,\tau_{-i}$, $i = \mathrm{H}, \mathrm{L}$, is the unique MPE in affine strategies with slopes in $(0, 1)$. Its steady state is $(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}}) = (\sigma + \eta,\ \sigma - \eta)$, and from any initial profile in $[\tau_0, 1]^2$ the tax profile converges to it with probability one.

Proof. See Appendix A.4. □

For these affine strategies, the steady state satisfies the wedge representation of Proposition 2 and its closed form for both players with $\hat{m}_{-i} = m$ (Appendix A.4). The slope m is the same for both players although their non-tax benefits differ: a_i enters the marginal returns r^i_i and r^i_{-i} only as a constant and drops out of all second derivatives (Equations (6) and (7)), so it shifts the intercept of a player's strategy but not its slope.

The constants σ and η of Equation (23) have a direct reading: $\sigma = \frac{1}{2}(\tau_{\mathrm{H}}^* + \tau_{\mathrm{L}}^*)$ is the mean of the two steady-state tax rates and $\eta = \frac{1}{2}(\tau_{\mathrm{H}}^* - \tau_{\mathrm{L}}^*)$ is half the gap between them. Since $a_{\mathrm{H}} \leq a_{\mathrm{L}}$, $\eta \geq 0$, so Player H's steady-state tax rate is at least as high as Player L's.

The slope m of the affine strategies is smaller than the slope $1/2$ of the static best responses (Equation (8)). The intuition for this is as follows. When player i moves and chooses a tax rate τ_i, its current return $r^i(\tau_i, \tau_{-i})$ depends on the opponent's *current* rate τ_{-i}; the static best response maximizes this return alone. However, the mover knows that once the opponent moves, the opponent's rate changes to $T_{-i}(\tau_i)$, which depends on τ_i and no longer on τ_{-i}; player i lives with the resulting return $r^i(\tau_i, T_{-i}(\tau_i))$ until it moves next (the source of the dynamic incentive in the steady-state first-order condition (20)). The first of these returns depends on τ_{-i}, the second one is fully independent of it. The second return is also concave in τ_i: since the opponent follows a higher rate only partly ($m < 1$), a higher τ_i still widens the tax differential and shrinks the base, so the marginal return to raising τ_i falls as τ_i rises. The mover's choice is therefore a compromise between the static best response to τ_{-i} and a rate that would maximize returns against the opponent's reply and does not depend on τ_{-i}. Hence the opponent's current rate carries less weight in the mover's choice than in the static best response, and the mover reacts less strongly to it.

In line with this intuition, reducing the importance of the opponent's future moves, by either decreasing the discount factor $\delta \to 0$ or the probability of moving $\lambda \to 0$, leads to $m \to 1/2$, the static best response's slope, monotonically.[10]

The affine steady state $(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$ of Proposition 3 is the benchmark with which the tax profile under a minimum tax is compared in Section 4. The following subsection first characterizes the wedges at this steady state and its distance from the static Nash equilibrium $(\tau_{\mathrm{H}}^{\mathrm{NE}}, \tau_{\mathrm{L}}^{\mathrm{NE}})$.

3.3 Incentive Wedge and Dynamic Premium

Proposition 3 established that the slope of the affine strategy is positive. I now show that this means that each player's wedge is positive and that its steady-state tax rate from Proposition 3 strictly exceeds the static Nash equilibrium rate from Equation (9). I call this difference between the two tax rates of player i, $\Delta_i \equiv \tau_i^* - \tau_i^{\mathrm{NE}}$, its *dynamic premium*. I further establish a relationship between each player's wedge and its dynamic premium.

At the affine MPE of Proposition 3, the wedge representation (18) holds for both players with $\hat{m}_{-i} = m$. Since $m > 0$, provided $\tau_i^* + a_i > 0$, each player's wedge is strictly positive:

$$w_i \equiv \tau_i^* - \tau_i^{\mathrm{BR}}(\tau_{-i}^*) = \frac{\omega m}{2}(\tau_i^* + a_i) > 0, \quad \text{equivalently} \quad r_i^i(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) < 0, \qquad i = \mathrm{H}, \mathrm{L}, \tag{25}$$

where the equivalence follows from Equation (20) and $r_{-i}^i(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) = (\tau_i^* + a_i)/\theta$ (Equation (5)).

Player L's wedge is never smaller than Player H's. By Equation (25), the two wedges differ only through the value of a firm to the player, $\tau_i^* + a_i$. With $\tau_{\mathrm{H}}^* - \tau_{\mathrm{L}}^* = 2\eta$ and η from

[10] The slope m depends on δ and λ only through $\omega = \delta\lambda D = \delta\lambda/\big(1 - \delta(1-\lambda)\big)$, which is increasing in both; by the implicit function theorem applied to Equation (22), m is decreasing in ω, and $m \to 1/2$ as $\omega \to 0$.

Equation (23),

$$\begin{aligned}(\tau_{\mathrm{L}}^{*}+a_{\mathrm{L}})-(\tau_{\mathrm{H}}^{*}+a_{\mathrm{H}}) &= a_{\mathrm{L}}-a_{\mathrm{H}}-2\eta=\frac{2\,(a_{\mathrm{L}}-a_{\mathrm{H}})}{3-\omega m}\geq 0,\\ w_{\mathrm{L}}-w_{\mathrm{H}} &= \frac{\omega m\,(a_{\mathrm{L}}-a_{\mathrm{H}})}{3-\omega m}\geq 0,\end{aligned} \tag{26}$$

where non-negativity follows from $a_{\mathrm{L}} \geq a_{\mathrm{H}}$ and $0 < \omega m < 1$. The intuition is that both players obtain the same response to an increase in their own rate — the opponent follows with slope m, at a date discounted by ω — but differ in what a returning firm is worth to them. Player L's higher non-tax benefit makes it the low-tax player: it sets the lower rate to attract firms. Although it thus taxes each firm less, its higher non-tax benefit more than makes up for it, so a firm is worth more to Player L than to Player H (first line of Equation (26)). Pulling the opponent's rate up, which brings firms back, therefore pays more for Player L, and it sets its rate further above its static best response.

Next, write the wedge representation (18) at the affine steady state as $2\tau_i^{*}-\tau_{-i}^{*}=\theta-a_i+2w_i$ and subtract $2\tau_i^{\mathrm{NE}}-\tau_{-i}^{\mathrm{NE}}=\theta-a_i$, obtained from the static Nash condition $\tau_i^{\mathrm{NE}}=(\theta-a_i+\tau_{-i}^{\mathrm{NE}})/2$. The difference is $2\Delta_i-\Delta_{-i}=2w_i$, which gives

$$\Delta_i = w_i+\tfrac{1}{2}\,\Delta_{-i} \quad\Longleftrightarrow\quad \Delta_i=\tfrac{2}{3}\,(2w_i+w_{-i}), \qquad i=\mathrm{H},\mathrm{L}. \tag{27}$$

Because both wedges w_i are positive, so are both dynamic premia Δ_i. This is summarized in the following corollary:

Corollary 4 (Dynamic premium)**.** *Under the conditions of Proposition 3, let* $(\tau_{\mathrm{H}}^{*},\tau_{\mathrm{L}}^{*})$ *denote the steady-state tax profile of the affine MPE of the same proposition, and let* $(\tau_{\mathrm{H}}^{\mathrm{NE}},\tau_{\mathrm{L}}^{\mathrm{NE}})$ *denote the static simultaneous-move Nash equilibrium tax profile given in Equation (9). Provided* $\tau_i^{*}+a_i>0$ *for* $i=\mathrm{H},\mathrm{L}$*, the dynamic incentive gives each player a strictly positive incentive wedge,* $w_i>0$*, and raises each player's steady-state tax rate above the static benchmark: both dynamic premia are strictly positive,* $\Delta_i=\tau_i^{*}-\tau_i^{\mathrm{NE}}>0$*,* $i=\mathrm{H},\mathrm{L}$*.*

The corollary formalizes a dynamic premium that the sequential structure of the game introduces, namely that the steady-state tax rates of the affine MPE exceed their static counterparts, a feature shared with Stackelberg models of tax competition (Wang 1999; Konrad 2009). Figure 1a in Section 4 draws the affine strategies and the static best responses at the calibration of Section 5.1. It illustrates Player H's wedge, the distance between the Markovian strategy and the steeper static best response at the affine steady-state tax profile $\tau^{*}\equiv(\tau_{\mathrm{H}}^{*},\tau_{\mathrm{L}}^{*})$, as well as its dynamic premium, the vertical distance between the affine steady state and the static Nash equilibrium.

The characterization of the wedge and the premium closes the analysis of the affine steady state. The next section analyzes the introduction of a minimum tax.

4 Introduction of a Minimum Tax

This section introduces a minimum tax that binds the low-tax player and studies its effect on the steady states of MPE. Throughout this section, I assume $\lambda_{\mathrm{H}}=\lambda_{\mathrm{L}}$, $k_{\mathrm{H}}+m\leq 1$ (Proposition 3)

and $\tau_i^* + a_i > 0$ for $i = \mathrm{H}, \mathrm{L}$, and compare two games that differ only in the lower bound τ_0 of the action space. In the benchmark game, $\tau_0 \leq k_{\mathrm{L}}/(1-m)$, so the affine MPE of Proposition 3 applies; its steady state $(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$ does not depend on τ_0. In the game with a minimum tax, the lower bound is $\tau_0 = \underline{\tau} \in [\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$.

The upper bound $\theta - a_{\mathrm{H}}$ is imposed because in the local MPE of Proposition 5, minimum taxes $\underline{\tau} \geq \theta - a_{\mathrm{H}}$ would be binding for both players. The lower bound τ_{L}^*, Player L's tax rate from the affine steady state of Proposition 3, is chosen because of its relevance for the policy question: "How to prevent a race to the bottom arising from tax competition in strategic complements?" A minimum tax below τ_{L}^* does not bind Player L at the affine steady state.[11]

The interval $[\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$ is non-empty if and only if Player L's non-tax benefit exceeds Player H's by enough:

$$\tau_{\mathrm{L}}^* = \sigma - \eta < \theta - a_{\mathrm{H}} \quad \Longleftrightarrow \quad a_{\mathrm{L}} - a_{\mathrm{H}} > \theta \frac{\omega m(3 - \omega m)}{(1 - \omega m)(2 - \omega m)}, \tag{28}$$

where the equivalence is obtained by plugging in the expressions for σ and η from Equation (23). Symmetric players fail this condition; the calibration of Section 5.1 satisfies it.

4.1 Steady States under a Minimum Tax Rate

This subsection constructs an MPE in which Player L's strategy is to play the floor $\underline{\tau}$, so that Player L does not respond to an increase of Player H's rate. In its steady state, Player H plays its static best response to the floor. In terms of the wedge representation (18), the slope of Player L's strategy is zero, $\hat{m}_{\mathrm{L}} = 0$, and the floor thus removes Player H's wedge, which is strictly positive at the affine MPE (Equation (25)). Intuitively, the minimum tax delivers what Player H's high rate previously did: it keeps Player L's rate high.

Player L stays at the floor only as long as Player H's rate is not too high: a higher rate of Player H enlarges Player L's tax base and raises its return to leaving the floor. The following proposition therefore considers a game in which both players' rates are bounded from above, Player H's by b and Player L's by $\underline{\tau} + \varepsilon$; I refer to an MPE of this game as a *local MPE*. The bound b can be chosen above Player H's rate in the affine steady state, τ_{H}^*.

Proposition 5 (Local MPE at the floor)**.** *Let the assumptions of this section hold, let* $\underline{\tau} \in [\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$*, and let*

$$\hat{\tau}_{\mathrm{H}} \equiv \frac{\theta - a_{\mathrm{H}} + \underline{\tau}}{2} \tag{29}$$

denote Player H's static best response to the floor. There exist $b \in (\hat{\tau}_{\mathrm{H}}, 1]$ *and* $\varepsilon > 0$ *such that, in the game with action sets* $\mathcal{A}_{\mathrm{H}} = [\underline{\tau}, b]$ *and* $\mathcal{A}_{\mathrm{L}} = [\underline{\tau}, \underline{\tau} + \varepsilon]$*, the strategies*

$$T_{\mathrm{H}}(\tau_{\mathrm{L}}) = \hat{\tau}_{\mathrm{H}} + \frac{\tau_{\mathrm{L}} - \underline{\tau}}{2(1 + \omega)}, \qquad T_{\mathrm{L}}(\tau_{\mathrm{H}}) = \underline{\tau} \tag{30}$$

form an MPE. Its steady state is $(\hat{\tau}_{\mathrm{H}}, \underline{\tau})$*, and from any initial profile in* $\mathcal{A}_{\mathrm{H}} \times \mathcal{A}_{\mathrm{L}}$ *the tax profile reaches it in finite time with probability one.*

[11] The restriction to minimum taxes not below τ_{L}^* is used for four results: Player L's optimality at the floor (Proposition 5), the sign of the complementarity effect (Corollary 6), Player H's welfare gain (Corollary 7) and the sign of aggregate tax revenue (Corollary 8).

Proof. See Appendix A.6. □

The proposition presupposes $\hat{\tau}_{\mathrm{H}} < 1$, so that the interval for b is not empty. This is without cost: a tax rate of one is prohibitive and would eliminate the incentive for economic activity.

In the local MPE of Proposition 5, Player L sits at the minimum tax $\underline{\tau}$, and Player H plays its static best response to it. The intuition is that, at the corner, Player L would like to undercut the minimum tax $\underline{\tau}$: its static best response lies below the floor. A higher rate of its opponent enlarges Player L's tax base, but, up to the bound b, not by enough to make leaving the floor worthwhile.

With b above τ^*_{H}, the state space $\mathcal{A}_{\mathrm{H}} \times \mathcal{A}_{\mathrm{L}}$ contains the profile $(\tau^*_{\mathrm{H}}, \underline{\tau})$ at which the game stands when a minimum tax $\underline{\tau} \leq \tau^*_{\mathrm{H}}$ pushes Player L's rate up from the affine steady state $(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}})$. From there, the game reaches the corner steady state: if Player H moves first, it sets $\hat{\tau}_{\mathrm{H}}$ at once; if Player L moves first, it stays at the floor until Player H moves. This provides an account of how the steady state of Proposition 5 can be reached after the introduction of the minimum tax.

The equilibrium is local in two senses. Player H's rates are bounded by b, and Player L's rates by $\underline{\tau} + \varepsilon$, so that large increases of Player L's rate above the floor are not available to it. Unlike Proposition 3, the proposition therefore does not establish an MPE of the game with the action space $[\underline{\tau}, 1]$ of Section 2.3, nor convergence from profiles outside $\mathcal{A}_{\mathrm{H}} \times \mathcal{A}_{\mathrm{L}}$, nor does it exclude other equilibria. Nevertheless, the local MPE is worth studying because it implies a range of minimum taxes for which Player H's steady-state rate *decreases* relative to the affine steady state of Proposition 3 (Corollary 6), a result from Stackelberg models of tax competition that simultaneous-move games have difficulties reconciling with the strategic complementarity of tax rates. Moreover, neither the bound b on τ_{H} nor the bound $\underline{\tau} + \varepsilon$ on τ_{L} binds along the path to the corner steady state after the minimum tax is introduced, so it might be possible to extend the local MPE to an MPE of the game with the action space $[\underline{\tau}, 1]$. This is not straightforward, because the bounds rule out large increases of Player L's rate above the floor and rates of Player H above b, which might induce Player L to leave the floor. I leave it for future work.

4.2 Complementarity- and Wedge Effect

Proposition 5 shows that in the steady state of the local MPE, in which Player L is pinned to the minimum tax $\underline{\tau}$ and does not respond to Player H's rate, Player H's wedge vanishes and Player H plays its static best response to the floor. I now establish under what circumstances this elimination of the wedge leads to Player H choosing a steady-state tax rate $\hat{\tau}_{\mathrm{H}}$ smaller than the steady-state tax rate τ^*_{H} from the MPE in affine strategies described in Proposition 3. To do so, I show that Player H's total tax rate change can be decomposed into two parts: the loss of the wedge, a negative pull to Player H's rate, which I call *wedge effect*; and the reaction to the minimum tax's mechanical effect on $\hat{\tau}_{\mathrm{L}}$, a positive push to Player H's rate, which I call the *complementarity effect*. The latter is increasing in $\underline{\tau}$ while the former is constant with respect to it. Thus, there is a threshold at which the two effects cancel, while the wedge effect dominates for values below the threshold and vice versa. This is made precise in the following corollary.

Corollary 6 (Decomposition and threshold)**.** *Let the hypotheses of Proposition 5 hold.*

(i) The difference between Player H's steady-state rate $\hat{\tau}_{\mathrm{H}}$ of Proposition 5 and its steady-state rate τ_{H}^* of Proposition 3 can be decomposed into a complementarity effect and a wedge effect

$$\hat{\tau}_{\mathrm{H}} - \tau_{\mathrm{H}}^* = \underbrace{\frac{\underline{\tau} - \tau_{\mathrm{L}}^*}{2}}_{\substack{\text{complementarity effect } (+): \\ \text{floor raises } \tau_{\mathrm{L}}}} - \underbrace{w_{\mathrm{H}}}_{\substack{\text{wedge effect } (-): \\ \text{wedge removed}}}, \tag{31}$$

where $w_{\mathrm{H}} = \tau_{\mathrm{H}}^* - \tau_{\mathrm{H}}^{\mathrm{BR}}(\tau_{\mathrm{L}}^*)$ is Player H's wedge at the affine MPE (Equation (25)).

(ii) Define the threshold

$$\underline{\tau}^\dagger \equiv \tau_{\mathrm{L}}^* + 2w_{\mathrm{H}} = 2\tau_{\mathrm{H}}^* - \theta + a_{\mathrm{H}} > \tau_{\mathrm{L}}^*, \tag{32}$$

the minimum tax at which the complementarity effect equals the wedge effect. For $\underline{\tau}$ larger than $\underline{\tau}^\dagger$, the complementarity effect dominates and Player H's tax rate with the minimum tax exceeds the rate without it, $\hat{\tau}_{\mathrm{H}} > \tau_{\mathrm{H}}^*$. For $\underline{\tau}$ smaller than $\underline{\tau}^\dagger$, the wedge effect dominates and the tax rate with the minimum tax falls short of the rate without it, $\hat{\tau}_{\mathrm{H}} < \tau_{\mathrm{H}}^*$. For $\underline{\tau} = \underline{\tau}^\dagger$, the two effects cancel out.

Proof. Part (i) follows from plugging Equation (29) from Proposition 5 into $\hat{\tau}_{\mathrm{H}} - \tau_{\mathrm{H}}^*$, adding and subtracting $\tau_{\mathrm{L}}^*/2$, and using the definition of w_{H} in Equation (25). The complementarity effect is non-negative because $\underline{\tau} \geq \tau_{\mathrm{L}}^*$ by assumption. The wedge effect is negative because $w_{\mathrm{H}} > 0$ by Equation (25), given $\tau_{\mathrm{H}}^* + a_{\mathrm{H}} > 0$.

For part (ii), $\underline{\tau}^\dagger > \tau_{\mathrm{L}}^*$ follows from $w_{\mathrm{H}} > 0$. By (i), $\hat{\tau}_{\mathrm{H}} - \tau_{\mathrm{H}}^* \gtreqless 0$ if and only if $(\underline{\tau} - \tau_{\mathrm{L}}^*)/2 \gtreqless w_{\mathrm{H}}$, that is, $\underline{\tau} \gtreqless \tau_{\mathrm{L}}^* + 2w_{\mathrm{H}} = \underline{\tau}^\dagger$. The second expression for $\underline{\tau}^\dagger$ follows from inserting the definition of w_{H}. □

Figure 1b illustrates the decomposition at the calibration of Section 5.1. The wedge effect is the vertical distance from Player H's strategy T_{H} down to its static best response at τ_{L}^*; the complementarity effect is the rise along this static best response as Player L's rate increases from τ_{L}^* to the floor $\underline{\tau}$.

The range of binding floors $[\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$ need not contain $\underline{\tau}^\dagger$; whether $\underline{\tau}^\dagger < \theta - a_{\mathrm{H}}$ depends on the parameters.[12] If it does not, the complementarity effect never dominates the wedge effect.

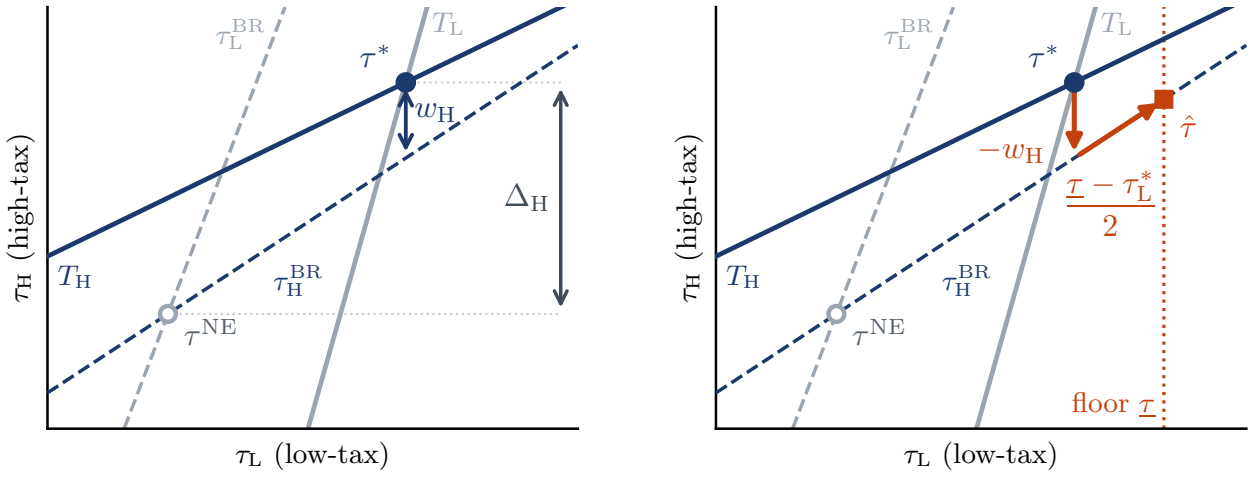
4.3 Welfare

Next, the analysis turns to the welfare of both players. I compare the steady-state flow returns of each player in the corner steady state from Proposition 5 with the flow returns in the affine

[12] Because $\underline{\tau}^\dagger = 2\tau_{\mathrm{H}}^* - \theta + a_{\mathrm{H}}$, the condition is equivalent to $\tau_{\mathrm{H}}^* < \theta - a_{\mathrm{H}}$ and hence, with $\tau_{\mathrm{H}}^* = \sigma + \eta$ from Proposition 3, to the asymmetry condition

$$a_{\mathrm{L}} - a_{\mathrm{H}} > \theta \, \frac{\omega m \, (3 - \omega m)}{1 - \omega m}.$$

This condition is strictly stronger than the non-emptiness condition for the binding floor interval in Equation (28); the two right-hand sides differ by the factor $2 - \omega m > 1$. Hence, if $a_{\mathrm{L}} - a_{\mathrm{H}}$ lies between the two thresholds, binding floors exist but every one of them lowers Player H's rate. In other words, the floor reaches Player H's own static best response before the complementarity effect catches up with the wedge effect.



(a) Without a minimum tax

(b) With a minimum tax

Figure 1: Strategies, static best responses, and steady states

Notes: Drawn to scale at the parameter values of Table 1; magnitudes in Table 2. Solid lines: strategies T_i of the affine MPE (Proposition 3); dashed lines: static best responses τ_i^{BR} (Equation (8)). Points: τ^{NE} static Nash equilibrium (Equation (9)); τ^* affine steady state (Proposition 3); $\hat{\tau}$ corner steady state (Proposition 5). Each τ denotes a tax profile, e.g., $\tau^* \equiv (\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$. Panel (a): Player H's incentive wedge w_{H} (Equation (25)) and dynamic premium Δ_{H} (Corollary 4). Panel (b): Player H's change from τ_{H}^* to $\hat{\tau}_{\mathrm{H}}$, as the wedge effect $-w_{\mathrm{H}}$ followed by the complementarity effect $(\underline{\tau} - \tau_{\mathrm{L}}^*)/2$ (Equation (31)).

steady state described in Proposition 3. In brief: in the corner steady state, Player H wins while Player L loses — at least for floors close to τ_{L}^*. The following corollary makes this precise.

Corollary 7 (Welfare effects of a binding minimum tax)**.** *Under the hypotheses of Proposition 5, steady-state flow returns compare to the affine steady state as follows:*

(i) the high-tax player strictly gains, $r^{\mathrm{H}}\big(\hat{\tau}_{\mathrm{H}}, \underline{\tau}\big) > r^{\mathrm{H}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$ *for all* $\underline{\tau} \in [\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$*;*

(ii) the low-tax player strictly loses, $r^{\mathrm{L}}\big(\hat{\tau}_{\mathrm{H}}(\underline{\tau}), \underline{\tau}\big) < r^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$*, at* $\underline{\tau} = \tau_{\mathrm{L}}^*$ *and, by continuity, for all* $\underline{\tau}$ *in a right neighborhood of* τ_{L}^**.*

Proof. (i) $r^{\mathrm{H}}\big(\hat{\tau}_{\mathrm{H}}, \underline{\tau}\big) \geq r^{\mathrm{H}}(\tau_{\mathrm{H}}^*, \underline{\tau})$ because $\hat{\tau}_{\mathrm{H}}$ is the unique maximizer of the strictly concave function $r^{\mathrm{H}}(\cdot, \underline{\tau})$ (Equation (29)), with equality only at $\underline{\tau} = \underline{\tau}^\dagger$, where $\hat{\tau}_{\mathrm{H}} = \tau_{\mathrm{H}}^*$ (Corollary 6). Moreover, $r^{\mathrm{H}}(\tau_{\mathrm{H}}^*, \underline{\tau}) \geq r^{\mathrm{H}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$ because $\underline{\tau} \geq \tau_{\mathrm{L}}^*$ and returns are strictly increasing in the opponent's rate at τ_{H}^*, $r_{\mathrm{L}}^{\mathrm{H}} = (\tau_{\mathrm{H}}^* + a_{\mathrm{H}})/\theta > 0$, independently of τ_{L}, with equality only at $\underline{\tau} = \tau_{\mathrm{L}}^*$. Since $\underline{\tau}^\dagger > \tau_{\mathrm{L}}^*$, at least one of the two inequalities is strict. The result follows by transitivity.

(ii) $\hat{\tau}_{\mathrm{H}}(\tau_{\mathrm{L}}^*) < \tau_{\mathrm{H}}^*$ by Equation (25), and r^{L} is strictly increasing in the opponent's rate ($r_{\mathrm{H}}^{\mathrm{L}} = (\tau_{\mathrm{L}} + a_{\mathrm{L}})/\theta > 0$), so at $\underline{\tau} = \tau_{\mathrm{L}}^*$ the drop of the first argument at unchanged own rate strictly lowers r^{L}; continuity of $\underline{\tau} \mapsto r^{\mathrm{L}}\big(\hat{\tau}_{\mathrm{H}}(\underline{\tau}), \underline{\tau}\big)$ extends the strict inequality to a right neighborhood of τ_{L}^*. □

Intuitively, if Player L's tax rate is pinned to the minimum tax as in the local MPE of Proposition 5, Player H no longer faces the dynamic incentive to pull up Player L's tax rate

in the future. Hence, instead of optimally balancing static and dynamic incentives, Player H can set its tax rate at its static best response to its opponent's rate, which maximizes its flow returns. In addition, the minimum tax may increase Player L's tax rate, thus narrowing the tax differential and further increasing Player H's returns. As for Player L, when the minimum tax is exactly at its tax rate without a floor, the only effect on Player L's returns r^{L} is through the change of its tax base. At the floor $\underline{\tau} = \tau_{\mathrm{L}}^*$, the complementarity effect is nil and Player H lowers its rate by its full wedge (Corollary 6), so Player L's tax base shrinks, and with it its returns, since each firm is worth $\tau_{\mathrm{L}}^* + a_{\mathrm{L}} > 0$ to it. By continuity, this extends to minimum taxes slightly above τ_{L}^*.

4.4 Aggregate Tax Revenue

The final piece of the analysis concerns aggregate tax revenue. While the welfare discussed in Section 4.3 is the relevant objective for the two players, aggregate tax revenue may be of interest from the perspective of a central government that imposes the minimum tax on them, for instance in a setting with tax competition between federal states or municipalities. Aggregate tax revenue is given by the sum of individual tax revenues $\tau_i\, x_i$. Plugging in the tax base x_i from Equation (3), this entails

$$R = \tau_{\mathrm{H}} \left(1 - \frac{g}{\theta}\right) + \tau_{\mathrm{L}} \left(1 + \frac{g}{\theta}\right) = \tau_{\mathrm{H}} + \tau_{\mathrm{L}} - \frac{g^2}{\theta}, \qquad g := \tau_{\mathrm{H}} - \tau_{\mathrm{L}}. \tag{33}$$

Intuitively, an equal split between the players, giving each of them a unit measure of firms, would give aggregate tax revenue $\tau_{\mathrm{H}} + \tau_{\mathrm{L}}$. Because the gap g between the two tax rates moves g/θ firms from H to L, each paying a tax rate lower by g, this reduces overall tax revenue by g^2/θ. In what follows, the expressions R^* and $\hat{R}(\underline{\tau})$ denote revenue in the special cases of the affine steady state and at the minimum-tax corner, respectively, with gaps $g^* := \tau_{\mathrm{H}}^* - \tau_{\mathrm{L}}^*$ and $\hat{g} := \hat{\tau}_{\mathrm{H}} - \underline{\tau} = \frac{1}{2}(\theta - a_{\mathrm{H}} - \underline{\tau})$.

Subtracting R^* from $\hat{R}(\underline{\tau})$ and adding and subtracting $\hat{\tau}_{\mathrm{H}}(1 - g^*/\theta)$ as well as $\tau_{\mathrm{L}}^*(1 + \hat{g}/\theta)$, I obtain the expression for the tax-revenue differential

$$\Delta R(\underline{\tau}) := \hat{R}(\underline{\tau}) - R^* = \underbrace{(\hat{\tau}_{\mathrm{H}} - \tau_{\mathrm{H}}^*)\left(1 - \frac{g^*}{\theta}\right)}_{(1)} + \underbrace{(\underline{\tau} - \tau_{\mathrm{L}}^*)\left(1 + \frac{\hat{g}}{\theta}\right)}_{(2)} + \underbrace{(\hat{\tau}_{\mathrm{H}} - \tau_{\mathrm{L}}^*)\,\frac{g^* - \hat{g}}{\theta}}_{(3)}, \tag{34}$$

which decomposes the change in tax revenue into three components. First and second, the inframarginal effects on the tax base that stays with Player H and Player L, respectively. Third, the effect on the tax base that Player H attracts from Player L. The first effect is negative, as long as the wedge effect dominates the complementarity effect, that is, as long as $\underline{\tau} < \underline{\tau}^\dagger$. In this case, all firms that located with Player H before the introduction of the minimum tax face a lower tax rate after the floor is installed. The second effect is positive for any minimum tax exceeding Player L's tax rate in the affine steady state, $\underline{\tau} > \tau_{\mathrm{L}}^*$, because all firms staying in Player L's jurisdiction now face a higher tax rate, and nil when $\underline{\tau} = \tau_{\mathrm{L}}^*$. The third effect is positive: the gap between the players' tax rates narrows, $g^* - \hat{g} = \frac{1}{2}(\underline{\tau} - \tau_{\mathrm{L}}^*) + w_{\mathrm{H}} > 0$ (Equation (31)), so that firms move to Player H, where they pay $\hat{\tau}_{\mathrm{H}} > \underline{\tau} \geq \tau_{\mathrm{L}}^*$. Hence, the total effect is ambiguous for minimum taxes $\underline{\tau} < \underline{\tau}^\dagger$. This is the subject of the following corollary.

Corollary 8 (Aggregate tax revenue)**.** *Under the hypotheses of Proposition 5, the tax-revenue differential ΔR of Equation (34) has the following properties:*

(i) it is strictly increasing in the minimum tax, $\Delta R'(\underline{\tau}) > 0$;

(ii) it is positive for $\underline{\tau} \geq \underline{\tau}^{\dagger}$;

(iii) it is negative for some binding minimum tax if and only if $\Delta R(\tau_{\mathrm{L}}^{*}) < 0$.

Proof. To derive Part (i), note that tax revenue at the corner is strictly increasing in the minimum tax: plugging $\hat{\tau}_{\mathrm{H}}$ from Equation (29), $\tau_{\mathrm{L}} = \underline{\tau}$, and $\hat{g}$ into Equation (33) gives

$$\hat{R}(\underline{\tau}) = \frac{\theta - a_{\mathrm{H}} + \underline{\tau}}{2} + \underline{\tau} - \frac{(\theta - a_{\mathrm{H}} - \underline{\tau})^2}{4\theta}, \qquad \hat{R}'(\underline{\tau}) = \frac{3}{2} + \frac{\hat{g}}{\theta},$$

where $\hat{R}'(\underline{\tau}) > 0$ because $\hat{g} > 0$ for any binding floor $\underline{\tau} < \theta - a_{\mathrm{H}}$. Since R^{*} does not depend on $\underline{\tau}$, the tax-revenue differential is strictly increasing in the minimum tax, too, $\Delta R'(\underline{\tau}) = \hat{R}'(\underline{\tau})$. Part (ii) follows from the analysis of the signs of the three components of the decomposition (34) above. Part (iii) follows from Part (i) because τ_{L}^{*} is the lowest binding minimum tax. □

Corollary 8 shows that a minimum tax raises aggregate tax revenue if it is high enough, at the latest from $\underline{\tau}^{\dagger}$ on. By Part (i), the tax-revenue differential changes sign at most once and is smallest at the lowest binding floor, $\underline{\tau} = \tau_{\mathrm{L}}^{*}$; a minimum tax below τ_{L}^{*} no longer binds Player L at the affine steady state. Intuitively, at this floor Player H cuts its tax rate most strongly because the complementarity effect is nil, maximally decreasing the first component of Equation (34) and letting its second component vanish.

5 Calibrated Model

In this section, the model is calibrated to the context of business tax competition between German municipalities around 2004, the year in which a minimum tax rate was enacted. It is the setting of the empirical analysis by Buettner and Poehnlein (2024) described in the introduction.

5.1 Calibration

Table 1 provides an overview of the calibrated parameter values; the details of the calibration are given in Appendix B. The model's period length is assumed to be one year. The discount factor δ is calibrated as the inverse of the gross real interest rate, $\delta = 1/(1 + \overline{r})$, where $\overline{r}$ is the average real interest rate over 1995–2001, computed from German data on 10-year federal securities and CPI inflation.

The remaining parameters are calibrated based on the so-called *Realsteuervergleich* (comparison of non-personal taxes) for the German federal state of Mecklenburg-Vorpommern (MV) for the years 2000–2007, the same data source that Buettner and Poehnlein (2024) use for seven German federal states, including MV, for the years 1999–2008. I use two of the reported variables: the business tax multiplier (*Hebesatz*, henceforth multiplier), a factor chosen by the

respective municipality that translates into the effective statutory business tax rate; and the revenue-implied tax base (*Grundbetrag*), a proxy for the sum of the assessed tax bases (*Steuermessbeträge*) of the firms located in the municipality.

Based on all municipalities in the MV panel, the probability of having the opportunity to move, λ, is computed as the share of year-on-year comparisons in which a municipality changed its multiplier. The tax base sensitivity θ is estimated based on the same data using a two-way fixed effects regression derived from a logit model of discrete firm location choice, whose linearization for two identical players yields the tax base of Section 2.1.

Non-tax benefits a_i are backed out from Equation (19) after identifying Players H and L with high-tax municipalities and tax havens, respectively, and taking their observed average tax rates in 2002 as the steady-state tax profile $(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$ of the affine MPE. To do so, I adopt the group definitions of Buettner and Poehnlein (2024). Every municipality with a 2002 multiplier below 200, the minimum value introduced in 2004 and announced in 2003, counts as a tax haven. I identify this group with Player L. A municipality counts as high-tax if its 2002 multiplier exceeds 350, the threshold of the top 15% in the seven-state sample of Buettner and Poehnlein (2024). I identify the subgroup of high-tax municipalities within a 15 km radius of a tax haven with Player H. Following Buettner and Poehnlein (2024), I assume that only municipalities within this radius compete with tax havens for firms.[13] The focus on tax havens and high-tax municipalities follows Buettner and Poehnlein (2024), who find no significant response to the minimum tax among neighbors of tax havens with 2002 multipliers between 200 and 350. Section 6 discusses how an extension to more than two players would affect the results.

The resulting non-tax benefit of Player H, $a_{\mathrm{H}} = -0.005$, is slightly negative. Economically, it can be interpreted as a disutility from hosting firms, for example from local negative externalities such as air pollution, for which Player H compensates by choosing a high tax rate.

Table 1: Calibrated parameter values

Parameter		Value	Source or target
Discount factor	δ	0.962	Real interest rate, 1995–2001
Probability of moving	λ	0.053	Multiplier changes, 2000–2007
Tax base sensitivity	θ	0.227	Regression (59), 2000–2007
Non-tax benefits	a_{H}	-0.005	$\tau_{\mathrm{H}}^* = 16.5\%$, high-tax municipalities, 2002
	a_{L}	0.350	$\tau_{\mathrm{L}}^* = 6.5\%$, tax havens, 2002
Minimum tax	$\underline{\tau}$	9.09%	Multiplier of 200, since 2004

Notes: One period is one year. δ is based on German data (10-year federal securities, CPI inflation); all other values are based on the *Realsteuervergleich* for MV. a_{H} and a_{L} are chosen such that the affine steady state $(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$ matches the mean tax rates in 2002 of the two groups in Table 4 in Appendix B.

[13] Assuming that the relevant competitors for a municipality are only its geographical neighbors is defensible for small municipalities, but less so for cities, which are also subject to inter-regional competition (Janeba and Osterloh 2013). In my MV panel, only one of the six high-tax municipalities in proximity of a tax haven is a city with more than 50,000 inhabitants (Neubrandenburg). All remaining municipalities have fewer than 2,500 inhabitants, four of them only a few hundred (all population figures are for 2002). Buettner and Poehnlein (2024, p. 8) argue that in their estimation, inter-regional effects cannot be identified but are likely absorbed by year-fixed effects.

5.2 Results

I start by analyzing the players' strategies, reported in Table 2. The dynamic slopes, $m = 0.370$, are smaller than the static ones, in line with the result $m \in (0, 1/2)$ from Proposition 3 and as illustrated in Figure 1. The intercepts of the strategies have the expected ordering, $0.36\% = k_{\mathrm{L}} < k_{\mathrm{H}} = 14.12\%$, and the parameter conditions of Proposition 3 hold: $k_{\mathrm{H}} + m = 0.51 \leq 1$, and the lower bound of the action space never binds for any $\tau_0 \leq k_{\mathrm{L}}/(1-m) \approx 0.57\%$. Because Player L has the higher non-tax benefit, its static best response has the smaller intercept, $(\theta - a_{\mathrm{L}})/2$. For the current calibration, this intercept is negative, leading to a negative tax rate in the static Nash equilibrium, a net subsidy.

Both incentive wedges are sizable, and Player L's is the larger one, 4.37 against 1.69 pp, although its tax rate is 10 percentage points below Player H's (Equation (26)). The dynamic premia, 5.16 and 6.95 pp, inherit this ordering but lie closer together: their gap is two thirds of the wedges' gap (Equation (27)).

Introducing the minimum tax. The effect of a minimum tax at 9.09%, above Player L's steady-state tax rate of $\tau_{\mathrm{L}}^* = 6.47\%$, is computed and decomposed based on the results from Proposition 5, and Corollary 6. The parameter conditions of these results hold: the minimum tax lies in $[\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}}) = [6.47\%, 23.2\%)$, the range of binding floors (Equation (28)) in which the high-tax player is not pinned to the minimum tax, and below the threshold $\underline{\tau}^\dagger = 9.84\%$ of Corollary 6, so that Player H lowers its rate.

The effect on Player L's rate is a sizable mechanical increase. In contrast, Player H *decreases* its tax rate by 0.38 percentage points. This is close to both the difference-in-differences (DiD) estimate of Buettner and Poehnlein (2024) and the raw DiD from the harmonized panel used for the calibration, both equal to -0.51 percentage points (MV is part of their sample). The estimated relative decrease in high-tax municipalities' tax rates is thus compatible, in principle, with strategic complementarity and upward sloping strategies. When comparing these figures, note that the model's prediction is a steady-state figure, while the empirically measured figures include transition dynamics.

The total change in Player H's tax rate can be decomposed into two effects. First, the complementarity effect is Player H's positive reaction to Player L's tax rate increase under strategic complementarity. Second, the wedge effect is the removal of Player H's incentive wedge. Tied to the floor, Player L no longer responds to Player H's rate (its strategy in the local MPE of Proposition 5). For the current calibration, the net effect is negative. Note that the wedge effect alone is more than four times as large as the model's net effect, and a minimum tax slightly above $\underline{\tau}^\dagger = 9.84\%$ would lead to a positive overall reaction by Player H *despite* the negative wedge effect.

Buettner and Poehnlein (2024, Fig. 2) show a steep upward trend of the havens' average multiplier from 2002 to 2004 from below 150 to almost 250, continuing on a slight upward trend for the rest of the observed years 2004–2008. The authors do not provide a numerical figure for this change. The harmonized MV panel I use for calibration features a similar pattern, with an average multiplier of 200 in 2003, when 7 of the 9 tax havens had already moved to or above the minimum, and of 257 in 2006–2007. The model does not account for this increase beyond the minimum tax: in the corner steady state of Proposition 5, the haven's rate is exactly $\underline{\tau}$, so

Table 2: Results of the calibrated model

		Player H (high-tax)		Player L (tax haven)	
		Dynamic	Static	Dynamic	Static
Strategies					
Intercept		14.12%	11.60%	0.36%	−6.16%
Slope		0.370	0.500	0.370	0.500
Without minimum tax					
Static Nash equilibrium	τ_i^{NE}	11.36%		−0.48%	
MPE steady state	τ_i^*	16.52%		6.47%	
Dynamic premium	$\tau_i^* - \tau_i^{\mathrm{NE}}$	5.16 pp		6.95 pp	
Incentive wedge	$\frac{\omega m}{2}(\tau_i^* + a_i)$	1.69 pp		4.37 pp	
With minimum tax $\underline{\tau} = 9.09\%$					
MPE steady state	$\hat{\tau}_i$	16.14%		9.09%	
Change	$\hat{\tau}_i - \tau_i^*$	−0.38 pp		+2.62 pp	
Complementarity effect		+1.31 pp			
Wedge effect		−1.69 pp			
Steady-state return		+20.84%		−3.42%	
Data: change in tax rates after the minimum tax					
Buettner and Poehnlein (2024)	DiD estimate	−0.51 pp			
MV panel	Raw change	−1.36 pp		+4.39 pp	
	Raw DiD	−0.51 pp			

Notes: Results are based on parameter values from Table 1. Dynamic: MPE strategy from Equation (21). Static: static best response from Equation (8). The steady state of the affine MPE is a calibration target. The static Nash equilibrium is given by Equation (9). The incentive wedge is the distance of τ_i^* above the static best response to τ_{-i}^* (Equation (18)). With the minimum tax, Player L sits at $\underline{\tau}$ and Player H at its static best response to it (Proposition 5). The complementarity effect is the shift of Player H's static best response as Player L's tax rate rises mechanically to $\underline{\tau}$. The wedge effect is the removal of Player H's incentive wedge, as Player L's strategy becomes locally flat at the floor. Steady-state return: change in the per-period return r^i (Equation (3)), $(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}})$ against $(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$. pp: percentage points. $\omega \equiv \delta\lambda D$. Data: Buettner and Poehnlein (2024): difference-in-differences estimate for high-tax municipalities within 15 km of a tax haven, −14.4 multiplier points, converted at their mean multiplier of 383. MV panel: mean over 2004–2007 against 2002 for the groups of Table 4; the raw DiD nets out the change of the high-tax municipalities without a tax haven within 15 km, converted at Player H's 2002 mean multiplier.

the model's change for Player L falls short of the observed one (Table 2). Section 6 discusses possible explanations.

Welfare and aggregate tax revenue. The change in steady-state returns is the percentage change of each player's per-period return r^i from the affine steady state $(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)$ to the steady state under the minimum tax $(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}})$. While sizable and positive for the high-tax player, it is negative for the tax haven. Buettner and Poehnlein (2024, p. 13) also interpret their evidence as suggesting that "the minimum tax-rate harms tax havens", but they trace this back to a loss in tax revenue. In the model, the finding is driven by the reduction in the tax base combined with the high non-tax benefits a_{L}. Haven tax revenues increase because, at the minimum tax introduced in 2004, the rise in the haven's tax rate outweighs the loss of tax base. The loss of tax base is a general result. By Corollary 6(i), Player H's rate changes by $(\underline{\tau} - \tau_{\mathrm{L}}^*)/2 - w_{\mathrm{H}}$, less than the increase $\underline{\tau} - \tau_{\mathrm{L}}^*$ in the haven's rate. Because of the smaller tax differential, the haven's tax base shrinks under every binding minimum tax, and for minimum taxes close to τ_{L}^*, at which its own rate hardly rises, so does its tax revenue. The model thus contains the mechanism behind the decline in havens' tax revenue documented by Buettner and Poehnlein (2024), but at the minimum tax introduced in 2004, the loss of tax base is outweighed by the rise in the haven's tax rate.

For Player H, the gain holds for every binding minimum tax (Corollary 7(i)). For Player L, the corollary guarantees the loss only for minimum taxes close to τ_{L}^* (Corollary 7(ii)); at the current calibration, it extends to the whole admissible range. More precisely, Player L is worse off not only for the steady state induced by the minimum tax $\underline{\tau} = 9.09\%$, but for all steady states induced by a minimum tax in the interval $[\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}}) = [6.47\%, 23.2\%)$, including those that make Player H raise its tax rate, that is, above the threshold of $\underline{\tau}^\dagger = 9.84\%$. To see this, note that Player L's flow return in the corner steady state of Proposition 5 is given by

$$f(\underline{\tau}) \equiv r^{\mathrm{L}}\big(\hat{\tau}_{\mathrm{H}}(\underline{\tau}), \underline{\tau}\big) = (\underline{\tau} + a_{\mathrm{L}}) \left(1 + \frac{\theta - a_{\mathrm{H}} - \underline{\tau}}{2\theta}\right),$$

where I plugged $\tau_{\mathrm{L}} = \underline{\tau}$ and $\tau_{\mathrm{H}} = \hat{\tau}_{\mathrm{H}}(\underline{\tau}) = \frac{1}{2}(\theta - a_{\mathrm{H}} + \underline{\tau})$, Player H's static best response, into (3). Because f is strictly concave,[14] it has a unique global maximizer and resulting returns at the maximum given by

$$f'(\underline{\tau}^{\max}) = 0 \quad \Longleftrightarrow \quad \underline{\tau}^{\max} = \frac{3\theta - a_{\mathrm{H}} - a_{\mathrm{L}}}{2}, \qquad f(\underline{\tau}^{\max}) = \frac{(3\theta - a_{\mathrm{H}} + a_{\mathrm{L}})^2}{8\theta}.$$

Since $f(\underline{\tau}^{\max})$ lies below the flow return at the affine steady state by 1.24% and is the maximal value over all $\underline{\tau}$, the low-tax player loses under every binding floor in $\big[\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}}\big)$.

A minimum tax of 9.09% raises aggregate tax revenue by 24.3%, although Player H lowers its tax rate and the first effect from the decomposition (34) is negative. In contrast, for a minimum tax equal to Player L's tax rate in the affine steady state, $\underline{\tau} = \tau_{\mathrm{L}}^*$, aggregate tax

[14] The first and second derivatives of f are given by

$$f'(\underline{\tau}) = 1 + \frac{\theta - a_{\mathrm{H}} - \underline{\tau}}{2\theta} - \frac{\underline{\tau} + a_{\mathrm{L}}}{2\theta} = \frac{3\theta - a_{\mathrm{H}} - a_{\mathrm{L}} - 2\underline{\tau}}{2\theta}, \qquad f''(\underline{\tau}) = -\frac{1}{\theta} < 0.$$

revenue decreases by 1.7%. By Corollary 8, the threshold above which the minimum tax raises aggregate tax revenue is unique; it is given by 6.64%, only 0.17 percentage points above the haven's rate in the affine steady state and comfortably below the minimum tax of 9.09% introduced in 2004.

6 Discussion

This section picks up discussion items from the main text. It discusses an observation from the harmonized MV panel that seems to be at odds with the model — the havens' observed rate increase *above* the minimum tax — and possible explanations for it, within the model and outside of it. Moreover, it discusses the insights the model offers for the global minimum tax and an alternative: tax clubs with a common club rate. It starts with two notes on the model, on why the affine MPE from Proposition 3 might be more than just an analytically convenient formulation and on its extension to more than two players.

The affine MPE and the finite-horizon limit. The MPE described in Section 3.2 is unique only among affine MPE with slopes in $(0, 1)$; non-affine MPE cannot be ruled out. This is the same gap that Maskin and Tirole (1987) acknowledge in their closely related analysis of quantity competition in a dynamic duopoly game. However, the authors describe another property of their affine MPE: it is the limit, as the horizon tends to infinity, of the unique backward-induction equilibrium of the truncated, finite-horizon version of the game. In the truncated game, the player who moves in the last period plays its static best response, its opponent takes this reaction into account in the penultimate period, and so on. The behavioral reading of this result is that players who reason backward from an increasingly distant but finite end of the game arrive at this MPE; it can be read as an account of how real-world actors may actually end up in this equilibrium.

Maskin and Tirole (1987, pp. 962–963) note that a game in which commitments are of random length, as in this model, gives rise to an equilibrium formally identical to that of their alternating-move game. Indeed, the slope equation (22) coincides with their Equation (20), with ω in place of the discount factor. A similar mapping of models is used in the proof of the global convergence property of the MPE from Proposition 3. Whether this formal identity is sufficient for the affine MPE to be the limit of the truncated game's perfect equilibrium is left for future work. If it is, this would be a compelling argument for studying this equilibrium that goes beyond its tractability.

More than two players. Municipalities, regional governments, and countries compete with many others, so that the two-player model is a simplification. Relaxing it would affect the quantitative results, because it would dilute the dynamic incentive, the core of the mechanism presented. If every player competes with every other for a common tax base, each player's tax rate enters every other player's decision problem only as one of $N - 1$ relevant state variables. The extension would thus reduce the incentive wedge and the dynamic premium. As the number of competitors increases, the dynamic incentive would tend to zero and only the static incentive would survive. So what is the number of relevant competitors? While I am not aware of any

estimate for the international or the German context, de Paula et al. (2025) estimate a network for tax competition between US states with only 49 ties among the 48 mainland states, so that a state's tax setting is influenced by about one other state on average. Most geographically adjacent states are irrelevant for each other's tax setting, and tax havens are less influential than other states. In this setting, the dilution of the dynamic incentive would be limited.

Beyond the quantitative results, an extension to more than two players would allow for players outside an agreement, who change both the effects of the agreement on its members and the members' alternative to joining. This applies to the minimum tax and to the common club rate discussed below alike, and is an important avenue for future research.

The havens' observed rate increases. In the MV panel as well as in the seven-state sample used by Buettner and Poehnlein (2024), the havens' average multiplier rises *above* the minimum of 200 and keeps increasing until 2007 (Section 5.2), so that the observed increase of their tax rate exceeds the model's (Table 2). Being a group average, the number masks that some tax havens *do* stay at the minimum in the period 2004–2007: Two out of the nine tax havens, namely Velgast and Süderholz, have a multiplier of 200 in the year 2007 (tax havens have a 2002 multiplier below 200, as defined in Buettner and Poehnlein (2024)); both increased their multiplier from 100 and 150, respectively, to 200 in 2003, the year in which the minimum multiplier was announced. When taking into account all municipalities with a multiplier equal to 200 in 2002 — those municipalities with a multiplier equal to the floor introduced in 2004 are constrained by the policy in the sense of Proposition 5, too — 21 out of 50 municipalities still have a multiplier of 200 in 2007. Hence, a non-negligible share of municipalities seems to stick with the minimum multiplier, in line with Proposition 5.

However, the majority does not. Their observed tax setting behavior is incompatible with the corner steady state described in Section 4. There are several possible explanations for this.

Within the model, the equilibrium of the game with a minimum tax need not be the local MPE of Proposition 5. The proposition bounds Player L's rates from above, so in the game with the action space $[\underline{\tau}, 1]$ a rate beyond this bound might be Player L's optimum, and its strategy T_{L} a different one.

Several other explanations are outside of the model framework presented above. First, for some municipalities the classification as a tax haven rests on a multiplier that was low only temporarily in 2002. A decrease in multiplier levels is observable for two havens between 2000 and 2002: Mölln and Kreien. For instance, Mölln was at 250 in 2000–2001, fell to 130 in 2002, and returned to 250 by 2006. For these municipalities, the increase could be a return to their normal multiplier, after it had been lowered in 2002 for reasons the model cannot capture.

Second, seven of the municipalities with multipliers not above 200 in 2002 increased their multiplier in the year of a merger with other municipalities. For example, Camin, the only one of the nine tax havens among them, had a multiplier of 100 until 2002, jumped to 200 in 2003 and above it in 2004, the year of its merger with Vellahn, a municipality whose multiplier had been 250 or more throughout the sample; from 2004 on, the panel records the multiplier of the merged municipality.

Third, when choosing its tax rate, a haven might base its decision on the mobility of the most mobile firm within its jurisdiction. Imagine the extreme case where two firms locate in a tax

haven's municipality: a perfectly mobile letterbox firm, and a highly non-mobile manufacturing firm with its only production site located in the haven's jurisdiction. Once the perfectly mobile firm is gone — potentially because of the dynamics described in this article — the haven might drastically increase the tax rate paid by the immobile manufacturer. Because relocation is costless for every firm ($C = 0$ in Equation (2)), the model cannot capture such dynamics.

Fourth, the previous competition in tax rates might now be continued with different instruments. In a German setting, Hauptmeier et al. (2012) find that municipalities respond to tax cuts of their neighbors with *both* tax cuts and higher spending on public inputs. However, increasing public inputs to compensate for higher business tax rates would explain why former tax havens could afford to choose higher tax rates, not why it would be optimal for them to do so.

Finally, a general trend is observable in the data: from 2004 on, low-tax municipalities in MV raised their multipliers. A potential explanation is a sequence of cuts in the transfers from the state of MV to its municipalities, which the state government justified with its own falling revenue (Landtag Mecklenburg-Vorpommern 2004b, p. 1). Transfers for infrastructure projects were reduced, and the guaranteed minimum funding of the municipalities was lowered for 2004 and abolished for 2005 (Landtag Mecklenburg-Vorpommern 2004b, p. 3). Taken together, the transfers planned for 2005 are 154.9 million euros lower than the transfers of 2002 (1,354.9 million euros), a reduction by 11.4%. Relative to the total receipts of the municipalities in MV in 2002 (3,353.3 million euros), this is a reduction by 4.6%.[15] These measures plausibly influenced municipal tax setting behavior. Establishing causality is beyond the scope of this paper.

Lessons for the global minimum tax. In their analysis of the Pareto-improving potential of the global minimum tax, Hebous and Keen (2023, p. 9) conclude that "[e]mpirical results which strongly suggest statutory tax rates to be strategic complements, and the likelihood in a context of profit shifting that each country benefits from a tax increase by the other, point to the most plausible outcome being a welfare gain for the low tax country". In their static setting, a minimum tax just above the low-tax player's Nash rate affects the low-tax player's welfare through two channels. First, its own rate rises. Since it sets its best response at the Nash equilibrium, this has no first-order effect on its welfare, because the gain from taxing its base at a higher rate and the loss from the narrower tax differential cancel. Second, the high-tax player responds to the higher rate, and with strategic complements, it raises its own rate, from which the low-tax player gains; on net, it gains. Only for higher minimum taxes does the loss from the own-rate increase become first-order. In their calibration, the minimum tax most preferred by the low-tax player still lies well above its Nash rate: for a slope of the high-tax player's best response of 0.4, around the midpoint of the empirical estimates they cite, it is 16.7% against a benchmark Nash rate of 12.5% (Hebous and Keen 2023, pp. 7–8).

[15] Own computation. Transfers in 2002 and total receipts are from Landtag Mecklenburg-Vorpommern (2004c, p. 6). Transfers planned for 2005: equalization payments of 1,121.9 million euros, a special payment of 43.0 million euros, and transfers for infrastructure projects of 35.1 million euros, as in the committee version of the amendment to the fiscal equalization law (Landtag Mecklenburg-Vorpommern 2004b, pp. 3, 7 and 24), passed in 2004 (Landtag Mecklenburg-Vorpommern 2004a).

However, their result hinges on the high-tax player's *positive response* to an increase in the low-tax player's tax rate, induced by a minimum tax. As I show, this reaction need not be positive *despite* strategic complementarity in tax rates because of the dynamic incentive and the elimination of the incentive wedge that can be caused by a minimum tax (Corollary 6). Even if the reaction is positive ($\underline{\tau} > \underline{\tau}^{\dagger}$ from Corollary 6), the low-tax player might lose, and in the calibration it does. The intuition is that at the affine steady state, the low-tax player already sets its rate above its static best response: it accepts a static marginal loss — unlike in the static case, the marginal base loss exceeds the gain from taxing the inframarginal base at a higher rate — in exchange for the discounted gain from pulling the opponent's rate up (Equation (20)). Raising its rate further, at the opponent's unchanged rate, therefore lowers its flow return already at the margin, and does so for every further increase.

As shown in Corollary 7, the low-tax player's steady-state flow return decreases for minimum taxes close to its affine steady-state rate τ_{L}^{*}. For these minimum taxes, its tax revenue decreases as well, and its tax base shrinks under every binding minimum tax (Section 5.2). If transferable to the international setting, this contradicts the result of Hebous and Keen (2023) that a floor marginally above the low-tax player's rate makes it better off.[16] For higher minimum taxes, the comparison depends on the objective. At the calibration, the low-tax player's flow return decreases for every minimum tax in the admissible range (Section 5.2), but at the minimum tax introduced in 2004, its tax revenue rises, and its loss stems from the non-tax benefits of the firms it loses.

As noted in the introduction, Hebous and Keen (2023, p. 3) anticipate these effects for the case in which one player moves first; the dynamic model recovers them without assigning the roles of leader and follower.

A single market levy as a common club rate. While this article casts doubt on the Pareto-improving potential of a global minimum tax for current tax havens — their participation constraint for such an agreement might thus not be met — it does not argue against the benefits of preventing race-to-the-bottom tax competition dynamics. To sketch an avenue for future research, I briefly discuss an alternative to a minimum tax rate that leaves the dynamic incentive and the incentive wedge of both players intact: a tax club whose members levy a common club rate on top of their domestic tax rates. It is inspired by the single market levy proposal by García Antón and Lejour (2023, p. 7), who recommend a low common rate, for example below 3%, levied on top of the member states' domestic tax rates to fund the EU budget. Given that tax legislation at the EU level must be adopted unanimously, making all member states weakly better off is of even greater importance in this context.

In the model framework introduced in Section 2, denote by $\tilde{\tau}_i$ Player i's domestic tax rate, that is, the rate determining the player's national tax revenue. Consider a tax club of both players that levies a common club rate $\phi \geq 0$ (e.g., the rate of the single market levy) on top

[16] The welfare concept of Hebous and Keen (2023) is the private after-tax income of a country's residents plus its tax revenue, the latter with a weight above one; it contains no non-tax benefits. Private income cannot rise under a binding minimum tax, because the minimum tax raises the rate that firms pay in the low-tax player's jurisdiction and reduces their number. For minimum taxes close to τ_{L}^{*}, tax revenue falls as well, so the low-tax player's welfare in the sense of Hebous and Keen (2023) falls, too.

of each player's domestic tax rate, and define $\tau_i \equiv \tilde{\tau}_i + \phi$ as the overall tax rate firms face. It is the rate that matters for firms' profits (Equation (1)) and thus for their location decisions. The location decisions determine the players' tax bases, denoted $x_i(\tau_\mathrm{H}, \tau_\mathrm{L})$, as in Equation (3). Assume that the revenues $\phi\, x_i(\tau_\mathrm{H}, \tau_\mathrm{L})$ of each player are used to finance a public good (e.g., one provided through the EU budget), where the constant $b_i \geq 0$ is the discounted sum of Player i's benefits from a marginal increase in the public good. Instantaneous returns are then given by

$$r^i(\tau_\mathrm{H}, \tau_\mathrm{L}; \phi) = (\tau_i - \phi + a_i) x_i(\tau_\mathrm{H}, \tau_\mathrm{L}) + b_i\, \phi(x_\mathrm{H} + x_\mathrm{L}), \tag{35}$$

where $x_\mathrm{H} + x_\mathrm{L} = 2$ is the constant mass of firms the two players compete for. The benefit from the public good thus depends only on the club rate ϕ, not on which player collects the revenue. Note that, if each player kept the revenue $\phi\, x_i$, returns would be those of the original game in the overall rates τ_i. Timing, information structure, and action space are those of Section 2.3, where players choose domestic rates $\tilde{\tau}_i \in \mathcal{A}(\tau_0)$.

Note that the tax bases depend on the two tax rates only through their difference, which the club rate leaves unchanged: $x_i(\tau_\mathrm{H}, \tau_\mathrm{L}) = x_i(\tilde{\tau}_\mathrm{H}, \tilde{\tau}_\mathrm{L})$. Together with $x_\mathrm{H} + x_\mathrm{L} = 2$, the instantaneous returns from Equation (35) can thus be written as

$$r^i(\tau_\mathrm{H}, \tau_\mathrm{L}; \phi) = (\tilde{\tau}_i + a_i)\, x_i(\tilde{\tau}_\mathrm{H}, \tilde{\tau}_\mathrm{L}) + 2\, b_i\, \phi = r^i(\tilde{\tau}_\mathrm{H}, \tilde{\tau}_\mathrm{L}) + 2\, b_i\, \phi, \tag{36}$$

where $r^i(\tilde{\tau}_\mathrm{H}, \tilde{\tau}_\mathrm{L})$ are the instantaneous returns of the benchmark model in Equation (3). In domestic rates, the game is therefore the benchmark game with a constant added to each player's instantaneous return. A constant does not affect the players' choices, so the game in domestic rates continues to have the affine MPE of Proposition 3, with the same strategies, for every ϕ.

Domestic tax rates in steady state are thus $(\tau_\mathrm{H}^*, \tau_\mathrm{L}^*)$ from Proposition 3, unaffected by the club rate. The same holds for each player's wedge w_i in Equation (25) and for its dynamic premium Δ_i in Corollary 4, because the static best responses and the static Nash equilibrium in domestic rates are those of the benchmark game, too. The overall tax rates firms face are $(\tau_\mathrm{H}^* + \phi, \tau_\mathrm{L}^* + \phi)$, a homogeneous shift by exactly ϕ as compared to the setting without the club rate. Because the shift is homogeneous, tax differentials and therefore tax bases are unchanged. Taken together, tax revenues are unchanged. If, in addition, benefits from the public good are strictly positive for at least one of the players, $b_\mathrm{H} > 0$ or $b_\mathrm{L} > 0$, then a positive club rate $\phi > 0$ is a Pareto improvement for the two players.

The main take-away of this discussion is the contrast with the minimum tax analyzed in Section 4: an instrument that bounds tax rates may remove the wedge, while an instrument that shifts rates does not. This is a valuable lesson for the design of international tax agreements and should be the starting point for further inquiry. In the EU context, it suggests that the single market levy is more likely to find unanimous support than a minimum tax above 15%.

7 Conclusion

This article provides a dynamic model of tax competition to analyze the effect of introducing a minimum tax rate on a low-tax player, bound by the floor, and a high-tax opponent. It

shows that a minimum tax can lead to a decrease in the high-tax player's rate despite strategic complementarity and upward sloping equilibrium strategies. The mechanism is that the floor does what the high-tax player's own high rate did before: it keeps the low-tax player's rate from falling. If the low-tax player no longer responds to the high-tax player's rate at the floor, the high-tax player's incentive wedge vanishes. This finding reconciles the decrease in the tax rates of high-tax German municipalities after the introduction of a minimum business tax in 2004, documented by Buettner and Poehnlein (2024), with the notion of strategic complementarity. The model does not account for the havens' observed tax rate increases beyond the minimum tax. The article further shows that the low-tax player loses in terms of both steady-state returns and tax revenue from minimum taxes just above its tax rate without the floor, because the high-tax player's lower rate takes tax base away from it. Aggregate tax revenue can decline after the introduction of a minimum tax if the floor is below a threshold whose existence depends on the model parameters. At the calibration, the low-tax player's welfare loss obtains over the whole range of admissible minimum taxes, and aggregate tax revenue declines only for floors just above the low-tax player's rate without one. A central government should thus introduce only minimum tax rates that are sufficiently high, if it wants to prevent aggregate tax revenue from falling. If transferable to the international setting, the low-tax player's loss casts doubt on whether tax havens' participation constraints are met with respect to the global minimum tax, which is currently being rolled out.

The article points toward several avenues for future research. First, two questions about the equilibria are open: the existence and the characterization of an MPE with a minimum tax when the players' tax rates are not bounded from above, and whether the affine MPE is the limit of the backward-induction equilibrium of the truncated, finite-horizon game, as Maskin and Tirole (1987) show for their alternating-move game. Second, extending the model to more than two players is a natural next step, since municipalities and national governments compete with many others. Third, firms relocate at no cost in the model ($C = 0$ in Equation (2)), so that the tax base adjusts to the tax differential at once. With positive relocation costs, firms would look ahead and the tax base would adjust gradually, as in the model of Paulus et al. (2026), at the price of an additional state variable. Moreover, the overall tax base is fixed. Making it elastic, for instance through firm entry and exit, would be a valuable addition. Fourth, a minimum tax bounds tax rates but not other instruments. When the global minimum tax was proposed, for instance, Switzerland drew up plans for subsidies to offset the higher tax rates (Jones 2021). Including a second instrument such as subsidies to firms, as in Janeba and Schjelderup (2023), would capture how a player at the floor may continue to compete. Fifth, and most promising from a policy perspective, the effects of a minimum tax on players' participation constraints could be compared to those of a tax club with a common club rate, such as the single market levy proposed for the European Union (García Antón and Lejour 2023). As discussed in Section 6, a common club rate shifts the tax rates firms face and leaves the wedges of both players intact, whereas a minimum tax bounds tax rates and may remove the incentive wedge of the high-tax player. In the model, a positive club rate makes no player worse off, which is what the unanimity requirement for tax legislation in the EU calls for. Whether this comparison survives once players outside the agreement and a distortive tax on firms are considered, and whether either agreement is self-enforcing, remains to be studied.

References

Agrawal, David R. and David E. Wildasin (2020). "Technology and tax systems". In: *Journal of Public Economics* 185, p. 104082.

Buettner, Thiess and Maximilian Poehnlein (2024). "Tax competition effects of a minimum tax rate: Empirical evidence from German municipalities". In: *Journal of Public Economics* 236, 105148.

Chiari, Alessandro (2024). "Revenue losses from corporate tax avoidance: Estimations from the UNUWIDER Government Revenue Dataset". In: *Review of Development Economics* 28.2, pp. 600–629.

de Paula, Auréo, Imran Rasul, and Pedro C. L. Souza (2025). "Identifying Network Ties from Panel Data: Theory and an Application to Tax Competition". In: *The Review of Economic Studies* 92.4, pp. 2691–2729.

Deutsche Bundesbank (2026). *Aus der Zinsstruktur abgeleitete Renditen für Bundeswertpapiere mit jährlichen Kuponzahlungen, Restlaufzeit 10 Jahre, Monatswerte*. Dataset. URL: https://api.statistiken.bundesbank.de/rest/download/BBSIS/M.I.ZAR.ZI.EUR.S1311.B.A604.R10XX.R.A.A._Z._Z.A?format=csv&lang=de (visited on 09/02/2026).

European Trade Union Confederation (2025). *ETUC resolution on a 3-year Taxation Action Plan (2025–2027)*. Resolution. Brussels: European Trade Union Confederation.

Eurostat (2026). *Gross domestic product (GDP) and main components (output, expenditure and income) - annual data (nama_10_gdp): Price index (implicit deflator), percentage change on previous period, national currency, Germany*. Dataset. URL: https://ec.europa.eu/eurostat/api/dissemination/sdmx/3.0/data/dataflow/ESTAT/nama_10_gdp/1.0/A.PD_PCH_PRE_NAC.B1GQ.DE?format=csvdata&formatVersion=2.0&labels=both (visited on 09/24/2026).

García Antón, Ricardo and Arjan Lejour (2023). *The Justification, Design and Revenues of a Single Market Levy*. Report. Socialists and Democrats in the European Parliament.

Garcia-Bernardo, Javier and Petr Janský (2024). "Profit shifting of multinational corporations worldwide". In: *World Development* 177, p. 106527.

Greene, William H. (2003). *Econometric Analysis*. 5th ed. Upper Saddle River, New Jersey: Pearson Education.

Greens/EFA Group in the European Parliament (July 9, 2021). *G20 finance ministers meeting – quote from Sven Giegold, MEP*. Press Release. Brussels: Greens/EFA Group in the European Parliament.

Hauptmeier, Sebastian, Ferdinand Mittermaier, and Johannes Rincke (2012). "Fiscal competition over taxes and public inputs". In: *Regional Science and Urban Economics* 42.3, pp. 407–419.

Hebous, Shafik and Michael Keen (2023). "Pareto-improving minimum corporate taxation". In: *Journal of Public Economics* 225, 104952.

Heimberger, Philipp (2021). "Corporate tax competition: A meta-analysis". In: *European Journal of Political Economy* 69, p. 102002.

Janeba, Eckhard and Steffen Osterloh (2013). "Tax and the city — A theory of local tax competition". In: *Journal of Public Economics* 106, pp. 89–100.

Janeba, Eckhard and Wolfgang Peters (1999). "Tax Evasion, Tax Competition and the Gains From Nondiscrimination: the Case of Interest Taxation in Europe". In: *The Economic Journal* 109.452, pp. 93–101.

Janeba, Eckhard and Guttorm Schjelderup (2023). "The global minimum tax raises more revenues than you think, or much less". In: *Journal of International Economics* 145, p. 103837.

Johannesen, Niels (2022). "The global minimum tax". In: *Journal of Public Economics* 212, p. 104709.

Jones, Sam (June 10, 2021). "Switzerland plans subsidies to offset G7 corporate tax plan". In: *Financial Times*.

Kanbur, Ravi and Michael Keen (1993). "Jeux Sans Frontières: Tax Competition and Tax Coordination When Countries Differ in Size". In: *The American Economic Review* 83.4, pp. 877–892.

Keen, Michael and Kai A. Konrad (2013). "The theory of international tax competition and coordination". In: *Handbook of Public Economics*. Vol. 5. Elsevier, pp. 257–328.

Kiss, Áron (2012). "Minimum taxes and repeated tax competition". In: *International Tax and Public Finance* 19.5, pp. 641–649.

Konrad, Kai A. (2009). "Non-binding minimum taxes may foster tax competition". In: *Economics Letters* 102.2, pp. 109–111.

KPMG (May 2026). *Clarity on Swiss Taxes: Finding a Balance*. Report. Zurich: KPMG.

Landesamt für innere Verwaltung Mecklenburg-Vorpommern, Statistisches Amt (2026). *Realsteuervergleich, Berichtsjahre 2000–2007 (Statistische Berichte L273)*. Dataset. URL: https://www.laiv-mv.de/Statistik/Zahlen-und-Fakten/Gesellschaft-&-Staat/Steuern/ (visited on 03/02/2026).

Landesamt für Statistik Niedersachsen (2026). *Realsteuervergleich in Niedersachsen: Methodische Hinweise*. URL: https://www.statistik.niedersachsen.de/startseite/themen/finanzen_steuern_personal/steuern_in_niedersachsen/realsteuervergleich_in_niedersachsen/realsteuervergleich-niedersachsen-methodische-hinweise-197959.html (visited on 09/01/2026).

Landtag Mecklenburg-Vorpommern (2004a). *32. Sitzung, 4. Wahlperiode*. Plenarprotokoll 4/32.

— (2004b). *Beschlussempfehlung und Bericht des Innenausschusses (2. Ausschuss) zu dem Gesetzesentwurf der Landesregierung - Drucksache 4/704: Entwurf eines Fünften Gesetzes zur Änderung des Finanzausgleichsgesetzes*. Drucksache 1038.

— (2004c). *Unterrichtung durch die Landesregierung: Bericht zur kommunalen Haushaltslage 2002/2003*. Drucksache 1381.

Leibrecht, Markus and Claudia Hochgatterer (2012). "Tax Competition as a Cause of Falling Corporate Income Tax Rates: A Survey of Empirical Literature". In: *Journal of Economic Surveys* 26.4, pp. 616–648.

Maskin, Eric and Jean Tirole (1987). "A theory of dynamic oligopoly, III". In: *European Economic Review* 31.4, pp. 947–968.

— (1988). "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs". In: *Econometrica* 56.3, p. 549.

Parchet, Raphaël (2019). "Are Local Tax Rates Strategic Complements or Strategic Substitutes?" In: *American Economic Journal: Economic Policy* 11.2, pp. 189–224.

Paulus, Nora, Weihua Ruan, and Benteng Zou (2026). *On the Desirability of the Global Minimum Tax - A Dynamic View*. Discussion Paper. University of Luxembourg – Department of Economics and Management.

Statistisches Bundesamt (2026). *Verbraucherpreisindex für Deutschland, Jahresdurchschnitte (GENESIS-Tabelle 61111-0001)*. Dataset. URL: https://www-genesis.destatis.de/datenbank/online/statistic/61111/table/61111-0001 (visited on 09/02/2026).

Thunecke, Georg U., Valeria Merlo, Andreas Schanbacher, and Georg Wamser (2026). "Identifying tax-setting responses from local fiscal policy programs". In: *Journal of Public Economics* 258, p. 105643.

Wang, You-Qiang (1999). "Commodity Taxes Under Fiscal Competition: Stackelberg Equilibrium and Optimality". In: *The American Economic Review* 89.4, pp. 974–981.

A Proofs and Derivations

A.1 Dwell-Factor Recursions for the Values of Moving and Waiting

Two derived valuations recur in the analysis of the equilibrium strategies. The first is the value of *moving* against opponent rate τ_{L}, which by the optimality condition (14) is

$$V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) = \max_{\tau'_{\mathrm{H}}} V_{\mathrm{H}}(\tau'_{\mathrm{H}},\tau_{\mathrm{L}}).$$

The second is the value of *watching* Player L move while Player H's own rate is held at τ_{H}, $W_{\mathrm{H}}(\tau_{\mathrm{H}}) = V_{\mathrm{H}}(\tau_{\mathrm{H}},T_{\mathrm{L}}(\tau_{\mathrm{H}}))$, defined in Equation (16) for an arbitrary MPE. I derive recursive formulations for both that are independent of the value at other states. Throughout, I use the dwell factor $D_i = D(\lambda_i)$ of Equation (11). Once Player H has best-responded to a fixed τ_{L}, the configuration is left unchanged by any further move of Player H — with τ_{L} fixed, the best response is again $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ — and by periods in which nobody moves, and is disturbed only when Player L moves; Player H's flow is therefore weighted by the discounted dwell $D_{\mathrm{L}} = \sum_{t\geq 0}[\delta(1-\lambda_{\mathrm{L}})]^t$ and its continuation by the discounted move weight $\delta\lambda_{\mathrm{L}}D_{\mathrm{L}} = \sum_{s\geq 1}\delta^s\lambda_{\mathrm{L}}(1-\lambda_{\mathrm{L}})^{s-1}$.

I first collect the $\lambda_\emptyset$ self-reference in the value recursion (13). Writing the mover branch as $\max_{\tau'_{\mathrm{H}}} V_{\mathrm{H}}(\tau'_{\mathrm{H}},\tau_{\mathrm{L}}) = V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}})$ and the opponent branch as $W_{\mathrm{H}}(\tau_{\mathrm{H}})$, the value recursion reads

$$V_{\mathrm{H}}(\tau_{\mathrm{H}},\tau_{\mathrm{L}}) = r^{\mathrm{H}}(\tau_{\mathrm{H}},\tau_{\mathrm{L}}) + \delta\left[\lambda_{\mathrm{H}}V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) + \lambda_{\mathrm{L}}W_{\mathrm{H}}(\tau_{\mathrm{H}}) + \lambda_\emptyset V_{\mathrm{H}}(\tau_{\mathrm{H}},\tau_{\mathrm{L}})\right],$$

and moving the $\lambda_\emptyset$ term to the left gives

$$(1-\delta\lambda_\emptyset)V_{\mathrm{H}}(\tau_{\mathrm{H}},\tau_{\mathrm{L}}) = r^{\mathrm{H}}(\tau_{\mathrm{H}},\tau_{\mathrm{L}}) + \delta\left[\lambda_{\mathrm{H}}V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) + \lambda_{\mathrm{L}}W_{\mathrm{H}}(\tau_{\mathrm{H}})\right]. \tag{37}$$

The value of watching the opponent move. Evaluate Equation (37) at the post-move state $(\tau_{\mathrm{H}},T_{\mathrm{L}}(\tau_{\mathrm{H}}))$. Its left-hand side is $(1-\delta\lambda_\emptyset)W_{\mathrm{H}}(\tau_{\mathrm{H}})$ by the definition (16); on the right, the mover branch becomes $V_{\mathrm{H}}(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})),T_{\mathrm{L}}(\tau_{\mathrm{H}}))$ and the opponent branch stays $W_{\mathrm{H}}(\tau_{\mathrm{H}})$, because Player L's move does not touch Player H's rate τ_{H}:

$$(1-\delta\lambda_\emptyset)W_{\mathrm{H}}(\tau_{\mathrm{H}}) = r^{\mathrm{H}}(\tau_{\mathrm{H}},T_{\mathrm{L}}(\tau_{\mathrm{H}})) + \delta\left[\lambda_{\mathrm{H}}V_{\mathrm{H}}(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})),T_{\mathrm{L}}(\tau_{\mathrm{H}})) + \lambda_{\mathrm{L}}W_{\mathrm{H}}(\tau_{\mathrm{H}})\right].$$

Collecting $W_{\mathrm{H}}(\tau_{\mathrm{H}})$ on the left and using $1-\delta\lambda_\emptyset-\delta\lambda_{\mathrm{L}} = 1-\delta(1-\lambda_{\mathrm{H}}) = 1/D_{\mathrm{H}}$,

$$W_{\mathrm{H}}(\tau_{\mathrm{H}}) = D_{\mathrm{H}}\, r^{\mathrm{H}}(\tau_{\mathrm{H}},T_{\mathrm{L}}(\tau_{\mathrm{H}})) + \delta\lambda_{\mathrm{H}}D_{\mathrm{H}}\, V_{\mathrm{H}}(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})),T_{\mathrm{L}}(\tau_{\mathrm{H}})). \tag{38}$$

The value of moving. Evaluate Equation (37) at the post-move state $(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}})$. Its left-hand side is $(1-\delta\lambda_\emptyset)V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}})$; on the right, the mover branch $V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}})$ is unchanged, since it depends on τ_{L} alone, and the opponent branch becomes $W_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}))$:

$$(1-\delta\lambda_\emptyset)V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) = r^{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) + \delta\left[\lambda_{\mathrm{H}}V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) + \lambda_{\mathrm{L}}W_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}))\right].$$

Collecting $V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}})$ on the left and using $1-\delta\lambda_\emptyset-\delta\lambda_{\mathrm{H}} = 1-\delta(1-\lambda_{\mathrm{L}}) = 1/D_{\mathrm{L}}$,

$$V_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) = D_{\mathrm{L}}\, r^{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}),\tau_{\mathrm{L}}) + \delta\lambda_{\mathrm{L}}D_{\mathrm{L}}\, W_{\mathrm{H}}(T_{\mathrm{H}}(\tau_{\mathrm{L}})). \tag{39}$$

Finally, substituting Equation (39) into Equation (38) — eliminating the value of moving — yields a "loop" equation that closes W_{H} on itself:

$$W_{\mathrm{H}}(\tau_{\mathrm{H}}) = D_{\mathrm{H}}\, r^{\mathrm{H}}(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})) + \delta\lambda_{\mathrm{H}} D_{\mathrm{H}} \left[D_{\mathrm{L}}\, r^{\mathrm{H}}(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})) + \delta\lambda_{\mathrm{L}} D_{\mathrm{L}}\, W_{\mathrm{H}}(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}}))) \right]. \tag{40}$$

The corresponding objects for Player L follow by relabeling $\mathrm{H} \leftrightarrow \mathrm{L}$ and $\lambda_{\mathrm{H}} \leftrightarrow \lambda_{\mathrm{L}}$; Player L's first-order condition is derived, by that symmetry, in Appendix A.3 below.

A.2 Proof of Lemma: Upward Sloping Equilibrium Strategies

Lemma 1: Let $(T_{\mathrm{H}}, T_{\mathrm{L}})$ be an MPE. Then $T_{\mathrm{H}}(\tau_{\mathrm{L}}) \geq T_{\mathrm{H}}(\tau'_{\mathrm{L}})$ for all $\tau_{\mathrm{L}}, \tau'_{\mathrm{L}} \in \mathcal{A}(\tau_0)$ with $\tau_{\mathrm{L}} > \tau'_{\mathrm{L}}$; the argument for Player L is symmetric.

Proof. Assume, toward a contradiction, that $\tau_{\mathrm{L}} > \tau'_{\mathrm{L}}$ but $T_{\mathrm{H}}(\tau_{\mathrm{L}}) < T_{\mathrm{H}}(\tau'_{\mathrm{L}})$. By the optimality condition (14), $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ maximizes $\tau'_{\mathrm{H}} \mapsto V_{\mathrm{H}}(\tau'_{\mathrm{H}}, \tau_{\mathrm{L}})$ over $\mathcal{A}(\tau_0)$ and $T_{\mathrm{H}}(\tau'_{\mathrm{L}})$ maximizes $\tau'_{\mathrm{H}} \mapsto V_{\mathrm{H}}(\tau'_{\mathrm{H}}, \tau'_{\mathrm{L}})$; comparing each maximizer against the other player's rate gives

$$V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) \geq V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau_{\mathrm{L}}\right), \tag{41}$$

$$V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau'_{\mathrm{L}}\right) \geq V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau'_{\mathrm{L}}\right). \tag{42}$$

Adding (41) and (42) and rearranging,

$$\mathcal{M} := V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) - V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau_{\mathrm{L}}\right) + V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau'_{\mathrm{L}}\right) - V_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau'_{\mathrm{L}}\right) \geq 0. \tag{43}$$

Substitute the collected value equation (37) for each of the four terms. Equation (37) has already absorbed the no-move ($\lambda_\emptyset$) self-reference into the factor $1-\delta\lambda_\emptyset$ on its left-hand side, so multiplying $\mathcal{M}$ by $1-\delta\lambda_\emptyset$ leaves, on the right, only the flow r^{H} and the two *move* continuations $\delta\lambda_{\mathrm{H}} V_{\mathrm{H}}(T_{\mathrm{H}}(\cdot), \cdot)$ and $\delta\lambda_{\mathrm{L}} W_{\mathrm{H}}(\cdot)$; the no-move continuation does not reappear as a separate term. These two move continuations cancel in the mixed difference: the λ_{H} term depends only on the second argument, so it enters $\mathcal{M}$ once with each sign at τ_{L} and once with each sign at τ'_{L}; the λ_{L} term depends only on the first argument, so it enters once with each sign at $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ and once with each sign at $T_{\mathrm{H}}(\tau'_{\mathrm{L}})$. What remains is

$$(1-\delta\lambda_\emptyset)\,\mathcal{M} = r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) - r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau_{\mathrm{L}}\right) + r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau'_{\mathrm{L}}\right) - r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau'_{\mathrm{L}}\right).$$

Since $1-\delta\lambda_\emptyset > 0$, the inequality (43) is equivalent to

$$r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) - r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau_{\mathrm{L}}\right) + r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau'_{\mathrm{L}}), \tau'_{\mathrm{L}}\right) - r^{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau'_{\mathrm{L}}\right) \geq 0,$$

which is equivalent to

$$\begin{aligned}
&\int_{T_{\mathrm{H}}(\tau'_{\mathrm{L}})}^{T_{\mathrm{H}}(\tau_{\mathrm{L}})} \frac{\partial r^{\mathrm{H}}(x, \tau_{\mathrm{L}})}{\partial x}\, dx + \int_{T_{\mathrm{H}}(\tau_{\mathrm{L}})}^{T_{\mathrm{H}}(\tau'_{\mathrm{L}})} \frac{\partial r^{\mathrm{H}}(x, \tau'_{\mathrm{L}})}{\partial x}\, dx \geq 0 \\
\iff \quad &\int_{T_{\mathrm{H}}(\tau_{\mathrm{L}})}^{T_{\mathrm{H}}(\tau'_{\mathrm{L}})} \left[\frac{\partial r^{\mathrm{H}}(x, \tau'_{\mathrm{L}})}{\partial x} - \frac{\partial r^{\mathrm{H}}(x, \tau_{\mathrm{L}})}{\partial x} \right] dx \geq 0 \\
\iff \quad &\int_{\tau'_{\mathrm{L}}}^{\tau_{\mathrm{L}}} \int_{T_{\mathrm{H}}(\tau_{\mathrm{L}})}^{T_{\mathrm{H}}(\tau'_{\mathrm{L}})} \frac{\partial^2 r^{\mathrm{H}}(x, y)}{\partial x\, \partial y}\, dx\, dy \leq 0.
\end{aligned}$$

But by assumption (game of strategic complements), $\partial^2 r^{\mathrm{H}}(x,y)/\partial x\,\partial y > 0$, while $\tau_{\mathrm{L}} > \tau'_{\mathrm{L}}$ and, by the contradiction hypothesis, $T_{\mathrm{H}}(\tau_{\mathrm{L}}) < T_{\mathrm{H}}(\tau'_{\mathrm{L}})$, so the double integral is strictly positive — a contradiction. Hence T_{H} is upward sloping. The argument compares maximizers only and uses neither differentiability of T_{H} or V_{H} nor the shape of the action set; it holds for any selection from the argmax in Equation (14). □

A.3 Difference–Differential Equations that characterize MPE

Next, characterize the policy functions T_{H} and T_{L} without recurring to the valuation functions V_i or W_i.

By the optimality condition (14), $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ maximizes $\tau'_{\mathrm{H}} \mapsto V_{\mathrm{H}}(\tau'_{\mathrm{H}}, \tau_{\mathrm{L}})$, so at the interior optimum

$$\frac{\partial V_{\mathrm{H}}}{\partial \tau_{\mathrm{H}}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) = 0. \tag{44}$$

To express this in terms of the primitives, differentiate the value recursion (13) with respect to τ_{H}. The mover branch $\lambda_{\mathrm{H}} \max_{\tau'_{\mathrm{H}}} V_{\mathrm{H}}(\tau'_{\mathrm{H}}, \tau_{\mathrm{L}})$ does not depend on τ_{H} and drops; the opponent branch is $\lambda_{\mathrm{L}} V_{\mathrm{H}}(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})) = \lambda_{\mathrm{L}} W_{\mathrm{H}}(\tau_{\mathrm{H}})$; hence

$$\frac{\partial V_{\mathrm{H}}}{\partial \tau_{\mathrm{H}}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) + \delta\left[\lambda_{\mathrm{L}} W'_{\mathrm{H}}(\tau_{\mathrm{H}}) + \lambda_{\emptyset}\frac{\partial V_{\mathrm{H}}}{\partial \tau_{\mathrm{H}}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}})\right],$$

and collecting the $\lambda_{\emptyset}$ term,

$$(1 - \delta\lambda_{\emptyset})\frac{\partial V_{\mathrm{H}}}{\partial \tau_{\mathrm{H}}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) + \delta\lambda_{\mathrm{L}} W'_{\mathrm{H}}(\tau_{\mathrm{H}}), \tag{15}$$

which is Equation (15). Evaluating (15) at $(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}})$ and imposing the first-order condition (44), the left-hand side vanishes; since $1 - \delta\lambda_{\emptyset} > 0$,

$$r^{\mathrm{H}}_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) + \delta\lambda_{\mathrm{L}}\, W'_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}})\right) = 0. \tag{45}$$

This is Equation (17) divided by D_{L}, since $\omega_{\mathrm{L}} = \delta\lambda_{\mathrm{L}} D_{\mathrm{L}}$ (Equation (12)).

Suppose now that T_{H} is strictly increasing, and hence invertible on its range, and that the floor does not bind, so that Equation (45) holds for every $\tau_{\mathrm{L}} \in \mathcal{A}(\tau_0)$. Substituting $\tau_{\mathrm{L}} = T_{\mathrm{H}}^{-1}(\tau_{\mathrm{H}})$ into Equation (45) then gives, for every τ_{H} in the range of T_{H},

$$r^{\mathrm{H}}_{\mathrm{H}}\left(\tau_{\mathrm{H}}, T_{\mathrm{H}}^{-1}(\tau_{\mathrm{H}})\right) + \delta\lambda_{\mathrm{L}}\, W'_{\mathrm{H}}\left(\tau_{\mathrm{H}}\right) = 0 \tag{46}$$

Similarly, once $\tau_{\mathrm{L}} = T_{\mathrm{L}}(\tau_{\mathrm{H}})$ is substituted, Equation (45) can be rearranged to obtain

$$r^{\mathrm{H}}_{\mathrm{H}}\left(T_{\mathrm{H}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) + \delta\lambda_{\mathrm{L}}\, W'_{\mathrm{H}}\left(T_{\mathrm{H}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right)\right) = 0. \tag{47}$$

The derivative of W_{H}, the value of watching the opponent move, is used in several places and thus formalized in the following lemma:

Lemma 9. *Let $(T_{\mathrm{H}}, T_{\mathrm{L}})$ be an MPE and $\tau_{\mathrm{H}} \in \mathcal{A}(\tau_0)$. Assume that T_{L} is differentiable at τ_{H}, T_{H} is differentiable at $\tau'_{\mathrm{L}} = T_{\mathrm{L}}(\tau_{\mathrm{H}})$, and W_{H} is differentiable at $T_{\mathrm{H}}(\tau'_{\mathrm{L}})$, all one-sided at the boundary*

of $\mathcal{A}(\tau_0)$; and that either Equation (47) holds or $T'_{\mathrm{H}}(\tau'_{\mathrm{L}}) = 0$. Then W_{H} is differentiable at τ_{H} and

$$W'_{\mathrm{H}}(\tau_{\mathrm{H}}) = D_{\mathrm{H}}\Big[r^{\mathrm{H}}_{\mathrm{H}}\big(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\big) + r^{\mathrm{H}}_{\mathrm{L}}\big(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\big)\, T'_{\mathrm{L}}(\tau_{\mathrm{H}})\Big] + \delta\lambda_{\mathrm{H}} D_{\mathrm{H}} D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}\big(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\big)\, T'_{\mathrm{L}}(\tau_{\mathrm{H}}). \tag{48}$$

The same applies to Player L after relabeling $\mathrm{H} \leftrightarrow \mathrm{L}$ *and* $\lambda_{\mathrm{H}} \leftrightarrow \lambda_{\mathrm{L}}$.

Proof. Start with the expression for W_{H} in Equation (40) and take the derivative with respect to τ_{H}, using the chain rule:

$$\begin{aligned}
W'_{\mathrm{H}}(\tau_{\mathrm{H}}) &= D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})) + D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
&\quad + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \delta\, D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{H}}\left(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{H}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
&\quad + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \delta\, D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
&\quad + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \lambda_{\mathrm{L}} D_{\mathrm{L}}\, \delta^2\, W'_{\mathrm{H}}\left(T_{\mathrm{H}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right)\right) T'_{\mathrm{H}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
&= D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})) + D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
&\quad + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \delta\, D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
&\quad + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \delta \underbrace{\left[\lambda_{\mathrm{L}} D_{\mathrm{L}}\, \delta\, W'_{\mathrm{H}}\left(T_{\mathrm{H}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right)\right) + D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{H}}\left(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right)\right] T'_{\mathrm{H}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right)}_{=0 \text{ by Equation (47), or } T'_{\mathrm{H}}(\tau'_{\mathrm{L}})=0} T'_{\mathrm{L}}(\tau_{\mathrm{H}})
\end{aligned}$$

Dropping the last term gives Equation (48). Because all elements of the right-hand side are differentiable at τ_{H} under the stated assumptions, so is W_{H}. □

Combining the result of Lemma 9 with Equation (46) yields

$$\begin{aligned}
-\frac{1}{\delta\,\lambda_{\mathrm{L}}} r^{\mathrm{H}}_{\mathrm{H}}\left(\tau_{\mathrm{H}}, T^{-1}_{\mathrm{H}}(\tau_{\mathrm{H}})\right) &= D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})) \\
&\quad + D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
&\quad + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \delta\, D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) T'_{\mathrm{L}}(\tau_{\mathrm{H}})
\end{aligned}$$

$$\begin{aligned}
\implies \quad & \left[D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \delta\, D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right)\right] T'_{\mathrm{L}}(\tau_{\mathrm{H}}) \\
= & -D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})) - \frac{1}{\delta\,\lambda_{\mathrm{L}}} r^{\mathrm{H}}_{\mathrm{H}}\left(\tau_{\mathrm{H}}, T^{-1}_{\mathrm{H}}(\tau_{\mathrm{H}})\right)
\end{aligned}$$

and thus

$$T'_{\mathrm{L}}(\tau_{\mathrm{H}}) = \frac{-D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{H}}\left(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) - \frac{1}{\delta\,\lambda_{\mathrm{L}}} r^{\mathrm{H}}_{\mathrm{H}}\left(\tau_{\mathrm{H}}, T^{-1}_{\mathrm{H}}(\tau_{\mathrm{H}})\right)}{D_{\mathrm{H}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) + \lambda_{\mathrm{H}} D_{\mathrm{H}}\, \delta\, D_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}\left(T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})), T_{\mathrm{L}}(\tau_{\mathrm{H}})\right)}. \tag{49}$$

This is the equivalent difference–differential condition for the Markov perfect strategy as given in Equation (13) in Maskin and Tirole (1987).

The same analysis applies to Player L after relabeling $\mathrm{H} \leftrightarrow \mathrm{L}$ and $\lambda_{\mathrm{H}} \leftrightarrow \lambda_{\mathrm{L}}$. Writing $W_{\mathrm{L}}(\tau_{\mathrm{L}}) := V_{\mathrm{L}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}})$ for Player L's value of watching Player H move, Player L's first-order condition is

$$r^{\mathrm{L}}_{\mathrm{L}}\left(\tau_{\mathrm{H}}, T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) + \delta\lambda_{\mathrm{H}}\, W'_{\mathrm{L}}\left(T_{\mathrm{L}}(\tau_{\mathrm{H}})\right) = 0, \tag{50}$$

and the same steps that led to Equation (49) yield

$$T'_{\mathrm{H}}(\tau_{\mathrm{L}}) = \frac{-D_{\mathrm{L}}\, r^{\mathrm{L}}_{\mathrm{L}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) - \frac{1}{\delta\,\lambda_{\mathrm{H}}} r^{\mathrm{L}}_{\mathrm{L}}\left(T^{-1}_{\mathrm{L}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right)}{D_{\mathrm{L}}\, r^{\mathrm{L}}_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\right) + \lambda_{\mathrm{L}} D_{\mathrm{L}}\, \delta\, D_{\mathrm{H}}\, r^{\mathrm{L}}_{\mathrm{H}}\left(T_{\mathrm{H}}(\tau_{\mathrm{L}}), T_{\mathrm{L}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}))\right)}.$$

A.4 Proof of Proposition 3 — Uniqueness of MPE

Proposition 3 (Section 3.2).

Proof. **Slope.** The proof is written for Player H; the steps for Player L follow by the relabeling $\mathrm{H} \leftrightarrow \mathrm{L}$. Let $T_i(\tau_{-i}) = k_i + m_i \tau_{-i}$, $i = \mathrm{H}, \mathrm{L}$, be affine strategies with slopes $m_i \in (0, 1)$ (Equation (21)). Since $m_{\mathrm{H}} > 0$, T_{H} is invertible, and

$$T^{-1}_{\mathrm{H}}(\tau_{\mathrm{H}}) = \frac{\tau_{\mathrm{H}} - k_{\mathrm{H}}}{m_{\mathrm{H}}}$$
$$T_{\mathrm{H}}(T_{\mathrm{L}}(\tau_{\mathrm{H}})) = k_{\mathrm{H}} + m_{\mathrm{H}}(k_{\mathrm{L}} + m_{\mathrm{L}}\tau_{\mathrm{H}})$$

Next, insert the affine strategies into the difference–differential equations of Appendix A.3 (Equation (49) and its analogue for Player L) to obtain two expressions that are linear in τ_i of the form

$$\phi_{i,0} + \phi_{i,1}\tau_i = 0,\ i = \mathrm{H}, \mathrm{L}. \tag{51}$$

Equation (51) holds on an interval of τ_i: an affine strategy with positive slope reaches the bounds of $\mathcal{A}(\tau_0)$ at most at an endpoint state, so the players are interior, and their first-order conditions hold, at all other states. An affine function of τ_i that is zero on an interval has both coefficients equal to zero, so four conditions are obtained, $\phi_{i,0} = 0$ and $\phi_{i,1} = 0$, for $i = \mathrm{H}, \mathrm{L}$. Multiplied by $\theta\delta\lambda_{\mathrm{L}} D_{\mathrm{L}} m_{\mathrm{H}} > 0$ and $\theta\delta\lambda_{\mathrm{H}} D_{\mathrm{H}} m_{\mathrm{L}} > 0$, respectively, the conditions $\phi_{\mathrm{H},1} = 0$ and $\phi_{\mathrm{L},1} = 0$ read

$$D_{\mathrm{L}} - 2D_{\mathrm{L}}m_{\mathrm{H}} - 2D_{\mathrm{L}}D_{\mathrm{H}}m_{\mathrm{H}}\delta\lambda_{\mathrm{L}} + 2D_{\mathrm{L}}D_{\mathrm{H}}m_{\mathrm{H}}m_{\mathrm{L}}\delta\lambda_{\mathrm{L}} + D^2_{\mathrm{L}}D_{\mathrm{H}}m^2_{\mathrm{H}}m^2_{\mathrm{L}}\delta^2\lambda_{\mathrm{H}}\lambda_{\mathrm{L}} = 0$$
$$D_{\mathrm{H}} - 2D_{\mathrm{H}}m_{\mathrm{L}} - 2D_{\mathrm{L}}D_{\mathrm{H}}m_{\mathrm{L}}\delta\lambda_{\mathrm{H}} + 2D_{\mathrm{L}}D_{\mathrm{H}}m_{\mathrm{H}}m_{\mathrm{L}}\delta\lambda_{\mathrm{H}} + D^2_{\mathrm{H}}D_{\mathrm{L}}m^2_{\mathrm{H}}m^2_{\mathrm{L}}\delta^2\lambda_{\mathrm{H}}\lambda_{\mathrm{L}} = 0.$$

Since $\lambda_{\mathrm{H}} = \lambda_{\mathrm{L}} \equiv \lambda$, $D_{\mathrm{H}} = D_{\mathrm{L}} \equiv D$, and $\delta\lambda D = \omega$ (Equations (11) and (12)). Subtracting the two conditions leaves $-2D(1+\omega)(m_{\mathrm{H}} - m_{\mathrm{L}}) = 0$, so $m_{\mathrm{H}} = m_{\mathrm{L}} \equiv m$, and either condition, divided by $D > 0$, becomes the slope equation (22).

Write $p(m)$ for the left-hand side of Equation (22). Since $p''(m) = 4\omega(1 + 3\omega m^2) > 0$, p is strictly convex and has at most two real roots. With $0 < \omega < 1$ (Equation (12)), $p(0) = 1 > 0$, $p(1/2) = \omega(\omega/16 - 1/2) < 0$, and $p(1) = \omega^2 - 1 < 0$. Hence p has exactly one root in $(0, 1)$, and it lies in $(0, 1/2)$; the other real root exceeds one and lies outside the class of slopes considered.

Intercepts. Given $\phi_{i,1} = 0$, Equation (51) holds for all τ_i if and only if it holds at a single τ_i. The equation is evaluated at the fixed point τ^* of $(T_{\mathrm{H}}, T_{\mathrm{L}})$, unique since $m^2 < 1$. At $\tau_{\mathrm{H}} = \tau^*_{\mathrm{H}}$, all composed arguments in Equation (49) collapse to τ^* (e.g., $T_{\mathrm{L}}(\tau^*_{\mathrm{H}}) = \tau^*_{\mathrm{L}}$), and the equation multiplied by its denominator and by ω becomes $D(1+\omega)\left[r^{\mathrm{H}}_{\mathrm{H}}(\tau^*) + \omega m\, r^{\mathrm{H}}_{\mathrm{L}}(\tau^*)\right] = 0$, that is, Equation (20) for $i = \mathrm{H}$ with $\hat{m}_{\mathrm{L}} = m$.

The same holds for Player L. Written out as in Equation (53) in the proof of Proposition 2, the two conditions read $(2-\omega m)\,\tau_i^* - \tau_{-i}^* = \theta - (1-\omega m)\,a_i$, $i = \mathrm{H}, \mathrm{L}$. Adding them gives $(1-\omega m)(\tau_\mathrm{H}^* + \tau_\mathrm{L}^*) = 2\theta - (1-\omega m)(a_\mathrm{H} + a_\mathrm{L})$, and subtracting them gives $(3-\omega m)(\tau_\mathrm{H}^* - \tau_\mathrm{L}^*) = (1-\omega m)(a_\mathrm{L} - a_\mathrm{H})$; hence $\tau_\mathrm{H}^* = \sigma + \eta$ and $\tau_\mathrm{L}^* = \sigma - \eta$, with σ and η of Equation (23). Since $T_i(\tau_{-i}^*) = \tau_i^*$, $k_i = \tau_i^* - m\,\tau_{-i}^*$, which gives Equation (24). Conversely, with these intercepts $\phi_{\mathrm{H},0} = \phi_{\mathrm{L},0} = 0$, so the strategies of the proposition satisfy Equation (49) and its analogue for Player L at every state.

Feasibility. The strategies map $\mathcal{A}(\tau_0)$ into itself. Since $a_\mathrm{H} \le a_\mathrm{L}$, $\eta \ge 0$ and hence $k_\mathrm{L} \le k_\mathrm{H}$ (Equation (24)). The assumption $\tau_0 \le k_\mathrm{L}/(1-m)$ gives $k_i \ge k_\mathrm{L} \ge (1-m)\,\tau_0$, so $T_i(\tau_{-i}) \ge k_i + m\,\tau_0 \ge \tau_0$ for every $\tau_{-i} \in \mathcal{A}(\tau_0)$; the assumption $k_\mathrm{H} + m \le 1$ gives $T_i(\tau_{-i}) \le k_i + m \le k_\mathrm{H} + m \le 1$. Hence play never leaves $[\tau_0, 1]^2$, and $T_\mathrm{H} \circ T_\mathrm{L}$ and $T_\mathrm{L} \circ T_\mathrm{H}$ map $\mathcal{A}(\tau_0)$ into itself.

Optimality. It remains to show that these strategies are optimal replies to each other. For affine strategies, the loop equation (40) reads $W_\mathrm{H} = q + \omega^2\, W_\mathrm{H} \circ h$, with $q(\tau_\mathrm{H}) = D\,r^\mathrm{H}\big(\tau_\mathrm{H}, T_\mathrm{L}(\tau_\mathrm{H})\big) + \omega D\,r^\mathrm{H}\big(h(\tau_\mathrm{H}), T_\mathrm{L}(\tau_\mathrm{H})\big)$ quadratic and $h \equiv T_\mathrm{H} \circ T_\mathrm{L}$ affine.

Write $q(\tau) = q_0 + q_1\tau + q_2\tau^2$ and $h(\tau) = \alpha + \beta\tau$, with $\beta = m^2$. For a quadratic $F(\tau) = c_0 + c_1\tau + c_2\tau^2$, $F \circ h$ is again quadratic, and matching the coefficients of τ^2, τ, and 1 in $F = q + \omega^2\, F \circ h$ gives

$$c_2(1-\omega^2\beta^2) = q_2, \qquad c_1(1-\omega^2\beta) = q_1 + 2\omega^2\alpha\beta\, c_2, \qquad c_0(1-\omega^2) = q_0 + \omega^2(\alpha c_1 + \alpha^2 c_2),$$

which determine c_2, c_1, and c_0 in turn, since $\omega^2\beta^j < 1$. No other bounded function solves the equation: if F and $\widetilde{F}$ both do, their difference Δ satisfies $\Delta = \omega^2\,\Delta \circ h$, and since h maps $\mathcal{A}(\tau_0)$ into itself, $\sup|\Delta| \le \omega^2 \sup|\Delta|$, so $\Delta = 0$ because $\omega < 1$ (Equation (12)). Since returns are bounded on the compact state space, W_H is bounded; hence it is this quadratic, and $W_\mathrm{H}'' = 2c_2$ is constant.

Let

$$R(\tau_\mathrm{L}) \equiv r_\mathrm{H}^\mathrm{H}\big(T_\mathrm{H}(\tau_\mathrm{L}), \tau_\mathrm{L}\big) + \delta\lambda\, W_\mathrm{H}'\big(T_\mathrm{H}(\tau_\mathrm{L})\big) \tag{52}$$

be the left-hand side of the first-order condition (45) for these strategies. I will now show that it vanishes for the strategies $T_i(\tau_{-i}) = k_i + m\,\tau_{-i}$ with the slope and intercepts found above.

For these strategies, the loop equation reads

$$W_\mathrm{H}(\tau_\mathrm{H}) = D\,r^\mathrm{H}\big(\tau_\mathrm{H}, T_\mathrm{L}(\tau_\mathrm{H})\big) + \omega D\,r^\mathrm{H}\big(h(\tau_\mathrm{H}), T_\mathrm{L}(\tau_\mathrm{H})\big) + \omega^2\, W_\mathrm{H}\big(h(\tau_\mathrm{H})\big).$$

Differentiating it, with $T_\mathrm{L}' = m$ and $h' = m^2$, and without using any first-order condition,

$$\begin{aligned} W_\mathrm{H}'(\tau_\mathrm{H}) = {} & D\Big[r_\mathrm{H}^\mathrm{H}\big(\tau_\mathrm{H}, T_\mathrm{L}(\tau_\mathrm{H})\big) + m\, r_\mathrm{L}^\mathrm{H}\big(\tau_\mathrm{H}, T_\mathrm{L}(\tau_\mathrm{H})\big)\Big] + \omega D\, m\, r_\mathrm{L}^\mathrm{H}\big(h(\tau_\mathrm{H}), T_\mathrm{L}(\tau_\mathrm{H})\big) \\ & + \omega m^2\Big[D\, r_\mathrm{H}^\mathrm{H}\big(h(\tau_\mathrm{H}), T_\mathrm{L}(\tau_\mathrm{H})\big) + \omega\, W_\mathrm{H}'\big(h(\tau_\mathrm{H})\big)\Big]. \end{aligned}$$

The first line is the right-hand side of Equation (48); since $h(\tau_\mathrm{H}) = T_\mathrm{H}(T_\mathrm{L}(\tau_\mathrm{H}))$ and $\omega = \delta\lambda D$, the bracket in the second line equals $D\,R\big(T_\mathrm{L}(\tau_\mathrm{H})\big)$, the term that Lemma 9 drops by assuming the first-order condition. Now evaluate at $\tau_\mathrm{H} = x \equiv T_\mathrm{H}(\tau_\mathrm{L})$ and insert into $R(\tau_\mathrm{L}) = r_\mathrm{H}^\mathrm{H}(x, \tau_\mathrm{L}) + \delta\lambda\, W_\mathrm{H}'(x)$, using $\tau_\mathrm{L} = T_\mathrm{H}^{-1}(x)$ and $\delta\lambda\,\omega D = \omega^2$:

$$\begin{aligned} R(\tau_\mathrm{L}) = {} & \underbrace{r_\mathrm{H}^\mathrm{H}\big(x, T_\mathrm{H}^{-1}(x)\big) + \delta\lambda\,\Big(D\big[r_\mathrm{H}^\mathrm{H}(x, T_\mathrm{L}(x)) + m\, r_\mathrm{L}^\mathrm{H}(x, T_\mathrm{L}(x))\big] + \omega D\, m\, r_\mathrm{L}^\mathrm{H}\big(h(x), T_\mathrm{L}(x)\big)\Big)}_{= \delta\lambda\, E_\mathrm{H}(x)} \\ & + \omega^2 m^2\, R\big(T_\mathrm{L}(x)\big). \end{aligned}$$

The underbraced term is $\delta\lambda$ times Equation (49) multiplied by its denominator and written as $E_{\mathrm{H}}(x) = 0$. It vanishes for every x: $E_{\mathrm{H}}(x) = \phi_{\mathrm{H},0} + \phi_{\mathrm{H},1}x$, and the slope and intercept parts chose m, k_{H}, and k_{L} such that $\phi_{\mathrm{H},1} = \phi_{\mathrm{H},0} = 0$.

Hence $R(\tau_{\mathrm{L}}) = \omega^2 m^2\, R\big(T_{\mathrm{L}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}))\big)$ for every τ_{L}. Since W_{H} was shown to be quadratic, W'_{H} is affine, and so is R, $R(\tau_{\mathrm{L}}) = \rho_0 + \rho_1\tau_{\mathrm{L}}$, and $T_{\mathrm{L}}(T_{\mathrm{H}}(\tau_{\mathrm{L}})) = \gamma + m^2\tau_{\mathrm{L}}$ with $\gamma = k_{\mathrm{L}} + mk_{\mathrm{H}}$; matching the coefficients of τ_{L} gives $\rho_1(1 - \omega^2 m^4) = 0$, so $\rho_1 = 0$, and matching the constants then gives $\rho_0(1 - \omega^2 m^2) = 0$, so $\rho_0 = 0$. Hence $R \equiv 0$: the first-order condition holds at every state.

Differentiating $R \equiv 0$ from Equation (52) with respect to τ_{L}, with $T'_{\mathrm{H}} = m$,

$$r^{\mathrm{H}}_{\mathrm{HL}} + \left[r^{\mathrm{H}}_{\mathrm{HH}} + \delta\lambda\, W''_{\mathrm{H}}\right] m = 0, \qquad \text{so} \qquad r^{\mathrm{H}}_{\mathrm{HH}} + \delta\lambda\, W''_{\mathrm{H}} = -\frac{r^{\mathrm{H}}_{\mathrm{HL}}}{m} = -\frac{1}{\theta m} < 0$$

by Equation (7). By Equation (37), Player H's objective when called to move against τ_{L} is, up to a positive factor and terms that do not depend on its own rate, $\tau_{\mathrm{H}} \mapsto r^{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) + \delta\lambda\, W_{\mathrm{H}}(\tau_{\mathrm{H}})$, a quadratic whose second derivative is this negative constant. Its unique stationary point $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ is therefore its unique maximizer over the real line, and hence over $\mathcal{A}(\tau_0)$, which contains it. The values of these strategies satisfy the value recursion (13) by construction, and now also the optimality condition (14); by the characterization in Section 2.3, the strategies form an MPE.

Steady state. The strategies were chosen such that $E_{\mathrm{H}} \equiv 0$, and likewise for Player L; at the fixed point $\tau_{\mathrm{H}} = \tau^*_{\mathrm{H}}$, this is Equation (20) for $i = \mathrm{H}$ with $\hat{m}_{\mathrm{L}} = m$, as shown in the intercept part, and likewise for $i = \mathrm{L}$. Written out as in Equation (53), the steady state therefore satisfies the wedge representation (18) and its closed form (19) for both players with $\hat{m}_{-i} = m$, without recourse to the hypotheses of Proposition 2.

Convergence.

For the last part of the proof, start by noticing that the linear–quadratic MPE $(T_{\mathrm{H}}, T_{\mathrm{L}})$ characterized above induces a sequence of tax profiles $(\tau_{t+s})_{s\in\mathbb{N}_0} \in \mathcal{S}$, where at each point in time, $\tau_{t+s} = (\tau_{\mathrm{H},t+s}, \tau_{\mathrm{L},t+s})$. Start by considering the related sequence $(\tilde{\tau}_{t+s})_{s\in\mathbb{N}_0}$, $\tilde{\tau}_{t+1} = \tau_{t+s}$ for the first s where one of the two players moves, $\tilde{\tau}_{t+2} = \tau_{t+s}$ for the second s where one of the two players moves, and so on. In other words, $(\tilde{\tau}_{t+s})_{s\in\mathbb{N}_0}$ is obtained from the original series by dropping the periods in which no player moves and those in which a player moves again before its opponent has moved; in both cases the tax profile does not change, since a player facing the same opponent rate re-chooses the same rate. Since in every period each player moves with probability $\lambda > 0$, independently across periods, both players move infinitely often with probability one, so the reduced sequence is infinite.

Hence, $(\tilde{\tau}_{t+s})_{s\in\mathbb{N}_0}$ tracks the moves in an alternating two-player game as in Maskin and Tirole (1988). Without loss of generality, I assume that Player H (Player L) moves at the odd (even) steps of the reduced sequence. Given linear strategies $T_{\mathrm{H}}, T_{\mathrm{L}}$, one can represent an iteration of the game in matrix form:

$$\begin{aligned}\tilde{\tau}_{d+1} &= A_{\mathrm{H}} \cdot \tilde{\tau}_d + \kappa_{\mathrm{H}}\\ \tilde{\tau}_{e+1} &= A_{\mathrm{L}} \cdot \tilde{\tau}_e + \kappa_{\mathrm{L}}\end{aligned}$$

where

$$A_{\mathrm{H}} = \begin{pmatrix} 0 & m_{\mathrm{H}} \\ 0 & 1 \end{pmatrix}, \; A_{\mathrm{L}} = \begin{pmatrix} 1 & 0 \\ m_{\mathrm{L}} & 0 \end{pmatrix}$$

and $\kappa_{\mathrm{H}} = (k_{\mathrm{H}}, 0)', \kappa_{\mathrm{L}} = (0, k_{\mathrm{L}})'$. The respective two-period ahead iterations

$$\begin{aligned} \tilde{\tau}_{d+2} &= A_{\mathrm{L}} \cdot A_{\mathrm{H}} \cdot \tilde{\tau}_d + A_{\mathrm{L}} \cdot \kappa_{\mathrm{H}} + \kappa_{\mathrm{L}} \\ &= A_{\mathrm{H,L}} \cdot \tilde{\tau}_d + \kappa_{\mathrm{H,L}} \end{aligned}$$

are given by the matrices

$$A_{\mathrm{H,L}} = A_{\mathrm{L}} \cdot A_{\mathrm{H}} = \begin{pmatrix} 0 & m_{\mathrm{H}} \\ 0 & m_{\mathrm{H}} \cdot m_{\mathrm{L}} \end{pmatrix}, \; A_{\mathrm{L,H}} = A_{\mathrm{H}} \cdot A_{\mathrm{L}} = \begin{pmatrix} m_{\mathrm{H}} \cdot m_{\mathrm{L}} & 0 \\ m_{\mathrm{L}} & 0 \end{pmatrix}.$$

and the translation vectors

$$\kappa_{\mathrm{H,L}} = A_{\mathrm{L}} \cdot \kappa_{\mathrm{H}} + \kappa_{\mathrm{L}} = \begin{pmatrix} k_{\mathrm{H}} \\ m_{\mathrm{L}} k_{\mathrm{H}} + k_{\mathrm{L}} \end{pmatrix}, \; \kappa_{\mathrm{L,H}} = \begin{pmatrix} m_{\mathrm{H}} k_{\mathrm{L}} + k_{\mathrm{H}} \\ k_{\mathrm{L}} \end{pmatrix}$$

Both two-step matrices have the eigenvalues 0 and $m_{\mathrm{H}} m_{\mathrm{L}}$. Since both $m_{\mathrm{H}}, m_{\mathrm{L}} \in (0, 1)$, both eigenvalues are less than one in absolute value. If Player H moves at step d, repeated forward iterations give

$$\tilde{\tau}_{d+2n} = (A_{\mathrm{H,L}})^n \cdot \tilde{\tau}_d + \sum_{j=0}^{n-1} (A_{\mathrm{H,L}})^j \, \kappa_{\mathrm{H,L}}.$$

Moreover,

$$(A_{\mathrm{H,L}})^n \cdot \tilde{\tau}_d = \begin{pmatrix} m_{\mathrm{H}} (m_{\mathrm{H}} m_{\mathrm{L}})^{n-1} \cdot \tilde{\tau}_{d,\mathrm{L}} \\ (m_{\mathrm{H}} m_{\mathrm{L}})^n \cdot \tilde{\tau}_{d,\mathrm{L}} \end{pmatrix}$$

which tends to 0 for $n \to \infty$. The same is true for $(A_{\mathrm{L,H}})^n \cdot \tilde{\tau}_e$, i.e., if Player L moves first. Hence, I can compute the limit of the sequence $\tilde{\tau}_{d+2n}$ using the matrix version of the geometric series, since both eigenvalues of $A_{\mathrm{H,L}}$ lie within the unit circle:

$$\tilde{\tau}_d^* := \lim_{n \to \infty} \tilde{\tau}_{d+2n} = \sum_{j=0}^{\infty} (A_{\mathrm{H,L}})^j \, \kappa_{\mathrm{H,L}} = (I - A_{\mathrm{H,L}})^{-1} \kappa_{\mathrm{H,L}},$$

and analogously $\tilde{\tau}_e^* = (I - A_{\mathrm{L,H}})^{-1} \kappa_{\mathrm{L,H}}$ along the steps at which Player H is next to move.

The two inverted matrices are given by

$$(I - A_{\mathrm{H,L}})^{-1} = \begin{pmatrix} 1 & \frac{m_{\mathrm{H}}}{1 - m_{\mathrm{H}} \cdot m_{\mathrm{L}}} \\ 0 & \frac{1}{1 - m_{\mathrm{H}} \cdot m_{\mathrm{L}}} \end{pmatrix}$$

$$(I - A_{\mathrm{L,H}})^{-1} = \begin{pmatrix} \frac{1}{1 - m_{\mathrm{H}} \cdot m_{\mathrm{L}}} & 0 \\ \frac{m_{\mathrm{L}}}{1 - m_{\mathrm{H}} \cdot m_{\mathrm{L}}} & 1 \end{pmatrix}$$

The limit is the same irrespective of the first mover and of the starting point. Since the original sequence of tax profiles coincides with the reduced one between moves, it converges to the same limit. With $m_{\mathrm{H}} = m_{\mathrm{L}} = m$, the limit is $\big((k_{\mathrm{H}} + m k_{\mathrm{L}})/(1 - m^2), \, (k_{\mathrm{L}} + m k_{\mathrm{H}})/(1 - m^2)\big) = (\sigma + \eta, \, \sigma - \eta) = \tau^*$, the fixed point of the intercept part.

□

A.5 Proof of Proposition 2 — Generalized Steady-State Characterization

Proposition 2 (Section 3.1).

Proof. The proof is written for player $i = \mathrm{H}$; the case $i = \mathrm{L}$ follows by the relabeling $\mathrm{H} \leftrightarrow \mathrm{L}$. With $\lambda_{\mathrm{H}} = \lambda_{\mathrm{L}} \equiv \lambda$, $D_{\mathrm{H}} = D_{\mathrm{L}} \equiv D$, and $\omega = \delta\lambda D$ (Equation (12)). By Equations (4) and (5), $r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = 1 + (\tau_{\mathrm{L}} - 2\tau_{\mathrm{H}} - a_{\mathrm{H}})/\theta$ and $r^{\mathrm{H}}_{\mathrm{L}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = (\tau_{\mathrm{H}} + a_{\mathrm{H}})/\theta$. All derivatives at the steady state are understood one-sidedly where $(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}})$ lies on the boundary of $[\tau_0, 1]^2$.

Since $\hat{\tau}_{\mathrm{H}}$ is interior for Player H, the first-order condition (17) holds at $\tau_{\mathrm{L}} = \hat{\tau}_{\mathrm{L}}$:

$$D\, r^{\mathrm{H}}_{\mathrm{H}}\big(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}}\big) + \omega\, W'_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}) = 0.$$

Lemma 9 gives an expression for $W'_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}})$ in Equation (48); its hypotheses hold at $\tau_{\mathrm{H}} = \hat{\tau}_{\mathrm{H}}$: T_{L} is differentiable at $\hat{\tau}_{\mathrm{H}}$ by (c), T_{H} at $T_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}) = \hat{\tau}_{\mathrm{L}}$ by (b), W_{H} at $T_{\mathrm{H}}(\hat{\tau}_{\mathrm{L}}) = \hat{\tau}_{\mathrm{H}}$ by (b) and (c), and Equation (47) holds since Player H is interior by (a). By the steady-state property ($T_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}) = \hat{\tau}_{\mathrm{L}}$, $T_{\mathrm{H}}(T_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}})) = T_{\mathrm{H}}(\hat{\tau}_{\mathrm{L}}) = \hat{\tau}_{\mathrm{H}}$), all composed arguments in Equation (48) evaluated at $\tau_{\mathrm{H}} = \hat{\tau}_{\mathrm{H}}$ collapse to $(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}})$.

With $T'_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}) = \hat{m}_{\mathrm{L}}$, Equation (48) becomes $W'_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}) = D\big[r^{\mathrm{H}}_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}}) + (1+\omega)\,\hat{m}_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}})\big]$. Substituting into the first-order condition and collecting terms,

$$0 = D(1+\omega)\big[r^{\mathrm{H}}_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}}) + \omega\hat{m}_{\mathrm{L}}\, r^{\mathrm{H}}_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{L}})\big],$$

which, divided by $D(1+\omega) > 0$, is Equation (20) for $i = \mathrm{H}$.

Writing out $r^{\mathrm{H}}_{\mathrm{H}} = \big(\theta + \hat{\tau}_{\mathrm{L}} - 2\hat{\tau}_{\mathrm{H}} - a_{\mathrm{H}}\big)/\theta$ and $r^{\mathrm{H}}_{\mathrm{L}} = (\hat{\tau}_{\mathrm{H}} + a_{\mathrm{H}})/\theta$ in Equation (20) and multiplying by θ,

$$\theta + \hat{\tau}_{\mathrm{L}} - 2\hat{\tau}_{\mathrm{H}} - a_{\mathrm{H}} + \omega\hat{m}_{\mathrm{L}}\big(\hat{\tau}_{\mathrm{H}} + a_{\mathrm{H}}\big) = 0 \qquad \Longleftrightarrow \qquad \hat{\tau}_{\mathrm{H}} = \frac{\theta - a_{\mathrm{H}} + \hat{\tau}_{\mathrm{L}}}{2} + \frac{\omega\hat{m}_{\mathrm{L}}}{2}\big(\hat{\tau}_{\mathrm{H}} + a_{\mathrm{H}}\big), \quad (53)$$

which is Equation (18); since $\omega < 1$ (Equation (12)) and $\hat{m}_{\mathrm{L}} < 1$, the coefficient $2 - \omega\hat{m}_{\mathrm{L}}$ of $\hat{\tau}_{\mathrm{H}}$ exceeds one, and solving for $\hat{\tau}_{\mathrm{H}}$ gives Equation (19).

For the comparative statics, write $g \equiv \omega\hat{m}_{\mathrm{L}} \in [0, 1)$ and multiply Equation (19) by $2 - g$:

$$(2-g)\,\hat{\tau}_{\mathrm{H}} = \theta - (1-g)\,a_{\mathrm{H}} + \hat{\tau}_{\mathrm{L}}.$$

Differentiating,

$$\frac{\partial\hat{\tau}_{\mathrm{H}}}{\partial g} = \frac{\hat{\tau}_{\mathrm{H}} + a_{\mathrm{H}}}{2-g}, \qquad \frac{\partial\hat{\tau}_{\mathrm{H}}}{\partial\hat{\tau}_{\mathrm{L}}} = \frac{1}{2-g} > 0,$$

where the first derivative is positive by the proposition's proviso $\hat{\tau}_{\mathrm{H}} + a_{\mathrm{H}} > 0$. Since $g = \omega\hat{m}_{\mathrm{L}}$ is strictly increasing in $\hat{m}_{\mathrm{L}}$, $\hat{\tau}_{\mathrm{H}}$ is strictly increasing in $\hat{m}_{\mathrm{L}}$; and the wedge $\frac{g}{2}(\hat{\tau}_{\mathrm{H}} + a_{\mathrm{H}})$ in Equation (18) is zero if and only if $g = 0$, that is, if and only if $\hat{m}_{\mathrm{L}} = 0$ (given $\hat{\tau}_{\mathrm{H}} + a_{\mathrm{H}} > 0$ and $\omega > 0$). □

A.6 Proof of Proposition 5 — Local MPE at the Floor

Proposition 5 (Section 4).

Proof. Guess and verify. Fix b and ε; the conditions they must satisfy arise in the first two parts and are collected in the third. Throughout, V_i and W_i denote player i's value and watch value under the strategy profile $(T_{\mathrm{H}}, T_{\mathrm{L}})$ of Equation (30) in the game with action sets $\mathcal{A}_{\mathrm{H}}$ and $\mathcal{A}_{\mathrm{L}}$. The guess is that Player L plays the floor at every $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$. The first part derives Player H's best reply to this guess on $\mathcal{A}_{\mathrm{L}}$. The second part verifies that, given this reply, the floor maximizes Player L's value on $\mathcal{A}_{\mathrm{L}}$ in every state $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$. By the one-shot deviation principle (Section 2.3), the two parts establish an MPE of the game with action sets $\mathcal{A}_{\mathrm{H}}$ and $\mathcal{A}_{\mathrm{L}}$, for the b and ε of the third part. The fourth part shows that the tax profile reaches the steady state.

Player H's strategy. Let $\tau_{\mathrm{L}} \in \mathcal{A}_{\mathrm{L}}$. Under the guess, $T_{\mathrm{L}}(\tau_{\mathrm{H}}) = \underline{\tau}$ for every $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$: whatever rate Player H sets, Player L returns to the floor at its next move, and Player H's own next move is then its reply to the floor, $T_{\mathrm{H}}(\underline{\tau})$. The derivations of Equations (15) and (40) carry over to the game with action sets $\mathcal{A}_{\mathrm{H}}$ and $\mathcal{A}_{\mathrm{L}}$ unchanged. With $D_{\mathrm{H}} = D_{\mathrm{L}} = D$ and $\delta\lambda D = \omega$, Equation (40) therefore reads

$$W_{\mathrm{H}}(\tau_{\mathrm{H}}) = D\, r^{\mathrm{H}}(\tau_{\mathrm{H}}, \underline{\tau}) + \omega \Big[D\, r^{\mathrm{H}}\big(T_{\mathrm{H}}(\underline{\tau}), \underline{\tau}\big) + \omega\, W_{\mathrm{H}}\big(T_{\mathrm{H}}(\underline{\tau})\big) \Big], \qquad \tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}.$$

The bracketed term is constant with respect to τ_{H} and thus drops out when taking the derivative:

$$W'_{\mathrm{H}}(\tau_{\mathrm{H}}) = D\, r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \underline{\tau}), \qquad \tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}.$$

In state τ_{L}, Player H maximizes $V_{\mathrm{H}}(\cdot, \tau_{\mathrm{L}})$ over $\mathcal{A}_{\mathrm{H}}$. Inserting W'_{H} into Equation (15), with $\delta\lambda D = \omega$,

$$(1 - \delta\lambda_\emptyset)\, \frac{\partial V_{\mathrm{H}}}{\partial \tau_{\mathrm{H}}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) + \omega\, r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \underline{\tau}), \qquad \tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}.$$

The right-hand side falls in τ_{H} with slope $-2(1+\omega)/\theta$, and $1 - \delta\lambda_\emptyset > 0$, so $V_{\mathrm{H}}(\cdot, \tau_{\mathrm{L}})$ is strictly concave on $\mathcal{A}_{\mathrm{H}}$. Hence, the first-order condition for Player H — sufficient for a maximum, due to strict concavity — is given by

$$r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) + \omega\, r^{\mathrm{H}}_{\mathrm{H}}(\tau_{\mathrm{H}}, \underline{\tau}) = 0.$$

Using $r^{\mathrm{H}}_{\mathrm{H}} = 1 + \frac{\tau_{\mathrm{L}} - 2\tau_{\mathrm{H}} - a_{\mathrm{H}}}{\theta}$ (Equation (4)), multiplying by θ, and collecting terms gives

$$0 = (1+\omega)(\theta - a_{\mathrm{H}}) + \tau_{\mathrm{L}} + \omega\underline{\tau} - 2(1+\omega)\,\tau_{\mathrm{H}},$$

and therefore

$$\begin{aligned} \tau_{\mathrm{H}} &= \frac{(1+\omega)(\theta - a_{\mathrm{H}}) + \tau_{\mathrm{L}} + \omega\underline{\tau}}{2(1+\omega)} \\ &= \hat{\tau}_{\mathrm{H}} + s_{\mathrm{H}}\,(\tau_{\mathrm{L}} - \underline{\tau}), \qquad s_{\mathrm{H}} \equiv \frac{1}{2(1+\omega)}, \end{aligned} \tag{54}$$

for $\tau_{\mathrm{L}} \in \mathcal{A}_{\mathrm{L}}$.

This rate lies in $[\hat{\tau}_{\mathrm{H}}, \hat{\tau}_{\mathrm{H}} + s_{\mathrm{H}}\varepsilon]$. Its lower end exceeds the floor, $\hat{\tau}_{\mathrm{H}} > \underline{\tau}$, because $\underline{\tau} < \theta - a_{\mathrm{H}}$. Its upper end does not exceed b if

$$\varepsilon \leq \frac{b - \hat{\tau}_{\mathrm{H}}}{s_{\mathrm{H}}} = 2(1+\omega)\,(b - \hat{\tau}_{\mathrm{H}}). \tag{C1}$$

Condition (C1) bounds Player L's action set $\mathcal{A}_{\mathrm{L}} = [\underline{\tau}, \underline{\tau} + \varepsilon]$ from above: the largest admissible ε depends on b, but it is strictly positive because $b > \hat{\tau}_{\mathrm{H}}$, so an $\varepsilon > 0$ that satisfies (C1) always exists. Under (C1) the rate in Equation (54) lies in $\mathcal{A}_{\mathrm{H}}$ and, by strict concavity, is the unique maximizer of $V_{\mathrm{H}}(\cdot, \tau_{\mathrm{L}})$ on $\mathcal{A}_{\mathrm{H}}$. Hence $T_{\mathrm{H}}(\tau_{\mathrm{L}}) = \hat{\tau}_{\mathrm{H}} + s_{\mathrm{H}}(\tau_{\mathrm{L}} - \underline{\tau})$ is Player H's best reply on $\mathcal{A}_{\mathrm{L}}$. For $\tau_{\mathrm{L}} = \underline{\tau}$ exactly, it reduces to $T_{\mathrm{H}}(\underline{\tau}) = \frac{1}{2}(\theta - a_{\mathrm{H}} + \underline{\tau}) = \hat{\tau}_{\mathrm{H}}$, which is Player H's static best response to the floor.

The slope of Player H's strategy on $\mathcal{A}_{\mathrm{L}}$ is thus $s_{\mathrm{H}} = 1/\big(2(1+\omega)\big)$. Note that $s_{\mathrm{H}} < m$, the slope of the affine MPE (Proposition 3). To see this, rearrange the quartic equation (22) that characterizes the slope m to

$$m = \frac{1 + 2\omega m^2 + \omega^2 m^4}{2(1+\omega)} > \frac{1}{2(1+\omega)} = s_{\mathrm{H}},$$

where the inequality follows from $m > 0$ and $\omega > 0$.

Player L at the floor. Now take Player H's strategy T_{H} as given. I will now show that, given T_{H} and Player L's own flat strategy at its later moves, the floor maximizes $V_{\mathrm{L}}(\tau_{\mathrm{H}}, \cdot)$ on $\mathcal{A}_{\mathrm{L}}$ in every state $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$: Player L's marginal return to raising its rate is negative for any $\tau_{\mathrm{H}} \leq \hat{\tau}_{\mathrm{H}}$, and this extends to values above $\hat{\tau}_{\mathrm{H}}$, up to b. First, start with Player L's marginal return to raising its rate just above the floor (Equation (15) with the players relabeled),

$$(1 - \delta\lambda_{\emptyset})\,\frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\tau_{\mathrm{H}}, \underline{\tau}) = r^{\mathrm{L}}_{\mathrm{L}}(\tau_{\mathrm{H}}, \underline{\tau}) + \delta\lambda\, W'_{\mathrm{L}}(\underline{\tau}),$$

which depends on the marginal effect on the watch value, $W'_{\mathrm{L}}(\underline{\tau})$. Since $\underline{\tau}$ is the lower end of $\mathcal{A}_{\mathrm{L}}$, all derivatives with respect to τ_{L} at the floor are right derivatives. To obtain $W'_{\mathrm{L}}(\underline{\tau})$, start from the watch value itself. For $\tau_{\mathrm{L}} \in \mathcal{A}_{\mathrm{L}}$, Equation (40) with the players relabeled reads

$$W_{\mathrm{L}}(\tau_{\mathrm{L}}) = D\, r^{\mathrm{L}}\big(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\big) + \omega\Big[D\, r^{\mathrm{L}}\big(T_{\mathrm{H}}(\tau_{\mathrm{L}}), T_{\mathrm{L}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}))\big) + \omega\, W_{\mathrm{L}}\big(T_{\mathrm{L}}(T_{\mathrm{H}}(\tau_{\mathrm{L}}))\big)\Big].$$

Two properties of the strategies enter. First, under (C1), Player H's reply $T_{\mathrm{H}}(\tau_{\mathrm{L}})$ lies in $\mathcal{A}_{\mathrm{H}}$, so $T_{\mathrm{L}}(T_{\mathrm{H}}(\tau_{\mathrm{L}})) = \underline{\tau}$. The deviation considered is a one-shot deviation, in which Player L raises its rate at the current move only and follows its flat strategy again at its next move. Second, Player H's reply to the floor is $T_{\mathrm{H}}(\underline{\tau}) = \hat{\tau}_{\mathrm{H}}$, and its strategy has slope $T'_{\mathrm{H}} = s_{\mathrm{H}}$. With the first property, the watch value becomes

$$W_{\mathrm{L}}(\tau_{\mathrm{L}}) = D\, r^{\mathrm{L}}\big(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \tau_{\mathrm{L}}\big) + \omega\Big[D\, r^{\mathrm{L}}\big(T_{\mathrm{H}}(\tau_{\mathrm{L}}), \underline{\tau}\big) + \omega\, W_{\mathrm{L}}(\underline{\tau})\Big], \qquad \tau_{\mathrm{L}} \in \mathcal{A}_{\mathrm{L}}.$$

Differentiate at $\tau_{\mathrm{L}} = \underline{\tau}$ and use the second property. The first term gives $D\big[r^{\mathrm{L}}_{\mathrm{L}} + r^{\mathrm{L}}_{\mathrm{H}}\, s_{\mathrm{H}}\big]$, the second gives $\omega D\, r^{\mathrm{L}}_{\mathrm{H}}\, s_{\mathrm{H}}$, and the third drops out because it does not change with τ_{L}. The marginal effect on the watch value is therefore

$$\begin{aligned} W'_{\mathrm{L}}(\underline{\tau}) = D\,\big[r^{\mathrm{L}}_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) + r^{\mathrm{L}}_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau})\, s_{\mathrm{H}}\big] \\ + \omega D\, r^{\mathrm{L}}_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau})\, s_{\mathrm{H}}. \end{aligned}$$

Hence, with $r^{\mathrm{L}}_{\mathrm{L}}$ and $r^{\mathrm{L}}_{\mathrm{H}}$ both evaluated at $(\hat{\tau}_{\mathrm{H}}, \underline{\tau})$,

$$W'_{\mathrm{L}}(\underline{\tau}) = D\, r^{\mathrm{L}}_{\mathrm{L}} + s_{\mathrm{H}}\, D\,(1+\omega)\, r^{\mathrm{L}}_{\mathrm{H}}.$$

Before signing the marginal return, note that $V_{\mathrm{L}}(\tau_{\mathrm{H}}, \cdot)$ is strictly concave on $\mathcal{A}_{\mathrm{L}}$. Up to a term that does not depend on τ_{L}, $(1-\delta\lambda_\emptyset)\, V_{\mathrm{L}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}})$ equals $r^{\mathrm{L}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) + \delta\lambda\, W_{\mathrm{L}}(\tau_{\mathrm{L}})$. Differentiating the watch value above twice, with $T'_{\mathrm{H}} = s_{\mathrm{H}}$, $r^{\mathrm{L}}_{\mathrm{LL}} = -2/\theta$, $r^{\mathrm{L}}_{\mathrm{LH}} = 1/\theta$, and $r^{\mathrm{L}}_{\mathrm{HH}} = 0$,

$$W''_{\mathrm{L}}(\tau_{\mathrm{L}}) = D\left[r^{\mathrm{L}}_{\mathrm{LL}} + 2 s_{\mathrm{H}}\, r^{\mathrm{L}}_{\mathrm{LH}} + s^2_{\mathrm{H}}\, r^{\mathrm{L}}_{\mathrm{HH}}\right] + \omega D\, s^2_{\mathrm{H}}\, r^{\mathrm{L}}_{\mathrm{HH}} = -\frac{2D\,(1-s_{\mathrm{H}})}{\theta},$$

so that, with $\delta\lambda D = \omega$,

$$(1-\delta\lambda_\emptyset)\,\frac{\partial^2 V_{\mathrm{L}}}{\partial \tau^2_{\mathrm{L}}}(\tau_{\mathrm{H}}, \tau_{\mathrm{L}}) = -\frac{2}{\theta} - \omega\,\frac{2(1-s_{\mathrm{H}})}{\theta} = -\frac{2}{\theta}\big[1+\omega(1-s_{\mathrm{H}})\big] < 0, \qquad \tau_{\mathrm{L}} \in \mathcal{A}_{\mathrm{L}}.$$

Hence, if Player L's marginal return at the floor is strictly negative, the floor, the lower end of $\mathcal{A}_{\mathrm{L}}$, is the unique maximizer of $V_{\mathrm{L}}(\tau_{\mathrm{H}}, \cdot)$ on $\mathcal{A}_{\mathrm{L}}$.

Taken together, in state $\tau_{\mathrm{H}} = \hat{\tau}_{\mathrm{H}}$, inserting $W'_{\mathrm{L}}(\underline{\tau})$ and using $\delta\lambda D = \omega$, with all derivatives of r^{L} evaluated at $(\hat{\tau}_{\mathrm{H}}, \underline{\tau})$,

$$\begin{aligned}(1-\delta\lambda_\emptyset)\,\frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) &= r^{\mathrm{L}}_{\mathrm{L}} + \omega\left[r^{\mathrm{L}}_{\mathrm{L}} + s_{\mathrm{H}}(1+\omega)\, r^{\mathrm{L}}_{\mathrm{H}}\right] \\ &= (1+\omega)\, r^{\mathrm{L}}_{\mathrm{L}} + s_{\mathrm{H}}\,\omega(1+\omega)\, r^{\mathrm{L}}_{\mathrm{H}} \\ &= (1+\omega)\, K(\underline{\tau}),\end{aligned}$$

where

$$K(\underline{\tau}) \equiv r^{\mathrm{L}}_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) + s_{\mathrm{H}}\,\omega\, r^{\mathrm{L}}_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}). \tag{55}$$

Hence, $\frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) < 0$ if and only if $K(\underline{\tau}) < 0$.

Second, I show that $K(\tau^*_{\mathrm{L}}) < 0$, that is, the condition holds at the lower bound of the admissible range for $\underline{\tau}$. It has already been established that, at the affine steady state, Player L's steady-state condition (20) holds with the slope m of Player H's affine strategy:

$$r^{\mathrm{L}}_{\mathrm{L}}(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}}) + \omega m\, r^{\mathrm{L}}_{\mathrm{H}}(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}}) = 0.$$

Compare the two terms of this expression with the two terms of $K(\tau^*_{\mathrm{L}})$. Concerning the first, $r^{\mathrm{L}}_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \tau^*_{\mathrm{L}}) < r^{\mathrm{L}}_{\mathrm{L}}(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}})$, because $\hat{\tau}_{\mathrm{H}}(\tau^*_{\mathrm{L}}) < \tau^*_{\mathrm{H}}$ and $r^{\mathrm{L}}_{\mathrm{LH}} = 1/\theta > 0$. Here $\hat{\tau}_{\mathrm{H}}(\tau^*_{\mathrm{L}})$ is Player H's static best response to τ^*_{L}, which lies below τ^*_{H} by Player H's wedge $w_{\mathrm{H}} > 0$ (Equation (25)). Concerning the second, $s_{\mathrm{H}}\,\omega\, r^{\mathrm{L}}_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}, \tau^*_{\mathrm{L}}) < \omega m\, r^{\mathrm{L}}_{\mathrm{H}}(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}})$, because $s_{\mathrm{H}} < m$ while $r^{\mathrm{L}}_{\mathrm{H}}(\hat{\tau}_{\mathrm{H}}, \tau^*_{\mathrm{L}}) = (\tau^*_{\mathrm{L}} + a_{\mathrm{L}})/\theta = r^{\mathrm{L}}_{\mathrm{H}}(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}}) > 0$. Hence, $K(\tau^*_{\mathrm{L}}) < r^{\mathrm{L}}_{\mathrm{L}}(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}}) + \omega m\, r^{\mathrm{L}}_{\mathrm{H}}(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}}) = 0$.

Third, I establish that $K'(\underline{\tau}) < 0$ for all $\underline{\tau} \geq \tau^*_{\mathrm{L}}$. This then establishes $\frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) < 0$ for all admissible $\underline{\tau} \geq \tau^*_{\mathrm{L}}$. Along $\hat{\tau}_{\mathrm{H}}(\underline{\tau}) = \frac{1}{2}(\theta - a_{\mathrm{H}} + \underline{\tau})$, Player H's steady-state rate moves with the floor at rate $d\hat{\tau}_{\mathrm{H}}/d\underline{\tau} = \frac{1}{2}$. Take the derivative

$$\begin{aligned}K'(\underline{\tau}) &= \underbrace{r^{\mathrm{L}}_{\mathrm{LL}}(\hat{\tau}_{\mathrm{H}}(\underline{\tau}), \underline{\tau})}_{-2/\theta} + \underbrace{r^{\mathrm{L}}_{\mathrm{LH}}(\hat{\tau}_{\mathrm{H}}(\underline{\tau}), \underline{\tau})}_{1/\theta}\cdot\frac{1}{2} + s_{\mathrm{H}}\,\omega\,\underbrace{r^{\mathrm{L}}_{\mathrm{HL}}(\hat{\tau}_{\mathrm{H}}(\underline{\tau}), \underline{\tau})}_{1/\theta} + s_{\mathrm{H}}\,\omega\,\underbrace{r^{\mathrm{L}}_{\mathrm{HH}}(\hat{\tau}_{\mathrm{H}}(\underline{\tau}), \underline{\tau})}_{0}\cdot\frac{1}{2} \\ &= \frac{1}{\theta}\left(\omega s_{\mathrm{H}} - \frac{3}{2}\right) < 0,\end{aligned}$$

where negativity follows from $\omega < 1$ and $s_{\mathrm{H}} < \frac{1}{2}$, so that $\omega s_{\mathrm{H}} < \frac{1}{2}$.

I have thus established that $\frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) < 0$ for all $\underline{\tau} \in [\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$.

Fourth, I show that $\frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\tau_{\mathrm{H}}, \underline{\tau}) < 0$ also for τ_{H} above $\hat{\tau}_{\mathrm{H}}$, that is, that the equilibrium does not rest on Player H being unable to raise its rate above $\hat{\tau}_{\mathrm{H}}$. To see this, return to Player L's marginal return in an arbitrary state $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$,

$$(1 - \delta\lambda_{\emptyset}) \frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\tau_{\mathrm{H}}, \underline{\tau}) = r_{\mathrm{L}}^{\mathrm{L}}(\tau_{\mathrm{H}}, \underline{\tau}) + \delta\lambda\, W_{\mathrm{L}}'(\underline{\tau}).$$

Because Player H's reply to $\underline{\tau}$ is still $\hat{\tau}_{\mathrm{H}}$, whatever its current rate τ_{H}, the marginal change in the watch value, that is, after Player H moved again, still is $W_{\mathrm{L}}'(\underline{\tau}) = D\, r_{\mathrm{L}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) + s_{\mathrm{H}}\, D\, (1 + \omega)\, r_{\mathrm{H}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau})$. Therefore, the marginal return is given by

$$\begin{aligned}
(1 - \delta\lambda_{\emptyset}) \frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\tau_{\mathrm{H}}, \underline{\tau}) &= r_{\mathrm{L}}^{\mathrm{L}}(\tau_{\mathrm{H}}, \underline{\tau}) + \omega\, r_{\mathrm{L}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) + s_{\mathrm{H}}\, \omega(1 + \omega)\, r_{\mathrm{H}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) \\
&= r_{\mathrm{L}}^{\mathrm{L}}(\tau_{\mathrm{H}}, \underline{\tau}) - r_{\mathrm{L}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) + (1 + \omega)\, K(\underline{\tau}) \\
&= \frac{\tau_{\mathrm{H}} - \hat{\tau}_{\mathrm{H}}}{\theta} + (1 + \omega)\, K(\underline{\tau}),
\end{aligned}$$

after plugging in the definition of $r_{\mathrm{L}}^{\mathrm{L}}$. In the previous steps, it was shown that $K(\underline{\tau}) < 0$ over the whole admissible range of $\underline{\tau}$. Hence, $\frac{\partial V_{\mathrm{L}}}{\partial \tau_{\mathrm{L}}}(\tau_{\mathrm{H}}, \underline{\tau}) < 0$ for all τ_{H} below a threshold:

$$\tau_{\mathrm{H}} < B \equiv \hat{\tau}_{\mathrm{H}} - \theta(1 + \omega)\, K(\underline{\tau}),$$

where the threshold B strictly exceeds $\hat{\tau}_{\mathrm{H}}$. Player L's marginal return at the floor is therefore strictly negative in every state $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$ if

$$b < B. \tag{C2}$$

Under (C1) and (C2), by strict concavity, the floor is the unique maximizer of $V_{\mathrm{L}}(\tau_{\mathrm{H}}, \cdot)$ on $\mathcal{A}_{\mathrm{L}}$ in every state $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$: $T_{\mathrm{L}}(\tau_{\mathrm{H}}) = \underline{\tau}$ is Player L's best reply on $\mathcal{A}_{\mathrm{H}}$.

Fifth, it is shown that the threshold B for values of τ_{H}, below which the floor is the unique maximizer, is larger than τ_{H}^*, Player H's rate from the affine steady state, that is, $\hat{\tau}_{\mathrm{H}} - \theta(1 + \omega)K(\underline{\tau}) > \tau_{\mathrm{H}}^*$ for all $\underline{\tau} \in [\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$. This step is not needed for the MPE; it shows that the bound b can be chosen above τ_{H}^*, as stated after the proposition. Define the shorthand

$$\begin{aligned}
G(\underline{\tau}) &\equiv \frac{\tau_{\mathrm{H}}^* - \hat{\tau}_{\mathrm{H}}}{\theta} + (1 + \omega)\, K(\underline{\tau}) \\
&= \frac{\tau_{\mathrm{H}}^* - \hat{\tau}_{\mathrm{H}}}{\theta} + (1 + \omega) \left(r_{\mathrm{L}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) + s_{\mathrm{H}}\, \omega\, r_{\mathrm{H}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \underline{\tau}) \right),
\end{aligned}$$

which by the fourth step is Player L's marginal return at the floor in state $\tau_{\mathrm{H}} = \tau_{\mathrm{H}}^*$, multiplied by $1 - \delta\lambda_{\emptyset}$. I will now show that $G(\tau_{\mathrm{L}}^*) < 0$ and that $G'(\underline{\tau}) < 0$ for $\underline{\tau} \in [\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$. Together they give $G(\underline{\tau}) < 0$, which is equivalent to $\hat{\tau}_{\mathrm{H}} - \theta(1 + \omega)K(\underline{\tau}) > \tau_{\mathrm{H}}^*$, on the same range.

Start by showing $G(\tau_{\mathrm{L}}^*) < 0$. For $\underline{\tau} = \tau_{\mathrm{L}}^*$, $\hat{\tau}_{\mathrm{H}}$ is the static best response to τ_{L}^*, and thus $\tau_{\mathrm{H}}^* - \hat{\tau}_{\mathrm{H}}$ is the wedge $w_{\mathrm{H}} > 0$ of Player H. From both components of $K(\tau_{\mathrm{L}}^*)$, subtract their affine steady-state counterparts to obtain

$$\begin{aligned}
r_{\mathrm{L}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \tau_{\mathrm{L}}^*) - r_{\mathrm{L}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) = \frac{\hat{\tau}_{\mathrm{H}} - \tau_{\mathrm{H}}^*}{\theta} &= -\frac{w_{\mathrm{H}}}{\theta} \\
\implies \quad r_{\mathrm{L}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \tau_{\mathrm{L}}^*) &= r_{\mathrm{L}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) - \frac{w_{\mathrm{H}}}{\theta}
\end{aligned}$$

and $r_{\mathrm{H}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \tau_{\mathrm{L}}^*) - r_{\mathrm{H}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) = 0$. From the steady-state condition (20) at the affine steady state, obtain

$$\begin{aligned} & r_{\mathrm{L}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) = -\omega m\, r_{\mathrm{H}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) \\ \Longrightarrow \quad & r_{\mathrm{L}}^{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}, \tau_{\mathrm{L}}^*) = -\omega m\, r_{\mathrm{H}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) - \frac{w_{\mathrm{H}}}{\theta}. \end{aligned}$$

Plugging this into $G(\tau_{\mathrm{L}}^*)$, obtain

$$\begin{aligned} G(\tau_{\mathrm{L}}^*) &= \frac{w_{\mathrm{H}}}{\theta} + (1+\omega)\left(-\omega m\, r_{\mathrm{H}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) - \frac{w_{\mathrm{H}}}{\theta} + s_{\mathrm{H}}\, \omega\, r_{\mathrm{H}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*)\right) \\ &= -\frac{\omega}{\theta}\, w_{\mathrm{H}} + (1+\omega)\, \omega\, (s_{\mathrm{H}} - m)\, r_{\mathrm{H}}^{\mathrm{L}}(\tau_{\mathrm{H}}^*, \tau_{\mathrm{L}}^*) < 0, \end{aligned}$$

where negativity follows from $s_{\mathrm{H}} < m$ and positivity of all other terms involved.

Next, $G'(\underline{\tau}) < 0$ for $\underline{\tau} \in [\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$. Take the derivative, with $d\hat{\tau}_{\mathrm{H}}/d\underline{\tau} = \frac{1}{2}$,

$$G'(\underline{\tau}) = \frac{-1/2}{\theta} + (1+\omega)\, K'(\underline{\tau}) < 0,$$

because $K'(\underline{\tau}) < 0$ by the third step. This completes the fifth step: for all minimum taxes in the admissible range $[\tau_{\mathrm{L}}^*, \theta - a_{\mathrm{H}})$, Player H's affine steady-state rate τ_{H}^* lies below the threshold B.

Conditions on b and ε. The two parts hold under (C1) and (C2). Since $K(\underline{\tau}) < 0$, the threshold B exceeds $\hat{\tau}_{\mathrm{H}}$, and $\hat{\tau}_{\mathrm{H}} < 1$ by assumption, so there is a b with $\hat{\tau}_{\mathrm{H}} < b < B$ and $b \leq 1$. Given b, there is an $\varepsilon > 0$ that satisfies (C1) and $\underline{\tau} + \varepsilon \leq 1$. For these b and ε, T_{H} is Player H's best reply on $\mathcal{A}_{\mathrm{L}}$ and T_{L} is Player L's best reply on $\mathcal{A}_{\mathrm{H}}$, so no player has a profitable one-shot deviation in any state in $\mathcal{A}_{\mathrm{H}} \times \mathcal{A}_{\mathrm{L}}$, and $(T_{\mathrm{H}}, T_{\mathrm{L}})$ is an MPE of the game with action sets $\mathcal{A}_{\mathrm{H}}$ and $\mathcal{A}_{\mathrm{L}}$. By the fifth step, B also exceeds τ_{H}^*, and $\tau_{\mathrm{H}}^* = k_{\mathrm{H}} + m\, \tau_{\mathrm{L}}^* < k_{\mathrm{H}} + m \leq 1$ because $\tau_{\mathrm{L}}^* \leq \underline{\tau} < 1$, so b can be chosen above τ_{H}^*.

Convergence to the steady state. First, note that, given the strategies described, the players remain in the state space $\mathcal{A}_{\mathrm{H}} \times \mathcal{A}_{\mathrm{L}}$ for any initial tax profile within it. If Player L moves, given a tax rate $\tau_{\mathrm{H}} \in \mathcal{A}_{\mathrm{H}}$, Player L sets its tax rate to $\underline{\tau}$ by definition. If Player H moves against a tax rate of Player L of $\tau_{\mathrm{L}} \in \mathcal{A}_{\mathrm{L}}$, Player H responds with $T_{\mathrm{H}}(\tau_{\mathrm{L}})$, which is in $\mathcal{A}_{\mathrm{H}}$ by (C1). If nobody moves, the profile persists.

I now show that the corner steady state $(\hat{\tau}_{\mathrm{H}}, \underline{\tau})$ is absorbing on this state space. It is a steady state because $T_{\mathrm{L}}(\hat{\tau}_{\mathrm{H}}) = \underline{\tau}$ and $T_{\mathrm{H}}(\underline{\tau}) = \hat{\tau}_{\mathrm{H}}$, and it is the only one because Player L's strategy is the floor in every state. To see that it is absorbing, notice that, once Player L moved, its tax rate is $\underline{\tau}$ and every later move keeps it there.

Once Player L moved, Player H plays $T_{\mathrm{H}}(\underline{\tau}) = \hat{\tau}_{\mathrm{H}}$ at its first move after that, and the game has reached its steady state. Both a move by Player L and a subsequent move by Player H happen with probability 1 in finite time because $\lambda > 0$: in every period each player moves with probability λ, so the probability that a given player has not moved within t periods is $(1-\lambda)^t$, which vanishes as t grows. This completes the proof. □

B Calibration Details

This appendix details the calibration summarized in Section 5.1.

B.1 Discount factor δ

The yearly discount factor δ is calibrated from the real interest rate according to

$$\delta = \frac{1}{1+\overline{r}} \qquad r_t = \frac{1+i_t}{1+\pi_t} - 1$$

where the real interest rate r_t is computed based on i_t, the nominal interest rate (computed from Bundesbank data on 10-year German federal securities (Deutsche Bundesbank 2026)), as well as π_t, the consumer price index (CPI) inflation from the German statistical office (Statistisches Bundesamt 2026).[17] The average real interest rate $\overline{r}$ is computed for the years 1995–2001, the period preceding 2002, the base year used by Buettner and Poehnlein (2024). This results in values of $\overline{r} = 4.0\%$ and $\delta = 0.962$.[18]

B.2 Data: The *Realsteuervergleich*

The remaining parameters are calibrated based on the same data source, the *Realsteuervergleich* (comparison of non-personal taxes) for the German federal state of Mecklenburg-Vorpommern (MV) for the years 2000–2007.[19] Buettner and Poehnlein (2024) use the same data source for seven German federal states, including MV, for the years 1999–2008.[20] I rely on two of the reported variables. First, I use the business tax multiplier (*Hebesatz*, henceforth multiplier), a factor chosen by the respective municipality that translates into the effective statutory rate of the local business tax (*Gewerbesteuer*) according to the formula

$$\tau = \frac{M\,h'}{1+M\,h'}, \qquad h' = \frac{h}{100}, \tag{56}$$

where h is the multiplier, $M = 5\%$ the federal basic rate (*Steuermesszahl*) in force until 2007. To understand the formula, it is helpful to know that the German business tax is levied in two steps: the tax office multiplies a firm's business profit by the federal basic rate M to obtain the assessed tax base (*Steuermessbetrag*, henceforth assessed base), to which the municipality then applies its multiplier to obtain the tax due. The denominator reflects that the business tax was

[17] For Germany, average inflation measured by the GDP-deflator (Eurostat 2026) over the same time window is 0.7% against 1.4% as measured by the CPI. Using it leads to a slightly lower discount factor of $\delta = 0.954$ but does not affect the results qualitatively.

[18] The average is used to smooth potential short term fluctuations in yearly figures. It starts in 1995 to exclude the high interest rates from the years following German reunification. Using the same method, the real interest rate for 2002 was 3.4%, resulting in $\delta = 0.967$ compared to $\delta = 0.962$ for the baseline specification.

[19] Data older than 2000 are available as PDF report only. In 2008, a business tax reform changed the way in which the rate chosen by a municipality translates into the effective tax rate (see below), so that tax rates before and after 2008 are not comparable.

[20] The German federal states are Schleswig-Holstein, Niedersachsen, Mecklenburg-Vorpommern, Brandenburg, Sachsen-Anhalt, Thüringen, and Sachsen.

deductible from its own base until 2007. A multiplier of 200, the minimum value introduced in 2004, thus corresponds to a business tax rate of $\tau = 9.09\%$.

Second, I use the revenue-implied tax base (*Grundbetrag*, henceforth implied base), which MV's regional statistical office (*Landesamt für innere Verwaltung, Statistisches Amt*) computes as the municipality's business-tax receipts divided by its h' (Landesamt für innere Verwaltung Mecklenburg-Vorpommern, Statistisches Amt 2026). It thus approximates the sum of the assessed bases of the firms located in the municipality, in euros.[21] As in Thunecke et al. (2026), I treat the implied base as the sum of the assessed bases.[22]

The data was collected for the single years from the website of MV's regional statistical office and harmonized following the procedure sketched in Buettner and Poehnlein (2024) to take into account communal mergers and separations as documented by the German Federal Statistical Office.[23]

B.3 Probability of moving λ

The probability of having the opportunity to move λ is computed from the multipliers of the municipalities as they report in the *Realsteuervergleich* for the years 2000–2007, before the harmonization described above.[24] Over this period, 1,023 municipalities report in at least two years. I count the changes of the multiplier relative to the previous year in 2001–2007 (334) and divide them by the number of possible year-on-year comparisons in the sample (6,317).[25] The resulting value for the probability to have the opportunity to move is $\lambda = 0.053$, reflecting one change every 19 years on average.

Note that the number of observed changes is only an imperfect measure for the probability

[21] Values for the municipalities' receipts for the year 2000 are converted from Deutsche Mark (DM) to euros using the conversion factor of 1.95583 DM/euro.

[22] The two measures would coincide if tax assessment and collection happened in the same year. They generally do not, since the receipts behind the implied base are a mix of advance payments based on earlier assessments and back payments and refunds for other years (Landesamt für Statistik Niedersachsen 2026). Thunecke et al. (2026) never state this identity explicitly but report negative values for their local business tax base in the article's online appendix (Table A.4). This rules out a sum of assessed bases, which cannot be negative since a loss yields an assessed base of zero, but is consistent with an implied base, which turns negative when refunds exceed receipts.

[23] Following them, I fix 2002 as base year — two years before the minimum tax was introduced in 2004 and one year before it was announced — and "freeze" the existing boundaries of the 979 municipalities of MV in 2002. Municipalities that merged after 2002 are reported as continuing to exist and inherit the new municipality's tax rate. The merged municipality's implied base is split back onto its 2002 predecessors, using their shares in the year before the merger. If agreed mergers are followed by a transition period, the respective municipalities continue to have their own data points until the merger factually takes place. No municipality was divided between 2002–2007. I abstracted from one minor partial transfer during the period. The values for municipalities merging before 2002 are dropped from the sample. Coverage after harmonization is thus 975 municipalities in 2000, 978 in 2001, and 979 from 2002 on.

[24] The harmonized panel carries each merged municipality forward with the tax rate of its successor. The merger itself then appears as a change in the tax rate of every predecessor whose rate differs from the successor's, and every later change in the successor's tax rate is counted once for each of its predecessors.

[25] The number of comparisons is the total count of observations from 2000–2007 (7,340) minus the number of municipalities in the sample (1,023), since a municipality's first observation has no previous year to compare with.

of having the *opportunity* to move. In the model, a player given the opportunity to move twice without any other player moving in between would not change its tax rate, the second opportunity thus staying unobserved. However, a municipality in MV has on average 26 other municipalities within a radius of 15 km, the distance Buettner and Poehnlein (2024) use to define the neighbors of a tax haven. With this many competitors, the probability that none of them moves between two opportunities of the same municipality is negligible.

B.4 Tax base sensitivity θ

The tax base sensitivity to the tax differential θ is estimated based on the harmonized panel for MV over the 2000–2007 time window. The estimating equation is derived from a logit model of discrete choice, standard in the discrete-choice literature, for the firm's location between N municipalities. After deriving the number of firms x_{it} locating in municipality i at time t in this model, I show that the baseline choice model from Section 2.1 is a linear approximation of this model for two players with identical characteristics, around equal tax rates. Using the logit model for the calibration has the advantage that it allows for any finite number of municipalities and is thus directly applicable to the data.

The preferences of firm k for location i at time t are given by

$$U_{kit} = -\tau_{it} + \zeta_i + \xi_{it} + \mu\,\nu_{kit},$$

where ν_{kit} is a randomly distributed taste parameter of firm k, whose dispersion is scaled by $\mu > 0$. Municipality-specific characteristics are collected in ζ_i, and ξ_{it} is a municipality-year shock to location attractiveness common to all firms.

Assuming that the taste parameters ν_{kit} are independently Gumbel distributed across firms and locations (Greene 2003, Ch. 21, sec. 7) makes it possible to write

$$x_{it} = N\,\mathbb{P}\left[U_{kit} \geq U_{kjt}\ \forall j\right] = N\,\frac{\exp\left((-\tau_{it} + \zeta_i + \xi_{it})/\mu\right)}{\sum_{j=1}^{N}\exp\left((-\tau_{jt} + \zeta_j + \xi_{jt})/\mu\right)}, \tag{57}$$

where, appealing to a law of large numbers, I identify the probability that municipality i offers firm k the highest utility at time t with i's realized share of firms, and the total mass of firms is N, one per municipality as in Section 2.1.

For two players with identical characteristics ($N = 2$, $\zeta_1 = \zeta_2$, $\xi_{1t} = \xi_{2t}$), Equation (57) gives

$$x_{1t} = \frac{2}{1 + e^{-d_t}}, \qquad d_t \equiv \frac{\tau_{2t} - \tau_{1t}}{\mu}.$$

Since the logistic function $1/(1 + e^{-d})$ has value $\frac{1}{2}$ and slope $\frac{1}{4}$ at $d = 0$, a first-order expansion around equal tax rates ($d = 0$) yields

$$x_{1t} \approx 1 + \frac{\tau_{2t} - \tau_{1t}}{2\mu},$$

the tax base of Section 2.1 with $\theta \equiv 2\mu$. Taking logs of Equation (57), I obtain

$$\ln x_{it} = \ln N - \frac{1}{\mu}\tau_{it} + \underbrace{\frac{\zeta_i}{\mu}}_{\kappa_i} + \tilde{\gamma}_t + \underbrace{\frac{\xi_{it}}{\mu}}_{\varepsilon_{it}}, \tag{58}$$

where time fixed effects $\tilde{\gamma}_t \equiv -\ln \sum_{j=1}^{N} \exp\left((-\tau_{jt} + \zeta_j + \xi_{jt})/\mu\right)$ absorb the logit equation's denominator, which is common to all municipalities in year t, while municipality fixed effects κ_i absorb municipality-specific characteristics driving firms' location decisions, and the error term ε_{it} is the scaled municipality-year shock. The latter also absorbs deviations of profits per firm from the homogeneous pre-tax profits $\bar{\pi}_t$, assumed below, and measurement error in the implied base. It is assumed to be uncorrelated with the municipality's tax rates in all years (strict exogeneity). This assumption is not tested: if municipalities adjust their multiplier in response to shocks to their tax base, $\hat{b}$ captures the co-movement of tax rate and tax base within municipalities, not the causal effect of the tax rate on the tax base, a limitation of the current approach.

The model variable x_{it} denotes the unobserved number of firms locating in municipality i. I will now derive its relation to the implied base I obtain from the data, remembering that I previously assumed it to be identical to the assessed base. To do so, I assume homogeneous pre-tax profits per firm $\bar{\pi}_t$, making profit at the municipality level $\Pi_{it} = \bar{\pi}_t \, x_{it}$.

To obtain the municipality's assessed base S_{it}, one must multiply the federal basic rate M by profit Π_{it} net of total tax liabilities G_{it} (in Germany, these could be deducted from profits before 2008), resulting in $S_{it} = M(\Pi_{it} - G_{it})$. Total tax revenue G_{it} for the municipality in turn is computed by applying the multiplier h' to the assessed base. Putting these together, I obtain

$$G_{it} = h'_{it} \, M(\Pi_{it} - G_{it}) \iff G_{it} = \underbrace{\frac{h'_{it} \, M}{1 + h'_{it} \, M}}_{\equiv \tau_{it}} \Pi_{it}.$$

Hence, I can express the assessed base S_{it} in terms of the number of firms x_{it} locating in i by rearranging

$$S_{it} = M(\Pi_{it} - \tau_{it}\Pi_{it}) = (1 - \tau_{it}) M \, \bar{\pi}_t \, x_{it} \quad \iff \quad \frac{S_{it}}{1 - \tau_{it}} = M \, \bar{\pi}_t \, x_{it}.$$

Taking logs results in

$$\ln x_{it} = \ln\left(\frac{S_{it}}{1 - \tau_{it}}\right) - \ln\left(M \, \bar{\pi}_t\right)$$

Substituting this into Equation (58), I obtain

$$\ln\left(\frac{S_{it}}{1 - \tau_{it}}\right) = b \, \tau_{it} + \kappa_i + \gamma_t + \varepsilon_{it}, \qquad b = -\frac{1}{\mu}, \tag{59}$$

where the time fixed effects $\gamma_t \equiv \tilde{\gamma}_t + \ln(NM\bar{\pi}_t)$ absorb the homogeneous, time-varying pre-tax profit per firm $\bar{\pi}_t$ as well as the constants N and M.

I thus use as dependent variable the natural logarithm of the assessed base, measured as implied base in the data, divided by 1 minus the effective tax rate, and as independent variable the effective tax rate. A regression with municipality and time fixed effects gives $b = -1/\mu$ as the coefficient of the tax rate τ_{it}.

Because I take the natural logarithm of the implied base, I drop 247 observations whose implied base is non-positive or missing, leaving a total of 7,580 observations. Because of the

municipality fixed effects, the variation exploited for the estimation stems only from the 313 municipalities that changed their multiplier within the time window, with 476 rate changes in total.

As I have shown before, the two-player benchmark model is a linearization of the logit model used to derive the regression equation with $\theta \equiv 2\mu$. Therefore, I calibrate $\theta = -2/b$. The estimation results are on display in Table 3 with an estimate for the pre-reform period 2000–2003 reported as robustness check.

Table 3: Estimation of the tax base sensitivity θ

	Baseline 2000–2007	Robustness 2000–2003
Coefficient $\hat{b}$ on τ_{it}	−8.83	−10.35
	(2.64)	(4.64)
Implied $\hat{\mu} = -1/\hat{b}$	0.113	0.097
Implied $\hat{\theta} = 2\hat{\mu}$	0.227	0.193
95% confidence interval for θ	[0.143, 0.547]	[0.103, 1.582]
Municipality and year fixed effects	yes	yes
Observations	7,580	3,739
Municipalities	979	979
Municipalities changing their multiplier	313	92
Multiplier changes	476	100

Notes: Dependent variable $\ln(S_{it}/(1-\tau_{it}))$, with S_{it} measured by the implied base. Standard errors clustered at the municipality level in parentheses. The confidence interval for θ maps the endpoints of the 95% interval for b through $\theta = -2/b$.

The standard errors are clustered at the municipality level to account for potential serial correlation of the error terms.

B.5 Non-tax benefits a_i

The non-tax benefits a_{H} and a_{L} are calibrated such that the steady-state tax profile $(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}})$ of the affine MPE matches the average tax rates observed in 2002 for two groups of municipalities: tax havens and high-tax municipalities. The group definitions follow Buettner and Poehnlein (2024) and are given in Section 5.1. In MV, only 24 of the 970 municipalities that are not tax havens exceed the high-tax threshold of 350; the 85th percentile of their 2002 multiplier is 300.

Table 4 summarizes the two groups in 2002 and compares them with the seven-state sample of Buettner and Poehnlein (2024).

The non-tax benefits a_i are then backed out from $(\tau^*_{\mathrm{H}}, \tau^*_{\mathrm{L}})$ by inverting the closed form of the steady-state tax rates in Equation (19), evaluated at the common slope $\hat{m}_{-i} = m$ of the affine MPE. Since neither ω nor m depends on a_{H} or a_{L}, each player's equation can be solved for its own a_i:

$$a_i = \frac{\theta + \tau^*_{-i} - (2 - \omega m)\,\tau^*_i}{1 - \omega m}, \qquad i = \mathrm{H}, \mathrm{L}.$$

Table 4: Tax havens and high-tax municipalities in 2002, and the non-tax benefits a_i

	Tax havens	High-tax municipalities		
		within 15 km	other	all
	(Player L)	(Player H)		
Mecklenburg-Vorpommern				
Municipalities	9	6	18	24
Mean multiplier	138.9	395.8	393.6	394.2
Mean tax rate τ	6.47%	16.52%	16.44%	16.46%
Non-tax benefit a_i	0.350	−0.005		
Buettner and Poehnlein (2024)				
Municipalities	30			782
Mean multiplier	126			383
Implied tax rate τ	5.9%			16.1%

Notes: Groups are defined by the 2002 multiplier: tax havens below 200, high-tax municipalities above 350. "Within 15 km" marks high-tax municipalities within 15 km of a tax haven. For MV, τ is the group mean of the effective tax rates implied by the multiplier. The values for Buettner and Poehnlein (2024) refer to their seven-state sample; their tax rates are not reported in the paper but implied by the mean multiplier they report: for tax havens in 2002, and for high-tax municipalities the value of about 383 at which they evaluate their estimate.

This requires the calibrated value of θ and the slope m, which depends only on δ and λ. It is the smaller root of the slope equation (22) of Proposition 3, which lies in $(0, 1/2)$; the other real root exceeds one (Appendix A.4). The slope equation depends on δ and λ only through $\omega = \delta\lambda/(1 - \delta + \delta\lambda)$ (Equation (12)), which gives $\omega = 0.570$ and $m = 0.370$ for the calibrated δ and λ.



Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim
ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.