Leader-Driven Collusion*

Juan Ortner†

Boston University

May 13, 2025

Abstract

Cartels often form among heterogeneous firms, with large firms acting as leaders and coordinating collusion. Motivated by these instances, I study optimal repeated-game equilibria in a market with a cartel leader and several less patient competitors. Optimal collusive equilibria may involve an initial phase with gradually increasing prices, followed by a stationary phase in which prices are set at the monopoly level and market shares remain constant. These dynamics align with the behavior of several detected cartels, which initially exhibited gradual price increases. Methodologically, I use Lagrangian techniques to characterize Pareto optimal repeated-game equilibria for fixed (and possibly heterogeneous) discount factors.

KEYWORDS: repeated games, cartels, collusion, optimal equilibria, heterogeneous discount factors, Lagrangian methods.

*I thank Joe Harrington, Marc Rysman and Alex Wolitzky for helpful comments.

†Contact information: jortner@bu.edu.

1 Introduction

Cartels often form among heterogeneous firms, with large firms acting as leaders and coordinating collusion. For example, Connor (2007) documents that in the lysine price-fixing conspiracy, executives from ADM – the largest US producer at the time – played an instrumental role in organizing meetings with their competitors and initiating the collusive agreement. Similar dynamics took place in the citric acid conspiracy and in several of the vitamins cartels, where the largest firms took the lead in organizing collusion (Connor, 2007). In the electrical equipment price-fixing conspiracy, the two dominant firms – General Electric and Westinghouse – were key in setting up the collusive scheme (Watkins, 1961).

Motivated by these instances of collusion, this paper studies optimal repeated-game equilibria in a market with a cartel leader and several less patient competitors. I show that optimal equilibria may involve an initial phase with gradually increasing prices. Moreover, play under optimal equilibria converges in finite time to a stationary profile in which prices are set at the monopoly level and market shares remain constant. Methodologically, I use Lagrangian techniques to characterize Pareto optimal repeated-game equilibria for fixed (and possibly heterogeneous) discount factors.

I study an infinitely repeated Bertrand oligopoly game in a market with a cartel leader and several competitors. The cartel leader is (weakly) more patient, while all other firms share the same, weakly higher discount rate. One possible interpretation is that the cartel leader is the largest firm in the industry, and faces a lower cost of capital. For simplicity, I assume that all firms have the same constant marginal cost, normalized to zero.[1] Firms face a unit mass of buyers with reservation price r. In each period, the firm with the lowest price serves the entire market. In case of ties, I allow firms to jointly determine market shares.

My first set of results concern the leader-optimal equilibrium. As an initial step, I characterize the leader-optimal equilibrium under the constraint that, along the path of play, prices are constant at all periods. I show that the optimal such equilibrium features constant mar-

[1]See the conclusion for a discussion on how the results in this paper extend to markets in which the more patient firm also has a cost advantage.

ket shares and prices at the monopoly level. I then characterize the overall leader-optimal equilibrium and show that it features two phases. In the first phase, which lasts until a finite date t^*, prices increase gradually and the leader captures the entire market. In the second phase, which starts at date t^*, play is equal to the leader-optimal stationary equilibrium: prices are set at the monopoly level and firms earn constant market shares.

Leader-optimal equilibria are consistent with the behavior of several detected cartels, including the citric acid and lysine cartels, as well as several of the vitamin conspiracies, which featured an initial phase with gradually increasing prices (Connor, 2007, Harrington, 2008, Marshall and Marx, 2012). In some of these cartels (e.g., beta carotene, vitamin B2 and biotin), the larger firms agreed to cede market share over time to smaller firms to entice them into participating in the collusive agreement, and to incentivize them not to defect from it (Connor, 2007). Previous work by Harrington (2004, 2005) showed that gradually increasing prices may arise when a newly formed cartel seeks to avoid detection by an antitrust authority. Instead, the price dynamics in my model are driven by maximization of the leader's profits. Indeed, the leader-optimal equilibrium features backloading of incentives, as in many dynamic principal-agent models (Ray, 2002). This backloading structure gives rise to gradually increasing prices.

My second set of results concern Pareto efficient equilibria that maximize the weighted sum of the leader's and competitors' profits. When the weight on the leader is sufficiently large, Pareto efficient equilibria retain the same key properties of leader-optimal equilibria: there is an initial phase with gradually increasing prices and with the leader capturing the entire market, and a second phase in which play is given by the leader-optimal equilibrium with stationary prices. When the weight on the leader is small, Pareto efficient equilibria also feature two phases: an initial phase in which the less patient firms capture the entire market and prices are set at the monopoly level, and a second phase in which play is given by the leader-optimal stationary equilibrium.

I also study an extension with stochastic demand, in the spirit of Rotemberg and Saloner (1986). I focus on the equilibrium that maximizes the leader's profits, and show that the key properties of the leader-optimal equilibrium with constant demand extend to this setting: the leader-optimal equilibrium also features an initial phase with gradually increasing prices, and a second phase with stationary prices and stationary market shares. In contrast to the stationary equilibrium analyzed by Rotemberg and Saloner (1986), the stationary phase of the leader-optimal equilibrium features market shares of fringe firms that vary pro-cyclically: fringe firms capture a larger share of the market during booms, which deters them from defecting.

From a methodological standpoint, I cast the problem of finding Pareto optimal equilibria as that of solving a Lagrangian with countably infinite constraints, corresponding to the firms' incentive compatibility constraints and the feasibility constraints on prices and market shares. I show that optimal Lagrange multipliers follow tractable recursive equations, which can be solved in closed form. These optimal multipliers, together with the complementary slackness conditions, can be used to back out optimal equilibrium play. Moreover, because the dual problem of finding multipliers that minimize the Lagrangian provides an upper bound on attainable payoffs, the methods in this paper could potentially be used more broadly to compute an upper bound to the set of equilibrium payoffs.

The paper is structured as follows. The rest of the introduction discusses related literature. Section 2 introduces the model. Section 3 derives preliminary results. Section 4 studies leader-optimal equilibria. Section 5 characterizes Pareto optimal equilibria. Section 6 generalizes the model to settings with stochastic demand. Section 7 concludes. All proofs are in the Appendix.

Related literature. This paper relates primarily to the literature that studies collusion in repeated oligopoly games (e.g., Abreu, 1986, Abreu et al., 1986, Fudenberg et al., 1994, Athey and Bagwell, 2001, Athey et al., 2004).[2] The closest paper within this literature is Obara and Zincenko (2017), who also study a repeated Bertrand oligopoly game with heterogeneous

[2]For more recent contributions, see Skrzypacz and Hopenhayn (2004), Harrington and Skrzypacz (2011), Bernheim and Madsen (2017), Sugaya and Wolitzky (2018), Chassang and Ortner (2019), Kawai et al. (2022), Ortner et al. (2024), Sugaya and Wolitzky (2024).

discounting. The paper has two main results. First, they show that collusion is possible if and only if the average discount factor across firms is larger than $(n-1)/n$.[3] Second, they characterize optimal cartel equilibria within the class of equilibria featuring prices equal to the monopoly price at all periods. They don't characterize the set of Pareto optimal equilibria, and don't consider settings with stochastic demand. Moreover, the two papers also differ in terms of the methods, as Obara and Zincenko (2017) don't use Lagrangian techniques.[4]

The current paper also relates to a set of papers studying repeated games with unequal discounting (Lehrer and Pauzner, 1999, Chen and Takahashi, 2012, Sugaya, 2015).[5] These papers focus on characterizing the limiting equilibrium payoff set as players' become infinitely patient. In contrast, I characterize optimal equilibria for fixed discount factors.

In terms of methods, the closest paper is Marcet and Marimon (2019), who study optimal dynamic contracts in settings with forward-looking constraints. Marcet and Marimon (2019) show that, under suitable conditions, the Lagrangian follows a recursive functional equation that characterizes payoffs under the optimal contract. Martimort (1999) also uses Lagrangian methods to characterize optimal collusion-proof contracts in a dynamic principal-monitor-agent setting.

Finally, as I highlighted above, the methods in this paper could potentially be applied more broadly to compute the set of equilibrium payoffs in repeated games. This relates my work to papers that develop methods to compute the set of repeated-game equilibrium payoffs (Abreu et al., 1990, Judd et al., 2003, Abreu and Sannikov, 2014, Abreu et al., 2020).

[3] In a previous paper, Harrington (1989) established the same result, restricting attention to stationary equilibria.

[4] In Section 5 I provide a more detailed discussion of the results in Obara and Zincenko (2017).

[5] See also Dasgupta and Ghosh (2022), who introduce the notion of *self-accessibility* to study equilibria in repeated games with unequal discounting. Mailath and Zemsky (1991) study optimal static collusion among firms with heterogeneous costs.

2 Model

Stage game. Consider a market with firms $N = \{1, ..., n\}$ producing a homogenous good and competing in prices. All firms share the same constant marginal cost c, which I normalize to $c = 0$.[6] There is a unit mass of buyers with reservation price $r = 1$. If $\mathbf{p} = (p_i)_{i \in N}$ are the prices chosen by firms in N, market shares are $\mathbf{x} = (x_i)_{i \in N}$, with $x_i = 1$ if $p_i < p_j$ for all $j \neq i$ and $p_i \leq r = 1$, $x_i = 0$ if there exists $j \neq i$ with $p_j < p_i$, or if $p_i > r$. In case of ties, I allow firms to jointly determine the allocation. Formally, each firm $i \in N$ chooses a number $\rho_i \in [0, 1]$ along with its price. In case of ties, the market share x_i of a firm i that posted the lowest price is $x_i = \frac{\rho_i}{\sum_{j:p_j = \min_k p_k} \rho_j}$, with the convention that $x_i = 1/|\{j : p_j = \min_k p_k\}|$ if $\rho_j = 0$ for all j with $p_j = \min_k p_k$.[7] Firm i's profits are $x_i p_i$.

Repeated game. Firms play the stage game described above at all periods $t \in \mathbb{N}_0$. Each firm i has a discount factor $\delta_i \in (0, 1)$, with $\delta_i = \delta$ for all $i \neq 1$ and $\delta_1 \geq \delta$. That is, firm 1 is (weakly) more patient than the other firms, and all other firms are homogenous. I sometimes refer to firm 1 as the leader, and firms $i \neq 1$ as fringe firms. One interpretation is that the leader is a large firm, facing a lower cost of capital, and fringe firms are smaller competitors. I maintain the following assumption throughout the paper.

Assumption 1. $\delta_1 + (n - 1)\delta > n - 1$.

Assumption 1 states that firms' average discount factor is larger than $(n - 1)/n$. This assumption is necessary for there to exist a collusive equilibrium: Obara and Zincenko (2017) show that no collusion is possible if the average discount factor is smaller than $(n - 1)/n$.

For each $s \in \mathbb{N}_0$, let $(\mathbf{p}_s, \rho_s) = (p_{i,s}, \rho_{i,s})_{i \in N}$ denote firms' actions in period s, and let $\mathbf{x}_s = (x_{i,s})_{i \in N}$ denote firms' market shares. The history at time t is $h_t = (\mathbf{p}_s, \rho_s)_{s<t}$. Let H_t denote the set of all period t histories and $H = \cup_{t \geq 0} H_t$ the set of all histories. A (pure) strategy σ_i for firm i is a mapping $\sigma_i : h_t \mapsto (p_{i,t}, \rho_{i,t})$. For each strategy profile $\sigma = (\sigma_i)_{i \in N}$,

[6] See the conclusion for a discussion of settings in which firms have heterogeneous costs.

[7] Obara and Zincenko (2017) and Chassang and Ortner (2019) assume a similar endogenous tie-breaking technology.

each player i and each history $h \in H$, let $V_i(\sigma|h)$ denote i's continuation payoff under σ at history h (evaluated using i's discount factor), and let $V_i(\sigma) = V_i(\sigma|h_0)$ denote i's payoff under σ at the start of the game.

Let Σ_{p} denote the set of pure strategy subgame perfect Nash equilibria (SPNE). The goal of the paper is to characterize Pareto optimal equilibria in Σ_{p}. In particular, for each $\alpha \in [0,1]$, I solve

$$\overline{V}(\alpha) \equiv \sup_{\sigma \in \Sigma_{\mathsf{p}}} \alpha V_1(\sigma) + (1-\alpha) \sum_{i \neq 1} V_i(\sigma). \tag{1}$$

That is, I focus on Pareto optimal equilibria that give the same weight to all fringe firms. Section 4 characterizes the solution to (1) for $\alpha = 1$; i.e., leader-optimal equilibria. Section 5 characterizes the solution to (1) for all $\alpha \in [0,1]$.

3 Preliminaries

This section derives some preliminary results that will be useful for the analysis of Pareto optimal equilibria. Say that a sequence of prices and market shares $(p_t, \mathbf{x}_t)_{t=0}^{\infty}$ is *implementable* if there exists $\sigma \in \Sigma_{\mathsf{p}}$ with the property that, at each period t, p_t and $\mathbf{x}_t = (x_{i,t})_{i \in N}$ are the lowest price and market shares on the equilibrium path under σ.

Lemma 1. *A sequence of prices and market shares $(p_t, \mathbf{x}_t)$ is implementable if and only if*

$$\forall i, \forall t, \quad (1 - x_{i,t}) p_t \leq \sum_{s=t+1}^{\infty} \delta_i^{s-t} x_{i,s} p_s. \tag{IC}$$

In words, a sequence $(p_t, \mathbf{x}_t)$ is implementable if and only if firms don't have an incentive to undercut the lowest price at any period.

For each $\sigma \in \Sigma_{\mathsf{p}}$, $h \in H$ and $i \in N$, let $\pi_i(\sigma|h)$ denote firm i's flow profits under σ at history h. The next Lemma shows that, when it comes to aggregate cartel profits, it is without loss to focus on equilibria under which all fringe firms obtain the same flow profits at every history.

Lemma 2. *For every $\sigma \in \Sigma_p$, there exists $\widehat{\sigma} \in \Sigma_p$ such that:*

(i) $V_1(\widehat{\sigma}) = V_1(\sigma)$ *and* $\sum_{i\neq 1} V_i(\widehat{\sigma}) = \sum_{i\neq 1} V_i(\sigma)$*, and*

(ii) for all $i, j \neq 1$ *and all histories* $h \in H$*,* $\pi_i(\widehat{\sigma}|h) = \pi_j(\widehat{\sigma}|h)$*.*

By Lemma 2, when looking for equilibria that solve (1), it is without loss to focus on equilibria under which all fringe firms obtain the same flow profits at every history.

4 Leader-Optimal Equilibria

This section characterizes the solution to (1) for $\alpha = 1$; i.e., firm 1's optimal equilibrium. Section 5 characterizes the solution to (1) for all $\alpha \in [0, 1]$.

As an initial step towards characterizing firm 1's optimal equilibrium, I characterize firm 1's optimal equilibrium under the constraint that prices are stationary on the equilibrium path; i.e., along the path of play, $\min_{i\in N} p_{i,t} = p \in [0, 1]$ for all t. Then, I characterize firm 1's overall optimal equilibrium. As I show below, under equilibria that solve (1), equilibrium play converges in finite time to play under the leader-optimal stationary equilibrium.

Let $\Sigma_{\mathsf{p}}^{\mathsf{st}} \subset \Sigma_{\mathsf{p}}$ denote the set of pure strategy equilibria with stationary prices. The problem of finding firm 1's optimal equilibrium with stationary prices is $\overline{V}_1^{\mathsf{st}} \equiv \sup_{\sigma\in\Sigma_{\mathsf{p}}^{\mathsf{st}}} V_1(\sigma)$, while the problem of finding firm 1's optimal equilibrium is $\overline{V}_1 \equiv \sup_{\sigma\in\Sigma_{\mathsf{p}}} V_1(\sigma)$.

4.1 Leader-Optimal Equilibria with Stationary Prices

The following result characterizes firm 1's optimal equilibrium with stationary prices.

Proposition 1. *There exists a leader-optimal equilibrium with stationary prices in which, at every period along the path of play, the price is equal to the reservation price of 1, each fringe firm earns a market share of* $1 - \delta$*, and firm 1 earns a market share of* $1 - (n - 1)(1 - \delta)$*.*

Firm 1's payoff under any leader-optimal equilibrium with stationary prices is

$$\overline{V}_1^{st} = \frac{1 - (n - 1)(1 - \delta)}{1 - \delta_1}.$$

Firm 1's optimal equilibrium with stationary prices features stationary market shares: each fringe firm sells $1-\delta$ of the market each period, with firm 1 selling the remaining $1-(n-1)(1-\delta)$ of the market. To see why market shares take this form, suppose that prices are constant and equal to $p \in (0,1]$ at each period. To satisfy the incentive compatibility constraints of fringe firms, for all t and all $i \neq 1$, it must be that

$$(1-x_{i,t})p \leq \sum_{s=t+1}^{\infty} \delta^{s-t} x_{i,s} p,$$

where $x_{i,s} \in [0,1]$ denotes i's on path market share at time s. Assume that these incentive constraints bind at every period, so that, for all t and $i \neq 1$, $1-x_{i,t} = \sum_{s=t+1}^{\infty} \delta^{s-t} x_{i,s}$. Hence, for all t and all $i \neq 1$,

$$1-x_{i,t} = \delta x_{i,t+1} + \sum_{s=t+2}^{\infty} \delta^{s-t} x_{i,s} = \delta x_{i,t+1} + \delta(1-x_{i,t+1}) = \delta,$$

and so $x_{i,t} = 1-\delta$ for all $i \neq 1$ and all t.

Finally, note that firm 1's IC constraints are satisfied with slack under the scheme in Proposition 1. Let $x^{\mathsf{st}}_{1,t} = 1-(n-1)(1-\delta)$ and $p^{\mathsf{st}}_t = 1$ denote, respectively, firm 1's market share and the price under the scheme in Proposition 1. Then, for all t,

$$(1-x^{\mathsf{st}}_{1,t})p^{\mathsf{st}}_t = (n-1)(1-\delta) < \frac{\delta_1}{1-\delta_1}(1-(n-1)(1-\delta)) = \sum_{s=t+1}^{\infty} \delta_1^{s-t} x^{\mathsf{st}}_{1,s} p^{\mathsf{st}}_s,$$

where the strict inequality follows from Assumption 1.

4.2 Leader-Optimal Equilibria

I now solve for firm 1's overall optimal equilibrium. Define

$$t^* \equiv \inf \left\{ t \in \mathbb{N}_0 : \sum_{\tau=0}^{t} \delta^{t-\tau} \delta_1^{\tau} \geq \delta_1^t (n-1) \right\}. \tag{2}$$

Note that $t^* > 0$ if and only if there are three or more firms (i.e., $n > 2$).[8] The following result characterizes the leader's optimal equilibrium.

Proposition 2. *There exists a leader-optimal equilibrium in which, at every period t along the path of play,*

(i) if $t < t^*$, the price is equal to δ^{t^*-t}, and firm 1 captures the entire market;

(ii) if $t \geq t^*$, the price is equal to the reservation price of 1, each fringe firm earns a market share of $1-\delta$, and firm 1 earns a market share of $1-(n-1)(1-\delta)$.

Firm 1's payoff under any leader-optimal equilibrium is $\overline{V}_1 = \sum_{t=0}^{t^*-1} \delta_1^t \delta^{t^*-t} + \delta_1^{t^*} \overline{V}_1^{st}$.

Firm 1's optimal equilibrium features two phases. In the second phase, which starts at t^*, the cartel arrangement is the same as the optimal stationary arrangement in Proposition 1: prices are set at the monopoly level and each fringe firm obtains a market share of $1-\delta$.

In the initial phase, which lasts until t^*, prices increase gradually over time, reaching the monopoly price at period t^*. Firm 1 serves the entire market during this initial phase. The prices during this initial phase are the highest prices that satisfy fringe firms' IC constraints. Indeed, for all $t < t^*$, the continuation value of a fringe firm under the leader-optimal equilibrium is $\delta^{t^*-t}\sum_{s=t^*}^{\infty} \delta^{s-t^*}(1-\delta) = \delta^{t^*-t}$. Hence, to satisfy fringe firms' incentive compatibility constraints, for all $t < t^*$ the price cannot be larger than δ^{t^*-t}. In Section 4.4 I provide a sketch of the proof of Proposition 2 and explain why the initial phase lasts exactly until t^*.

Under the leader-optimal equilibrium, the incentives of fringe firms are backloaded: fringe firms' continuation profits are increasing over time until time t^*, and the leader's flow profits are increasing until time t^*-1. This is reminiscent of the backloading structure that appears in many dynamic principal-agent models (e.g., Ray, 2002).

To understand why the leader-optimal equilibrium features an initial phase in which the leader serves the entire market, consider increasing the market share of fringe firms in period t. Note that this increase in the market share relaxes the IC constraint of fringe firms at t, allowing for a higher price at t. The overall effect is that fringe firms' flow profits increase

[8] Lemma A.2 in Appendix A.2 implies that $t^* < \infty$ whenever Assumption 1 holds.

at t, and, in spite of the price increase, firm 1's flow profits decrease at t if and only if there are two or more fringe firms (i.e., $n > 2$).

At the same time, because it increases period t profits for fringe firms, an increase in the market share of fringe firms at time t relaxes the incentive constraints of fringe firms for all periods before t. For t small, the benefit from relaxing the IC constraints of fringe firms at all $s < t$ is small, so the optimal scheme sets the market share of fringe firms at zero when $n > 2$. As t increases this benefit increases, and it becomes optimal to set the market share of fringe firms at positive levels.

4.3 Empirical Implications

Proposition 2 shows that, under the leader-optimal equilibrium, prices start at a relatively low level and increase gradually over time until they reach the monopoly price. Thus, a period of gradual price increases, followed by a period of constant high prices, is a potential marker of collusion.

It is worth highlighting that many detected cartels, like citric acid, lysine, and many of the vitamin cartels, featured gradually increasing prices at the outset of collusion (see, e.g., Harrington, 2008, Marshall and Marx, 2012). Previous work by Harrington (2004, 2005) showed that slowly increasing prices may arise when a newly formed cartel seeks to avoid detection by an antitrust authority. Instead, the price dynamics in Proposition 2 are driven by maximization of the leader's equilibrium profits. Proposition 3 below shows that similar price dynamics arise under all Pareto efficient equilibria that put enough weight on firm 1's profits.

The length t^* of the initial phase depends on the number of firms and their discount factors. In particular, t^* is increasing in the number of firms, with $t^* > 0$ if and only if there are three or more firms. With three or more firms, t^* is decreasing in δ, attaining a minimum of $n - 2$ when fringe firms are as patient as firm 1. The following Corollary summarizes this:

Corollary 1. *The length t^* of the initial phase is increasing in the number of firms and is decreasing in the discount factor of fringe firms.*

With two firms, there is no initial phase, and the leader-optimal equilibrium coincides with the equilibrium with stationary prices in Proposition 1.

Another notable feature of the optimal equilibrium in Proposition 2 is that the market share of fringe firms during the second phase of the equilibrium is decreasing in δ: the less patient fringe firms are, the larger their market share has to be to prevent them from defecting. For some parameter values, fringe firms' market shares may be larger than the market share of firm 1: $1 - \delta > 1 - (n-1)(1-\delta) \Longleftrightarrow \delta < \frac{n-1}{n}$.[9]

Lastly, under firm 1's optimal equilibrium, the leader's market share decreases over time: the leader serves the entire market until t^*, and obtains the stationary market share of Proposition 1 from t^* onwards.[10] Hence, a declining market share of the largest firms, accompanied by rising prices, is also a potential marker of collusion.

I note that these dynamics are also broadly consistent with the way some of the vitamin cartels were structured. Connor (2007) documents that the smaller firms in some of the vitamin cartels (beta carotene, vitamin B2 and biotin) were promised increasing market shares to entice them to join the collusive agreement, and to incentivize them not to defect from it. These increasing market shares came at the expense of the largest firms in the industry.[11]

Another case that featured a declining market share by a leading firm is United States v. United States Steel Corporation (1920). In this landmark case, the US government alleged that the US Steel Corporation had monopolized the market for steel. The US Steel Corporation participated in regular meetings with its competitors, in which pricing and production trends were discussed. These meetings "were instituted first in stress of panic, but, their potency being demonstrated, they were afterwards called to control prices in

[9]Note that this inequality is consistent with Assumption 1. Indeed, Assumption 1 gives $\delta > \frac{n-1-\delta_1}{n-1}$, and $\frac{n-1-\delta_1}{n-1} < \frac{n-1}{n}$ for all $\delta_1 \in ((n-1)/n, 1)$.

[10]In the Conclusion I discuss how these results change in the presence of capacity constraints. In particular, if firm 1 does not have enough capacity to serve the entire market, fringe firms capture a positive fraction of the market during the initial phase of the leader-optimal equilibrium.

[11]For example, in the beta carotene cartel, Roche, the largest producer with a 79% market share, agreed to cede 1% of the market each year to its competitor, BASF (Connor, 2007). As I discuss in the Conclusion, in an extension of my model in which firm 1 has a lower marginal cost than its competitors, the leader-optimal equilibrium features the two phases described in Proposition 2 even if there are two firms in the market.

periods of industrial calm" (United States v. United States Steel Corporation, 1920). In spite of this, the lower court ruled against the government, and this ruling was later affirmed by the US Supreme Court. One of the arguments used against the government's claim was that, during the period of alleged monopolization, the market share of the US Steel Corporation declined steadily, from around 50% to around 41%. This was seen as evidence that there was competition in the market. The model in the current paper rationalizes a declining market share by a cartel leader as a way of backloading the incentives of competitors to incentivize them to adhere to the collusive scheme.

4.4 Proof Sketch of Proposition 2

I now present a sketch of the proof of Proposition 2. Note first that, by Lemma 2, it is without loss to focus on equilibria under which fringe firms obtain the same flow profits each period. Hence, the problem of finding firm 1's optimal equilibrium can be written as finding market shares and prices $(x,p) = (x_t, p_t)_{t=0}^{\infty}$ that solve

$$\overline{V}_1 \equiv \sup_{(x,p)} \sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t)p_t \tag{P}$$

$$\text{s.t. } \forall t, \quad (1-x_t)p_t \leq \sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s \tag{IC-fringe}$$

$$\forall t, \quad (n-1)x_t p_t \leq \sum_{s=t+1}^{\infty} \delta_1^{s-t}(1-(n-1)x_s)p_s \tag{IC-leader}$$

$$\forall t, \quad p_t \in [0,1], x_t \in [0, 1/(n-1)], \tag{Feasibility}$$

where $x_t \in [0, 1/(n-1)]$ is the period t market share of each fringe firm, and so $1-(n-1)x_t$ is firm 1's market share. Consider the following relaxed program:

$$\widehat{V}_1 \equiv \sup_{(x,p)} \sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t) p_t \tag{P'}$$

$$\text{s.t. } \forall t, \quad (1-x_t)p_t \leq \sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s \tag{IC-fringe}$$

$$\forall t, \quad p_t \leq 1, p_t x_t \geq 0. \tag{Feasibility'}$$

Relative to program (P), program (P') ignores firm 1's IC constraints and the constraints that $x_t \leq (n-1)/n$ for all t, and imposes the weaker constraints $p_t x_t \geq 0$ for all t instead of $x_t \geq 0, p_t \geq 0$ for all t. Hence, $\widehat{V}_1 \geq \overline{V}_1$.

The Lagrangian associated with program (P') is:

$$\mathcal{L}(x,p,\lambda,\mu,\gamma) = \sum_{t=0}^{\infty} \delta_1^t \left(1-(n-1)x_t\right) p_t + \sum_{t=0}^{\infty} \lambda_t \left(\sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s - (1-x_t)p_t \right) + \sum_{t=0}^{\infty} \mu_t p_t x_t + \sum_{t=0}^{\infty} \gamma_t (1-p_t),$$

where $(\lambda, \mu, \gamma) = (\lambda_t, \mu_t, \gamma_t)_{t=0}^{\infty}$ are non-negative multipliers associated with the constraints in (P'). By standard results in Lagrangian duality (e.g., Boyd and Vandenberghe, 2004)

$$\inf_{(\lambda,\mu,\gamma)\geq 0} \sup_{(x,p)\in(\mathbb{R}\times\mathbb{R})^{\mathbb{N}_0}} \mathcal{L}(x,p,\lambda,\mu,\gamma) \geq \sup_{(x,p)\in(\mathbb{R}\times\mathbb{R})^{\mathbb{N}_0}} \inf_{(\lambda,\mu,\gamma)\geq 0} \mathcal{L}(x,p,\lambda,\mu,\gamma) = \widehat{V}_1. \tag{3}$$

The inequality in (3) is the standard max-min inequality. The equality in (3) follows since $\inf_{\lambda,\mu,\gamma\geq 0} \mathcal{L}(x,p,\lambda,\mu,\gamma) = -\infty$ for any sequence (x,p) that doesn't satisfy the constraints in (P'), and $\inf_{(\lambda,\mu,\gamma)\geq 0} \mathcal{L}(x,p,\lambda,\mu,\gamma) = \sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t)p_t$ for any sequence (x,p) that satisfies the constraints in (P').

By equation (3), the Lagrangian dual $\inf_{\lambda,\mu,\gamma\geq 0} \sup_{(p,x)} \mathcal{L}(x,p,\lambda,\mu,\gamma)$ gives an upper bound to firm 1's optimal equilibrium profits. Moreover, if there exists non-negative multipliers $(\lambda^*, \mu^*, \gamma^*)$ and market shares and prices (x^*, p^*) satisfying the constraints in (P') such that,

$$\forall(\lambda,\mu,\gamma)\geq 0, \forall(x,p)\in(\mathbb{R}\times\mathbb{R})^{\mathbb{N}_0},$$
$$\mathcal{L}(x^*,p^*,\lambda,\mu,\gamma)\geq\mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*)\geq\mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*), \tag{4}$$

then $(\lambda^*,\mu^*,\gamma^*)$ is a solution to $\inf_{(\lambda,\mu,\gamma)\geq 0}\sup_{(p,x)\in(\mathbb{R}\times\mathbb{R})^{\mathbb{N}_0}}\mathcal{L}(x,p,\lambda,\mu,\gamma)$, and (x^*,p^*) is a solution to program (P').[12]

To establish Proposition 2, I find multipliers $(\lambda^*,\mu^*,\gamma^*)\geq 0$ that, together with the market shares and prices (x^*,p^*) in the statement of the proposition, satisfy the saddle-point condition in (4). I then show that (x^*,p^*) satisfies the constraints in (P'), and so from the arguments above, (x^*,p^*) solves program (P'). To complete the proof, I verify that (x^*,p^*) also satisfies the remaining constraints in program (P).

I now describe how I construct the optimal multipliers $(\lambda^*,\mu^*,\gamma^*)$. Note first that the Lagrangian $\mathcal{L}(x,p,\lambda,\mu,\gamma)$ can be re-written as

$$\mathcal{L}(x,p,\lambda,\mu,\gamma)=\sum_{t=0}^{\infty}p_t x_t\left(-\delta_1^t(n-1)+\sum_{\tau=0}^{t}\delta^{t-\tau}\lambda_\tau+\mu_t\right)+p_t(\delta_1^t-\lambda_t-\gamma_t)+\sum_{t=0}^{\infty}\gamma_t. \tag{5}$$

To see the intuition behind this expression, consider the effect of increasing the market share x_t of fringe firms in period t. First, this change decreases firm 1's payoff by $(n-1)\delta_1^t p_t$. Second, since it increases fringe firm's profits in t, this change relaxes the IC constraint of fringe firms in all periods $\tau\leq t$, increasing the Lagrangian by $p_t\sum_{\tau=0}^{t}\delta^{t-\tau}\lambda_\tau$. Third, this change relaxes the non-negativity constraint on $p_t x_t$, increasing the Lagrangian by $p_t\mu_t$.

[12] If there exist $(\lambda^*,\mu^*,\gamma^*)\geq 0$ and (x^*,p^*) satisfying the constraints in (P') such that (4) holds, then

$$\begin{aligned}\sup_{(x,p)}\inf_{(\lambda,\mu,\gamma)\geq 0}\mathcal{L}(x,p,\lambda,\mu,\gamma)&\geq\inf_{(\lambda,\mu,\gamma)\geq 0}\mathcal{L}(x^*,p^*,\lambda,\mu,\gamma)=\mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*)\\&=\sup_{(x,p)}\mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*)\geq\inf_{(\lambda,\mu,\gamma)\geq 0}\sup_{(x,p)}\mathcal{L}(x,p,\lambda,\mu,\gamma),\end{aligned}$$

which, together with (3), implies that

$$\inf_{(\lambda,\mu,\gamma)\geq 0}\sup_{(x,p)}\mathcal{L}(x,p,\lambda,\mu,\gamma)=\sup_{(x,p)}\inf_{(\lambda,\mu,\gamma)\geq 0}\mathcal{L}(x,p,\lambda,\mu,\gamma)=\mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*)=\widehat{V}_1.$$

Since (x^*,p^*) satisfies the constraints in (P'), $\widehat{V}_1=\mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*)=\sum_{t=0}^{\infty}\delta_1^t(1-(n-1)x_t^*)p_t^*$.

Consider next the effect of increasing the price p_t in period t. In addition to the effects of increasing x_t described above, this change also (i) increases firm 1's profits by δ_1^t; (ii) tightens fringe firms' IC constraints at t, lowering the Lagrangian by λ_t; and (iii) tightens the constraint that $p_t \leq 1$, lowering the Lagrangian by γ_t.

Since (x^*, p^*) has to be a maximizer of $\mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*)$, the terms in brackets in (5) must all be equal to zero. Hence, the optimal multipliers $(\lambda^*, \mu^*, \gamma^*)$ solve

$$\forall t, \quad \sum_{\tau=0}^{t} \lambda_\tau^* \delta^{t-\tau} + \mu_t^* = \delta_1^t (n-1), \tag{6}$$

$$\delta_1^t - \lambda_t^* - \gamma_t^* = 0. \tag{7}$$

Next, note that under the scheme in Proposition 2, prices are below the reservation price of 1 for all $t < t^*$. Hence, the multiplier γ_t^* on the constraint $p_t \leq 1$ must be zero for all $t < t^*$. Equation (7) then implies that the multiplier λ_t^* on fringe firms' IC constraint satisfies $\lambda_t^* = \delta_1^t$ for all $t < t^*$. Using this in equation (6), for all $t < t^*$, the multiplier μ_t^* on the non-negativity constraint $p_t x_t \geq 0$ satisfies

$$\mu_t^* = \delta_1^t (n-1) - \sum_{\tau=0}^{t} \delta_1^\tau \delta^{t-\tau}.$$

The definition of t^* in (2) guarantees that $\mu_t^* > 0$ for all $t < t^*$. Intuitively, multiplier μ_t^* is the shadow value of relaxing the non-negativity constraint on $p_t x_t$. A relaxation of this constraint permits a lower market share x_t of fringe firms at time t, increasing firm 1's profits by $\delta_1^t(n-1)$. At the same time, lowering x_t tightens fringe firms' IC constraints for all $t \leq \tau$, and has a cost equal to $\sum_{\tau=0}^{t} \lambda_\tau \delta^{t-\tau} = \sum_{\tau=0}^{t} \delta_1^\tau \delta^{t-\tau}$. For $t < t^*$, the benefit outweighs the cost, and so the leader-optimal equilibrium sets $x_t = 0$.

For $t \geq t^*$, the scheme in Proposition 2 sets $x_t^* = 1 - \delta > 0$ and $p_t^* = 1$. As shown in Section 4.1, these market shares and prices guarantee that fringe firms' IC constraints bind at all periods $t \geq t^*$. Hence, the multiplier μ_t^* on the non-negativity constraint $p_t x_t \geq 0$ is equal to zero for all $t \geq t^*$. Using this in equation (6), for all $t \geq t^*$ multipliers λ_t^* solve the recursive equation $\sum_{\tau=0}^{t} \lambda_\tau^* \delta^{t-\tau} = (n-1)\delta_1^t$, and γ_t^* solves (7).

Finally, note that under the optimal multipliers $(\lambda^*, \mu^*, \gamma^*)$, for all (x,p) the Lagrangian satisfies

$$\mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*) = \sum_{t=0}^{\infty} \gamma_t^*,$$

where the equality uses (5) and the fact that the optimal multipliers $(\lambda^*, \mu^*, \gamma^*)$ satisfy (6) and (7). Intuitively, γ_t^* is the shadow value of relaxing the constraint that the price in period t be less than 1. Since firm 1's profits and fringe firms' IC constraints are both linear in prices, relaxing the constraints that $p_t \leq 1$ for all t uniformly by $y\%$ allows prices to be uniformly increased by $y\%$, raising firm 1's discounted profits by $y\%$.

5 Pareto Optimal Equilibria

I now turn to study equilibria that maximize a weighted sum of leader profits and fringe firms' profits. By Lemma 2, it is without loss to restrict attention to equilibria under which all fringe firms obtain the same flow profits in every period. Hence, program (1) can be written as:

$$\overline{V}(\alpha) \equiv \sup_{(x,p)} \alpha \left(\sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t) p_t \right) + (1-\alpha)(n-1) \left(\sum_{t=0}^{\infty} \delta^t x_t p_t \right) \qquad (\mathrm{P}_\alpha)$$
$$\text{s.t. (IC-fringe), (IC-leader) and (Feasibility).}$$

For $\alpha = 1$ this corresponds to program (P) in Section 4.

For $\alpha \in [1/2, 1]$, define

$$t^*(\alpha) \equiv \inf \left\{ t : \alpha \sum_{\tau=0}^{t} \delta_1^\tau \delta^{t-\tau} \geq \alpha \delta_1^t (n-1) - (1-\alpha)\delta^t (n-1) \right\}.$$

Lemma A.2 in the Appendix shows that under Assumption 1, $t^*(\alpha) < \infty$ for all $\alpha \in [1/2, 1]$. Note that $t^*(\alpha)$ is increasing in α, with $t^*(1/2) = 0$, and that $t^*(1)$ is equal to t^* in equation (2). The following result characterizes the solution to (P_α) for $\alpha \in [1/2, 1]$.

Proposition 3. *For $\alpha \in [1/2, 1]$, there exists an equilibrium that attains $\overline{V}(\alpha)$ in which, at every period t along the path of play,*

(i) if $t < t^*(\alpha)$, the price is equal to $\delta^{t^*(\alpha)-t}$, and firm 1 captures the entire market;

(ii) if $t \geq t^*(\alpha)$, the price is equal to the reservation price of 1, each fringe firm earns a market share of $1-\delta$, and firm 1 earns a market share of $1-(n-1)(1-\delta)$.

Proposition 3 shows that, when the weight α on firm 1 is sufficiently large, the equilibrium that attains $\overline{V}(\alpha)$ shares the same features as firm 1's optimal equilibrium. In particular, this equilibrium also features two phases: an initial phase in which prices slowly increase over time and in which firm 1 captures the entire market, and a second stationary phase in which prices are equal to the monopoly price and in which fringe firms capture their stationary market share of $1-\delta$.

The length $t^*(\alpha)$ of the initial phase is increasing in the number of firms and in the weight on firm 1's payoff, and is decreasing in fringe firms' discount factor. Lastly, there is no initial stage when there are two firms or when $\alpha = 1/2$: in these cases, the equilibrium in Proposition 3 coincides with the stationary equilibrium in Proposition 1.

The equilibria in Proposition 3 have the property that firm 1's IC constraints are never binding. Indeed, in the initial phase they are not binding since firm 1 serves the entire market. And in the stationary phase they are not binding, as shown in Section 4.1. I now characterize the solution to (P_α) for $\alpha \in [0, 1/2)$. A key difference with the equilibria in Proposition 3 is that firm 1's IC constraints may bind when $\alpha < 1/2$.

Define the following two times:

$$t_1^*(\alpha) \equiv \sup\left\{t \in \mathbb{N}_0 : -\alpha\delta_1^t + (1-\alpha)\delta^t > 0\right\}, \text{ and}$$
$$t_2^* \equiv \sup\left\{t \in \mathbb{N}_0 : \frac{\delta_1^t}{1-\delta_1}(1-(n-1)(1-\delta)) \geq 1\right\}.$$

For $\alpha \in [0, 1/2)$, let $t^*(\alpha) \equiv \min\{t_1^*(\alpha), t_2^*\}$. Note that $t_2^* \in [0, \infty)$, since, by Assumption 1, $\frac{1}{1-\delta_1}(1-(n-1)(1-\delta)) > 1$. Hence, $t^*(\alpha) < \infty$. Note further that $t^*(\alpha)$ is decreasing in α. Moreover, when $\delta_1 > \delta$, $t^*(\alpha) = 0$ for $\alpha < 1/2$ close enough to $1/2$.

Proposition 4. *For $\alpha \in [0, 1/2)$, there exists an equilibrium that attains $\overline{V}(\alpha)$ in which, at every period t along the path of play, the price is equal to the reservation price of 1, and*

(i) if $t < t^*(\alpha)$, each fringe firm earns a market share of $1/(n-1)$, and firm 1 earns zero market share;

(ii) if $t = t^*(\alpha)$, each fringe firm earns a market share $x^* \in [(1-\delta), 1/(n-1)]$, and firm 1 earns a market share of $1-(n-1)x^*$;

(iii) if $t > t^*(\alpha)$, each fringe firm earns a market share $1-\delta$, and firm 1 earns a market share of $1-(n-1)(1-\delta)$.

The equilibria in Proposition 4 also feature two phases: an initial phase in which fringe firms capture the entire market, and a second phase in which fringe firms capture their stationary market share of $1-\delta$. In contrast to the equilibria in Proposition 3, prices are constant and equal to the monopoly price throughout the two phases. The length $t^*(\alpha)$ of the initial phase is decreasing in α, and is equal to zero for values of α close to $1/2$. Hence, for $\alpha \approx 1/2$, the optimal equilibrium again coincides with the stationary equilibrium in Proposition 1.

To understand why equilibria that maximize the weighted sum of leader's and fringe firms' payoffs take this form for $\alpha < 1/2$, note that in this case the first-best outcome has: (i) prices always equal to the reservation price of 1, (ii) fringe firms capturing the entire market at all periods $t \leq t_1^*(\alpha)$, and (iii) firm 1 capturing the entire market at all periods $t > t_1^*(\alpha)$. When $t^*(\alpha) = t_1^*(\alpha) \leq t_2^*$, optimal equilibria introduce one modification relative to the first-best outcome: fringe firms obtain their stationary market share of $1-\delta$ at all periods $t > t_1^*(\alpha)$, so as to satisfy their IC constraints. When $t_1^*(\alpha) > t_2^* = t^*(\alpha)$, setting firm 1's market share equal to 0 at all periods $t \leq t_1^*(\alpha)$ and equal to $1-(n-1)(1-\delta)$ at all periods $t > t_1^*(\alpha)$ would violate firm 1's IC constraint at $t = 0$. Hence, in this case, firm 1 starts capturing a positive market share at the earlier time $t_2^* < t_1^*(\alpha)$.

The dynamics of Pareto optimal equilibria. Figure 1 plots the Pareto frontier for parameter values $\delta_1 = 0.92$, $\delta = 0.9$ and $n = 7$, with fringe firms' profits on the horizontal

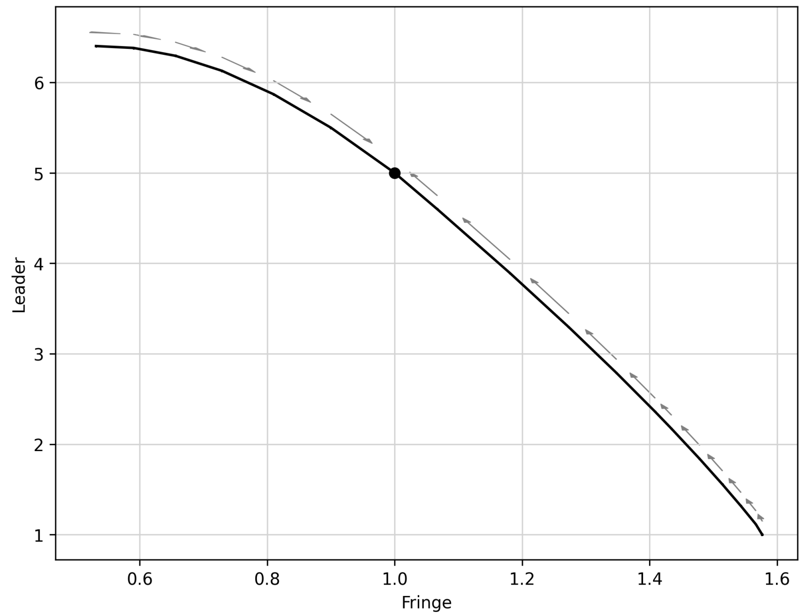

Figure 1: Pareto frontier, with fringe firms' payoffs on the horizontal axis and the leader's payoffs on the vertical axis. Parameter values: $\delta_1 = 0.92, \delta = 0.9, n = 7$. The black circle corresponds to the payoffs under the stationary equilibrium in Proposition 1.

axis and the leader's profits on the vertical axis. The black circle corresponds to profits under the stationary equilibrium in Proposition 1, which maximizes the weighted sum of profits for $\alpha = 1/2$. Points to the left of the circle correspond to Pareto optimal equilibria for values of α larger than 1/2, and points to the right of the circle correspond to Pareto optimal equilibria for values of α smaller than 1/2.

Consider first Pareto optimal equilibria for $\alpha > 1/2$. On the path of play under these equilibria, firms' continuation payoffs move along the frontier towards the stationary payoffs marked with the black circle. Indeed, at time $t^*(\alpha)$, continuation play under the equilibrium that attains $\overline{V}(\alpha)$ coincides with equilibrium play under the leader-optimal stationary equilibrium. Consider next Pareto optimal equilibria with $\alpha < 1/2$. Under such an equilibrium, firms' continuation payoffs on the path of play also move along the frontier towards the stationary payoffs marked with the black circle, reaching that point at time $t^*(\alpha)$. The gray arrows in Figure 1 illustrate these equilibrium dynamics.

I end this section with a discussion of the relation between my results and those in Obara and Zincenko (2017), who also study a repeated Bertrand oligopoly game with heterogeneous discount factors. Obara and Zincenko (2017) focus on equilibria with the property that prices are always equal to the monopoly price. They show that optimal equilibria within this class have the same properties as the equilibria described in Proposition 4: less patient firms enter the market weakly earlier than more patient firms, and equilibrium play converges to the stationary equilibrium in Proposition 1. Propositions 3 and 4 show that equilibria with these properties are Pareto efficient for $\alpha \in [0, 1/2]$, but not for $\alpha > 1/2$.

Although they don't characterize all Pareto optimal equilibria, Obara and Zincenko (2017) do establish two additional results. First, they show that Pareto optimal equilibria that don't have constant prices feature industry profits that increase over time during the first periods. Second, they show that under any Pareto efficient equilibrium, play eventually converges to a stationary profile. Propositions 3 and 4 strengthen these results by characterizing the set of Pareto efficient equilibria.

6 Stochastic Demand

I now study an extension of the model with stochastic demand, in the spirit of Rotemberg and Saloner (1986). The purpose of this section is two-fold. First, I show that optimal equilibria under stochastic demand retain key properties of the equilibria in Sections 4 and 5. Second, the results in this section illustrate how Lagrangian methods can also be used to characterize optimal equilibria in this stochastic setting.

The model is the same as in Section 2, with the only difference that at each time t the buyers' reservation price r_t is drawn i.i.d. from a discrete distribution F_r supported on $\{r_1, ..., r_K\}$, with $0 < r_1 < r_2 < ... < r_K$. For $k = 1, ..., K$, let $q_k = \mathsf{prob}_{F_r}(r = r_k)$. At each period t, firms publicly observe the realization of the reservation price r_t, and then simultaneously choose actions $(\mathbf{p}_t, \rho_t) = (p_{i,t}, \rho_{i,t})_{i \in N}$. Market shares are determined in the same way as in the baseline model in Section 2. As in the model in Section 2, there is perfect monitoring: the public history at time t is $h_t = (r_s, \mathbf{p}_s, \rho_s)_{s<t}$.

For conciseness, I focus on leader-optimal equilibria. Let Σ_{p} denote the set of pure strategy SPNE of this game. The leader-optimal equilibrium payoff is $\overline{V}_1 \equiv \sup_{\sigma \in \Sigma_p} V_1(\sigma)$, where, for each strategy profile σ, $V_1(\sigma)$ is player 1's expected discounted payoff under σ.

Let $\overline{r} \equiv \sum_k q_k r_k$ denote the expected reservation price. I make the following assumption:

Assumption 2. *(i)* $\delta \overline{r} < r_1$*;*

(ii) $(n-1)(r_K - \delta \overline{r}) < \min\{r_K, \delta_1((n-1)r_K - (n-2)\overline{r})\}$.

Assumption 2 implies that demand shocks are not too large: Assumption 2(i) puts a lower bound on the lowest demand shock r_1, and Assumption 2(ii) implies that the highest demand shock r_K is such that $r_K(n-2) < \delta \overline{r}(n-1)$. It is worth noting that, while Assumption 2 implies Assumption 1, the two assumptions are equivalent when demand is constant. Indeed, if $r_k = \overline{r}$ for all k, then Assumption 2(i) holds for all $\delta < 1$, and Assumption 2(ii) becomes $(n-1)(1-\delta) < \delta_1$, which is Assumption 1. In the end of this section I describe the leader-optimal equilibrium when Assumption 2 does not hold.

Note that Lemma 2 extends to this environment. Hence, it is without loss to focus on equilibria under which fringe firms earn the same flow profits at all periods. The problem of finding the leader-optimal equilibrium can then be written as

$$\begin{aligned} \overline{V}_1 \equiv \sup_{(x,p)} \mathbb{E}_0 &\left[\sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t) p_t \right] \qquad (\mathrm{P}_s) \\ \text{s.t. } \forall t, \quad (1-x_t)p_t &\leq \mathbb{E}_t \left[\sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s \right] \\ \forall t, \quad (n-1)x_t p_t &\leq \mathbb{E}_t \left[\sum_{s=t+1}^{\infty} \delta_1^{s-t} (1-(n-1)x_s) p_s \right] \\ \forall t, \quad p_t &\in [0, r_t], x_t \in [0, 1/(n-1)], \end{aligned}$$

where $(x,p) = (x_t, p_t)_{t=0}^{\infty}$ is a stochastic process adapted to the filtration generated by (r_t).

As a first step towards solving (P_s), consider a stationary equilibrium in which, at each time t, prices are equal to the reservation price r_t, and the market share of each fringe firm is equal to $x(r_t) \in [0, 1/(n-1)]$ (with firm 1 obtaining market share $1-(n-1)x(r_t)$). Pick market shares $(x(r_k))_{k=1}^K$ so that the IC constraints of fringe firms bind at all periods under this stationary equilibrium: i.e., $(x(r_k))_{k=1}^K$ solves

$$\text{for } k = 1, ..., K, \quad r_k(1 - x(r_k)) = \frac{\delta}{1-\delta} \sum_{k'} q_{k'} r_{k'} x(r_{k'}). \tag{8}$$

The right-hand side of (8) is the expected continuation payoff of fringe firms under this stationary equilibrium: at each period, when the reservation price is $r_{k'}$, fringe firms serve a fraction $x(r_{k'})$ of the market at price $r_{k'}$. The left-hand side is the deviation gain: when buyers' reservation price is r_k, a fringe firm can undercut the price and serve the entire market.

The solution to (8) has $x(r_k) = \frac{r_k - \delta \overline{r}}{r_k}$ for $k = 1, ..., K$. Assumption 2 guarantees that $x(r_k) \in (0, 1/(n-1))$ for all k. Moreover, Assumption 2 also guarantees that, under this profile, firm 1's IC constraints are also satisfied.[13] Market shares $(x(r_k))_{k=1}^K$ generalize the stationary market shares of Proposition 1 to the current setting: when $r_k = \overline{r}$ for all k, $x(r_k) = 1 - \delta$ for all k.

It is worth highlighting that this stationary equilibrium differs from the symmetric stationary equilibrium in Rotemberg and Saloner (1986): instead of featuring countercyclical pricing, this stationary equilibrium features pro-cyclical pricing (i.e., prices always equal to the reservation price) and counter-cyclical market shares for the leader. Indeed, market share $x(r)$ is increasing in r, to satisfy fringe firms' IC constraints; and so the leader's market share, $1-(n-1)x(r)$, is decreasing in r.

Let V_f denote fringe firm's expected discounted profits under this stationary equilibrium:

$$V_f = \frac{1}{1-\delta} \sum_k q_k r_k x(r_k) = \overline{r},$$

where the second equality uses $x(r_k) = \frac{r_k - \delta \overline{r}}{r_k}$ for all k and $\overline{r} = \sum_k q_k r_k$.

[13]The proof of Proposition 5 in Appendix A.3 shows this.

Recall the definition of t^* in equation (2):

$$t^* \equiv \inf \left\{ t \in \mathbb{N}_0 : \sum_{\tau=0}^{t} \delta^{t-\tau} \delta_1^{\tau} \geq \delta_1^t (n-1) \right\},$$

The following result characterizes the solution to (P_s).

Proposition 5. *With stochastic demand, there exists a leader-optimal equilibrium in which, at every period t along the path of play:*

(i) if $t < t^*$, the price is equal to $\delta^{t^*-t}\overline{r}$, and firm 1 captures the entire market;

(ii) if $t \geq t^*$, the price is equal to the reservation price of r_t, each fringe firm earns a market share of $x(r_t)$, and firm 1 earns a market share of $1-(n-1)x(r_t)$.

Proposition 5 shows that the key properties of leader-optimal equilibria in Proposition 2 extend to this setting with stochastic demand. In particular, leader-optimal equilibria have two phases: an initial phase with slowly increasing prices and with firm 1 capturing the entire market, and the second phase with prices at the monopoly price and with stationary market shares. As highlighted above, the market share of fringe firms during the stationary phase is pro-cyclical, and the market share of the leader is counter-cyclical.

I end this section by discussing the role that Assumption 2 plays in Proposition 5, and how the results generalize when the assumption does not hold. Note that Assumption 2(i) guarantees that the price $\delta^{t^*-t}\overline{r} = \mathbb{E}_t[\delta^{t^*-t}V_f]$ during the initial phase is always below the reservation price (recall that V_f is fringe firm's profits under the stationary equilibrium). When Assumption 2 doesn't hold, the leader-optimal equilibrium still features two phases: an initial phase with prices equal to $\min\{\mathbb{E}_t[\delta^{t^*-t}V_f], r_t\}$ (and with the leader capturing the entire market), and a second stationary phase.[14]

Interestingly, it can be shown that the length of the initial phase is random when Assumption 2 doesn't hold. To see why, recall from the proof sketch of Proposition 2 in Section

[14]When Assumption 2 doesn't hold, fringe firms' market shares during the stationary phase may be equal to zero at states with very low demand, and may be equal to $1/(n-1)$ in states with very high demand. Moreover, in states with very high demand, prices may be set below the reservation price to satisfy firms' IC constraints.

4.4 that the initial phase ends the first period t such that $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau$ is sufficiently large, where λ_τ is the multiplier on the IC constraint of fringe firms at time τ. When Assumption 2 doesn't hold, fringe firms' IC constraints don't bind in periods $t < t^*$ with $r_t < \mathbb{E}_t[\delta^{t^*-t}V_f]$. Hence, the multipliers (λ_t) (which are now a stochastic process adapted to (r_t)) may be zero at some periods during the initial phase, depending on the realized demand shocks. As a result, the end of the initial phase, which is given by the first time $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau$ is large enough, will be random.

7 Conclusion

This paper studies optimal collusive equilibria in a market with a cartel leader and several less patient competitors. These equilibria have two key features. First, they may involve an initial phase in which prices increase gradually over time. Second, under every optimal equilibrium, play converges in finite time to a stationary profile with prices set at the monopoly level and with constant market shares. From a methodological standpoint, the paper uses Lagrangian methods to characterize Pareto efficient equilibria in a repeated Bertrand oligopoly game.

Perhaps an unsatisfactory feature of the equilibria in Propositions 2 and 3 is that firm 1 captures the entire market in the initial phase of the collusive scheme. The reason for this is that, in this model, each firm is assumed to have enough capacity to serve the entire market at every period. When firms are capacity constrained, leader-optimal equilibria have the property that the leader sells at full capacity during the initial phase, and fringe firms serve the remaining fraction of the market.

Throughout the paper, I maintained the assumption that firms have different time preferences but share the same marginal cost. Another natural extension is to consider settings in which firm 1, the most patient firm, also has a cost advantage relative to fringe firms. It is possible to show that, when firm 1's cost advantage is not too large, firm 1's optimal equilibrium shares the same features of the equilibrium in Proposition 2: there is an initial phase in which prices rise slowly and in which firm 1 captures the entire market, followed by a stationary phase in which prices are equal to the reservation price and in which market shares are stationary. Moreover, the length t^* of the initial phase is positive even with two firms (i.e., even if $n = 2$).

A Appendix

A.1 Proofs of Section 3

Proof of Lemma 1. Suppose $(p_t, \mathbf{x}_t)$ is implementable by some strategy profile $\sigma \in \Sigma_{\mathsf{p}}$. Then, for every history h_t and each $i \in N$, player i's continuation payoff at h_t under σ is $V_i(\sigma|h_t) = \sum_{s=t}^{\infty} \delta_i^{s-t} x_{i,s} p_s$. Since $\sigma \in \Sigma_{\mathsf{p}}$, it must be that for all i and all on-path histories h_t

$$\sum_{s=t}^{\infty} \delta_i^{s-t} x_{i,s} p_s \geq p_t + \delta 0 \Longleftrightarrow (1 - x_{i,t}) p_t \leq \sum_{s=t+1}^{\infty} \delta_i^{s-t} x_{i,s} p_s.$$

The first inequality follows since firm i can get $p_t + \delta 0$ at time t by undercutting price p_t. Hence, if $(p_t, \mathbf{x}_t)$ is implementable by some strategy profile $\sigma \in \Sigma_{\mathsf{p}}$, (IC) must hold.

Next, let $(p_t, \mathbf{x}_t)$ be a sequence satisfying (IC), and consider the following strategy profile σ. On the path of play, at each time t firm $i \in N$ chooses $p_{i,t} = p_t$ and $\rho_{i,t} \in [0, 1]$, where $\rho_t = (\rho_{j,t})_{j \in N} \in [0, 1]^n$ is such that, for all $j \in N$, $\frac{\rho_{j,t}}{\sum_{k \in N} \rho_{k,t}} = x_{j,t}$. Deviations by any player i are punished by Nash reversion (i.e., $p_{i,t} = 0$ and $\rho_{i,t} = 1$ for all i and all t following a deviation). Since $(p_t, \mathbf{x}_t)$ satisfies (IC), no player has an incentive to deviate at any history under σ, and so $\sigma \in \Sigma_{\mathsf{p}}$. Hence, if $(p_t, \mathbf{x}_t)$ satisfies (IC), it is implementable by some strategy profile $\sigma \in \Sigma_{\mathsf{p}}$. ■

Proof of Lemma 2. Fix an equilibrium σ, and let $(p_t, \mathbf{x}_t)$ be the sequence of prices and allocations implemented on the equilibrium path under σ. By Lemma 1, $(p_t, \mathbf{x}_t)$ satisfies

$$\forall t, \forall i, \quad (1 - x_{i,t}) p_t \leq \sum_{s=t-1}^{\infty} \delta_i^{s-t} x_{i,s} p_s. \tag{9}$$

Summing (9) across all $i \neq 1$ and diving by $n-1$, we get

$$\forall t, \quad \left(1 - \frac{\sum_{i\neq 1} x_{i,t}}{n-1}\right) p_t \leq \sum_{s=t-1}^{\infty} \delta^{s-t} \left(\frac{\sum_{i\neq 1} x_{i,s}}{n-1}\right) p_s \tag{10}$$

Consider the sequence of allocations $(\widehat{\mathbf{x}}_t)$ such that $\widehat{x}_{1,t} = x_{1,t}$ for all t, and, for all $i \neq 1$, $\widehat{x}_{i,t} = \frac{\sum_{i\neq 1} x_{i,s}}{n-1}$. By (9) and (10), $(p_t, \widehat{\mathbf{x}}_t)$ satisfies (IC). By Lemma 1, there exists an equilibrium $\widehat{\sigma} \in \Sigma_p$ that implements $(p_t, \widehat{\mathbf{x}}_t)$. By construction $V_1(\widehat{\sigma}) = V_1(\sigma)$, $\sum_i V_i(\widehat{\sigma}) = \sum_i V_i(\sigma)$ and, for all on path histories h and all $i, j \neq 1$, $\pi_i(\widehat{\sigma}|h) = \pi_j(\widehat{\sigma}|h)$. Moreover, without loss we can choose $\widehat{\sigma}$ such that, off the equilibrium path, firms play the static Nash equilibrium. Hence, for all h and all $i, j \neq 1$, $\pi_i(\widehat{\sigma}|h) = \pi_j(\widehat{\sigma}|h)$. ■

A.2 Proofs of Sections 4 and 5

Proof of Proposition 1. I first show that any equilibrium $\sigma \in \Sigma_{\mathsf{p}}^{\mathsf{st}}$ that attains $\overline{V}_1^{\mathsf{st}} > 0$ must have the property that, along the path of play, prices are equal to the monopoly price of 1 at every period. To see why, consider an equilibrium $\sigma \in \Sigma_{\mathsf{p}}^{\mathsf{st}}$ with $V_1(\sigma) > 0$. Hence, under σ, along the path of play the lowest price p_t is equal to some $p \in (0, 1]$ at every period.

Let $(x_{i,t})_{i\in N, t=0}^{\infty}$ denote the market shares along the path of play under σ. Then, by Lemma 1, it must be that

$$\forall t, \forall i, \quad (1 - x_{i,t}) \leq \sum_{s=t+1}^{\infty} \delta_i^{s-t} x_{i,s} \tag{IC-st}$$

Since the price level $p \in (0, 1]$ doesn't affect the constraints in (IC-st), it is optimal to set $p_t = r = 1$ for all t.

Next, note that by the same argument as in the proof of Lemma 2, it is without loss to consider equilibria such that, for all $i \neq 1$ and all t, $x_{i,t} = x_t \in [0, 1/(n-1)]$. Hence, I can re-write player 1's problem as

$$\overline{V}_1^{\mathsf{st}} = \sup_{x \in [0,1/(n-1)]^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \delta_1^t \left(1-(n-1)x_t\right) \text{ s.t. (IC-}\mathsf{st}\text{).} \tag{P_{st}}$$

Ignore for now the constraints that $x_t \in [0, 1/(n-1)]$ for all t and the incentive compatibility constraints of firm 1, and consider the following relaxed program:

$$\widehat{V}_1^{\mathsf{st}} = \sup_{x \in \mathbb{R}^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \delta_1^t \left(1-(n-1)x_t\right) \tag{P'_{st}}$$

$$\text{s.t. } \forall t, \quad (1-x_t) \leq \sum_{s=t+1}^{\infty} \delta^{s-t} x_s \tag{IC'-$i \neq 1$}$$

The Lagrangian associated with this program is

$$\mathcal{L}(x,\lambda) = \sum_{t=0}^{\infty} \delta_1^t \left(1-(n-1)x_t\right) + \sum_{t=0}^{\infty} \lambda_t \left(\sum_{s=t+1}^{\infty} \delta^{s-t} x_s - (1-x_t) \right).$$

where $\lambda = (\lambda_t)_{t=0}^{\infty}$ are non-negative multipliers associated with the constraints in (P'_{st}). Note then that,

$$\inf_{\lambda \geq 0} \sup_{x \in \mathbb{R}^{\mathbb{N}_0}} \mathcal{L}(x,\lambda) \geq \sup_{x \in \mathbb{R}^{\mathbb{N}_0}} \inf_{\lambda \geq 0} \mathcal{L}(x,\lambda) = \widehat{V}_1^{\mathsf{st}},$$

where the inequality is the max-min inequality, and the equality follows since $\inf_{\lambda \geq 0} \mathcal{L}(x,\lambda) = \sum_{t=0}^{\infty} \delta_1^t \left(1-(n-1)x_t\right)$ for any x that satisfies the constraints in (P'_{st}), and $\inf_{\lambda \geq 0} \mathcal{L}(x,\lambda) = -\infty$ for any x that doesn't satisfy these constraints.

Moreover, if we can find $\lambda^* \geq 0$ and $x^* \in [0, 1/(n-1)]^{\mathbb{N}_0}$ such that

$$\forall \lambda \geq 0, \forall x \in \mathbb{R}^{\mathbb{N}_0}, \qquad \mathcal{L}(x^*,\lambda) \geq \mathcal{L}(x^*,\lambda^*) \geq \mathcal{L}(x,\lambda^*), \tag{11}$$

then λ^* solves the dual program $\inf_{\lambda \geq 0} \sup_{x \in \mathbb{R}^{\mathbb{N}_0}} \mathcal{L}(x,\lambda)$, and x^* solves program (P'_{st}). In addition, if x^* also satisfies the remaining constraints in (P_{st}), then x^* also solves (P_{st}). Hence, to establish the Proposition, I find multipliers λ^* that, together with the allocations x^* in the statement of the Proposition, satisfy the saddle-point condition (11). I then verify that x^* satisfies the remaining constraints in (P_{st}).

Let $\lambda^* = (\lambda^*_t)$ be given by,

$$\lambda^*_t = \begin{cases} n-1 & \text{if } t = 0, \\ (n-1)\delta_1^{t-1}(\delta_1 - \delta) & \text{if } t > 0. \end{cases}$$

Clearly, $\lambda^*_t \geq 0$ for all t. I use the following Lemma.

Lemma A.1. *For all* $t \geq 0$, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda^*_\tau = (n-1)\delta_1^t$.

Proof. The proof is by induction. Note that for $t = 0$, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda^*_t = \lambda^*_0 = (n-1)\delta_1^0$. Next, let $t \geq 1$ and suppose $\sum_{\tau=0}^{s} \delta^{s-\tau}\lambda^*_\tau = (n-1)\delta_1^s$ for $s = 0, ..., t-1$. Then,

$$\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda^*_\tau = \lambda^*_t + \delta \sum_{\tau=0}^{t-1} \delta^{t-1-\tau}\lambda^*_\tau = (n-1)\delta_1^{t-1}(\delta_1 - \delta) + \delta(n-1)\delta_1^{t-1} = (n-1)\delta_1^t,$$

where the second equality uses the induction hypothesis. Hence, for all $t \geq 0$, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda^*_\tau = (n-1)\delta_1^t$. $\square$

Let x^* be the allocation in the statement of the Proposition: i.e., $x^*_t = 1 - \delta$ for all t. I now show that (x^*, λ^*) satisfies the saddle-point condition (11). Note first that, for all t,

$$\sum_{s=t+1}^{\infty} \delta^{s-t} x^*_s = \frac{\delta}{1-\delta}(1-\delta) = \delta = 1 - x^*_t.$$

Hence, (x^*, λ^*) satisfy the following complementary slackness conditions:

$$\forall t, \quad \lambda^*_t \left(\sum_{s=t+1}^{\infty} \delta^{s-t} x^*_s - (1 - x^*_t) \right) = 0.$$

This implies that, for all $\lambda \geq 0$, $\mathcal{L}(x^*, \lambda) = \mathcal{L}(x^*, \lambda^*)$, so the first inequalities in (11) hold.

Note next that the Lagrangian can be re-written as

$$\mathcal{L}(x, \lambda^*) = \sum_{t=0}^{\infty} (\delta_1^t - \lambda^*_t) + \sum_{t=0}^{\infty} x_t \left(-(n-1)\delta_1^t + \sum_{\tau=0}^{t} \delta^{t-\tau}\lambda^*_\tau \right) = \sum_{t=0}^{\infty} (\delta_1^t - \lambda^*_t),$$

where the last equality follows from Lemma A.1. Hence, $\mathcal{L}(x^*, \lambda^*) = \mathcal{L}(x, \lambda^*)$ for all x, and so x^* solves (P'_{st}).

Finally, I show that x^* also satisfies the remaining constraints in (P_{st}). By Assumption 1, $x_t^* = 1 - \delta < 1/(n-1)$. Next, note that the IC constraints of firm 1 hold since, for all t,

$$\sum_{s=t+1}^{\infty} \delta_1^t (1 - (n-1)x_t^*) = \frac{\delta_1}{1-\delta_1}(1 - (n-1)(1-\delta)) > (n-1)(1-\delta) = (n-1)x_t^*,$$

where the strict inequality follows from Assumption 1. ■

Proof of Proposition 2. Follows from Proposition 3. ■

Lemma A.2. *Under Assumption 1, $t^*(\alpha) < \infty$ for all $\alpha \in [1/2, 1]$.*

Proof. Let $\beta \equiv \frac{\delta_1}{\delta} \geq 1$, and note that

$$\alpha \sum_{\tau=0}^{t} \delta_1^\tau \delta^{t-\tau} \geq \alpha \delta_1^t (n-1) - (1-\alpha)\delta^t (n-1) \iff \alpha \sum_{\tau=0}^{t} \beta^\tau \geq \alpha \beta^t (n-1) - (1-\alpha)(n-1).$$

When $\beta = 1$, the inequality holds if and only if $t \geq n - 2 - \frac{1-\alpha}{\alpha}(n-1)$, and so $t^*(\alpha) < \infty$.

For $\beta > 1$, $\sum_{\tau=0}^{t} \beta^\tau = \frac{\beta^{t+1}-1}{\beta-1}$, and so

$$\begin{aligned} \alpha \sum_{\tau=0}^{t} \beta^\tau &\geq \alpha \beta^t (n-1) - (1-\alpha)(n-1) \\ \iff \alpha \beta^t ((n-1) - \beta(n-2)) &\geq \alpha - (\beta - 1)(1-\alpha)(n-1). \end{aligned}$$

By Assumption 1, $(n-1)\delta > n - 1 - \delta_1 > n - 2 > \delta_1(n-2)$, and so $n - 1 - \beta(n-2) = \frac{1}{\delta}(\delta(n-1) - \delta_1(n-2)) > 0$. Note that this implies that $t^*(\alpha) < \infty$. □

Proof of Proposition 3. Ignore firm 1's IC constraints and the constraints that $x_t \leq 1/(n-1)$ for all t. Moreover, replace the constraints that $x_t \geq 0$ and $p_t \geq 0$ for all t with the weaker constraints $p_t x_t \geq 0$. Consider then the following relaxation to program (P_α):

$$\widehat{V}(\alpha) \equiv \sup_{(x_t,p_t)_{t=0}^{\infty}} \sum_{t=0}^{\infty} \alpha\delta_1^t(1-(n-1)x_t)p_t + \sum_{t=0}^{\infty}(1-\alpha)(n-1)\delta^t x_t p_t \tag{P'_α}$$

$$\text{s.t. } \forall t, \quad (1-x_t)p_t \leq \sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s \tag{IC-$i \neq 1$}$$

$$\forall t, \quad p_t \leq 1, p_t x_t \geq 0. \tag{Feasibility'}$$

Since this is a relaxation of program (P_α), $\widehat{V}(\alpha) \geq \overline{V}(\alpha)$ for all $\alpha \in [1/2, 1]$.

The Lagrangian associated with (P'_α) is

$$\begin{aligned}\mathcal{L}(x,p,\lambda,\mu,\gamma) = &\sum_{t=0}^{\infty} \alpha\delta_1^t\left(1-(n-1)x_t\right)p_t + \sum_{t=0}^{\infty}(1-\alpha)(n-1)\delta^t x_t p_t \\ &+ \sum_{t=0}^{\infty}\lambda_t\left(\sum_{s=t+1}^{\infty}\delta^{s-t}x_s p_s - (1-x_t)p_t\right) \\ &+ \sum_{t=0}^{\infty}\mu_t p_t x_t + \sum_{t=0}^{\infty}\gamma_t(1-p_t),\end{aligned}$$

where (λ, μ, γ) are non-negative multipliers associated with the constraints in (P'_α). To prove the result, I construct multipliers $(\lambda^*, \mu^*, \gamma^*) \geq 0$ that, together with the prices and allocations (p_t^*, x_t^*) in the statement of the Proposition, satisfy the saddle-point condition

$$\begin{aligned}&\forall(\lambda,\mu,\gamma) \geq 0, \forall(x,p), \\ &\quad \mathcal{L}(x^*,p^*,\lambda,\mu,\gamma) \geq \mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*) \geq \mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*),\end{aligned} \tag{12}$$

As explained in Section 4.4, this implies that (p_t^*, x_t^*) is a solution to the relaxed program (P'_α). Lastly, I show that (p_t^*, x_t^*) also satisfy the remaining constraints in program (P_α), so it is also a solution that program.

Consider the following multipliers λ^*, μ^* and γ^*:

$$\lambda_t^* = \begin{cases} \alpha\delta_1^t & \text{if } t < t^*(\alpha) \\ \alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1) - \alpha\sum_{\tau=0}^{t-1}\delta_1^\tau\delta^{t-\tau} & \text{if } t = t^*(\alpha) \\ \alpha\delta_1^{t-1}(n-1)(\delta_1-\delta) & \text{if } t > t^*(\alpha), \end{cases}$$

$$\mu_t^* = \mathbf{1}_{t<t^*(\alpha)}\left(\alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1) - \alpha\sum_{\tau=0}^{t}\delta_1^\tau\delta^{t-\tau}\right)$$

$$\gamma_t^* = \alpha\delta_1^t - \lambda_t^*,$$

with the convention that $\sum_{\tau=0}^{-1}\delta_1^\tau\delta^{t-\tau} = 0$. I use the following Lemma:

Lemma A.3. *(i)* $(\lambda^*, \mu^*, \gamma^*) \geq 0$*;*

(ii) For all $t \geq 0$, $\sum_{\tau=0}^{t}\lambda_\tau^*\delta^{t-\tau} = \alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1) - \mu_t^*$.

Proof. I start by proving part (i). To economize on notation, in what follows I write t^* instead of $t^*(\alpha)$. Clearly, $\lambda_t^* \geq 0$ for all $t \neq t^*$. If $t^* = 0$, then $\lambda_{t^*}^* = \alpha(n-1) - (1-\alpha)(n-1) \geq 0$, where the inequality uses $\alpha \geq 1/2$. Suppose next that $t^* > 0$, and note that

$$\begin{aligned} \alpha\sum_{\tau=0}^{t^*-1}\delta_1^\tau\delta^{t^*-\tau} &= \alpha\delta\sum_{\tau=0}^{t^*-1}\delta_1^\tau\delta^{t^*-1-\tau} < \delta\left(\alpha\delta_1^{t^*-1}(n-1) - (1-\alpha)\delta^{t^*-1}(n-1)\right) \\ &\leq \delta_1^{t^*}\alpha(n-1) - (1-\alpha)\delta^{t^*}(n-1) \end{aligned}$$

where the strict inequality follows from the definition of t^*, and the weak inequality follows since $\delta \leq \delta_1$. Hence, $\lambda_t^* \geq 0$ for all t.

Next, note that, by the definition of t^*, $\mu_t^* \geq 0$ for all $t \geq 0$. Lastly, note that $\gamma_t^* = 0$ for all $t < t^*$. For $t = t^*$,

$$\begin{aligned} \gamma_{t^*}^* &= \alpha\delta_1^{t^*} - \left(\alpha\delta_1^{t^*}(n-1) - (1-\alpha)\delta^{t^*}(n-1) - \alpha\sum_{\tau=0}^{t^*-1}\delta_1^\tau\delta^{t^*-\tau}\right) \\ &= \alpha\sum_{\tau=0}^{t^*}\delta_1^\tau\delta^{t^*-\tau} - \left(\alpha\delta_1^{t^*}(n-1) - (1-\alpha)\delta^{t^*}(n-1)\right) \geq 0 \end{aligned}$$

where the inequality follows from the definition of t^*. For all $t > t^*$,

$$\gamma_t^* = \alpha\delta_1^t - \alpha\delta_1^{t-1}(n-1)(\delta_1 - \delta) = \alpha\delta_1^{t-1}(\delta(n-1) - \delta_1(n-2)) > 0$$

where the strict inequality follows since, by Assumption 1, $(n-1)\delta > n-1-\delta_1$, and since $n-1-\delta_1 > \delta_1(n-2)$ for all $\delta_1 \in (0,1)$. Hence, $(\lambda^*, \mu^*, \gamma^*) \geq 0$.

I now prove part (ii). For all $t < t^*$, the equality follows from the definition of μ_t^* and from $\lambda_\tau^* = \alpha\delta_1^\tau$ for all $\tau < t^*$. For $t = t^*$,

$$\lambda_t^* = \alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1) - \alpha\sum_{\tau=0}^{t-1}\delta_1^\tau\delta^{t-\tau} \Longleftrightarrow \sum_{\tau=0}^{t}\lambda_\tau^*\delta^{t-\tau} = \alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1) - \mu_t^*,$$

where I used $\lambda_\tau^* = \alpha\delta_1^\tau$ for all $\tau < t^*$ and $\mu_{t^*}^* = 0$. I now show that the equality also holds for all $t > t^*$. Towards an induction, suppose the equality holds for $s = 0, .., t$, with $t \geq t^*$. Then,

$$\begin{aligned}
&\alpha\delta_1^{t+1}(n-1) - (1-\alpha)\delta^{t+1}(n-1) - \sum_{\tau=0}^{t+1}\lambda_\tau^*\delta^{t+1-\tau} - \mu_{t+1}^* \\
=&\alpha\delta_1^{t+1}(n-1) - (1-\alpha)\delta^{t+1}(n-1) - \lambda_{t+1}^* - \delta\left(\alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1)\right) = 0,
\end{aligned}$$

where the first equality follows since $\mu_{t+1}^* = 0$ and since, by the induction hypothesis and using $\mu_t^* = 0$, $\sum_{\tau=0}^{t+1}\lambda_\tau^*\delta^{t+1-\tau} = \lambda_{t+1}^* + \delta\sum_{\tau=0}^{t}\lambda_\tau^*\delta^{t-\tau} = \lambda_{t+1}^* + \delta(\alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1))$, and the last equality uses $\lambda_{t+1}^* = \alpha\delta_1^t(n-1)(\delta_1 - \delta)$. Hence, $\sum_{\tau=0}^{t+1}\lambda_\tau^*\delta^{t+1-\tau} = \alpha\delta_1^{t+1}(n-1) - (1-\alpha)\delta^{t+1}(n-1) - \mu_{t+1}^*$. □

Let (x^*, p^*) be the allocations and prices in the statement of the Proposition: i.e., $x_t^* = \mathbf{1}_{t \geq t^*}(1-\delta)$, $p_t^* = \delta^{t^*-t}$ for all $t < t^*$, and $p_t^* = 1$ for all $t \geq t^*$. I now show that (x^*, p^*) and $(\lambda^*, \mu^*, \gamma^*)$ satisfy the saddle-point condition (12). I use the following Lemma.

Lemma A.4. (x^*, p^*) *satisfies the constraints in program* (P'_α). *Moreover, for all* t,

$$\lambda_t^* \left(\sum_{s=t+1}^{\infty} \delta^{s-t} x_s^* p_s^* - (1 - x_t^*) p_t^* \right) = 0,$$

$$\mu_t^* p_t^* x_t^* = 0,$$

$$\gamma_t^* (1 - p_t) = 0.$$

Proof. Note first that, under (x^*, p^*), incentive constraints for fringe firms bind for all t. Indeed, for all $t < t^*$,

$$\sum_{s=t+1}^{\infty} \delta^{s-t} x_s^* p_s^* = \delta^{t^*-t} \frac{1}{1-\delta}(1-\delta) = \delta^{t^*-t} = p_t^* = p_t^*(1 - x_t^*).$$

For all $t \geq t^*$,

$$\sum_{s=t+1}^{\infty} \delta^{s-t} x_s^* p_s^* = \frac{\delta}{1-\delta}(1-\delta) = \delta = (1 - x_t^*) p_t^*.$$

Moreover, the non-negativity constraint on $p_t^* x_t^* \geq 0$ holds with equality for $t < t^*$, and is slack for $t \geq t^*$. Finally, the constraint on prices $p_t^* \leq 1$ holds with equality for $t \geq t^*$, and is slack for $t < t^*$.

Since fringe firms' IC constraints bind at all t, the first equalities in the Lemma hold for all t. The second equalities hold for all t since $x_t^* = 0$ for all $t < t^*$, and since $\mu_t^* = 0$ for all $t \geq t^*$. The third equalities hold since $\gamma_t^* = 0$ for all $t < t^*$, and since $p_t^* = 1$ for all $t \geq t^*$. $\square$

Note that the complementary slackness conditions in Lemma A.4, and the fact that (x^*, p^*) satisfies all the constraints in (P'_α), implies that for all $(\lambda, \mu, \gamma) \geq 0$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (12) hold.

I now show that the second set of inequalities in (12) also hold. Note that the Lagrangian

$\mathcal{L}(x, p, \lambda, \mu, \gamma)$ can be re-written as

$$\mathcal{L}(x, p, \lambda, \mu, \gamma) = \sum_{t=0}^{\infty} p_t x_t \left(-\alpha(n-1)\delta_1^t + (1-\alpha)(n-1)\delta^t + \sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau + \mu_t \right)$$
$$+ \sum_{t=0}^{\infty} p_t(\alpha\delta_1^t - \lambda_t - \gamma_t) + \sum_{t=0}^{\infty} \gamma_t.$$

Using Lemma A.3(ii) and $\gamma_t^* = \alpha\delta_1^t - \lambda_t^*$ for all t, it follows that, for all (x, p), $\mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*) = \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*) = \sum_{t=0}^{\infty} \gamma_t^*$. Hence, the second set of inequalities in (12) also hold, and so (x^*, p^*) solves (P'_α).

Lastly, I show that (x^*, p^*) also satisfies the constraints in program (P_α) that I ignored. Clearly, $p_t^* \geq 0$ for all t. Moreover, note that $x_t^* = 0$ for all $t < t^*$, and $x_t^* = 1 - \delta$ for all $t \geq t^*$. By Assumption 1, $(n-1)(1-\delta) < \delta_1 < 1$. Hence, for all t, $x_t^* < 1/(n-1)$. To see that (x^*, p^*) also satisfies the incentive compatibility constraints of firm 1, consider $t < t^*$, and note that

$$(n-1)x_t^* p_t^* = 0 < \sum_{s=t+1}^{\infty} \delta_1^{s-t}(1 - (n-1)x_s^*)p_s^*,$$

where the equality uses $x_t^* = 0$ for all $t < t^*$. For $t \geq t^*$, $x_t^* = 1 - \delta$ and $p_t^* = 1$, and so

$$(n-1)x_t^* p_t^* = (n-1)(1-\delta) < \frac{\delta_1}{1-\delta_1}(1 - (n-1)(1-\delta)) = \sum_{s=t+1}^{\infty} \delta_1^{s-t}(1 - (n-1)x_s^*)p_s^*,$$

where the strict inequality follows from Assumption 1. Hence, (x^*, p^*) also solves (P_α). ■

Proof of Proposition 4. Fix $\alpha \in [0, 1/2)$, and consider the following relaxation to program (P_α):

$$\widehat{V}(\alpha) \equiv \sup_{(x_t,p_t)_{t=0}^{\infty}} \sum_{t=0}^{\infty} \alpha\delta_1^t(1-(n-1)x_t)p_t + \sum_{t=0}^{\infty}(1-\alpha)(n-1)\delta^t x_t p_t \tag{P$''_\alpha$}$$

$$\text{s.t. } \forall t, \quad (1-x_t)p_t \le \sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s \tag{IC-$i \neq 1$}$$

$$\forall t, \quad (n-1)x_t p_t \le \sum_{s=t+1}^{\infty} \delta_1^{s-t}(1-(n-1)x_s)p_s \tag{IC-1}$$

$$\forall t, \quad p_t x_t \le p_t/(n-1), p_t \le 1. \tag{Feasibility'}$$

Relative to program (P_α), (P''_α) replaces the constraints that $p_t \ge 0$ and $x_t \in [0, 1/(n-1)]$ with the weaker constraints $p_t x_t \le p_t/(n-1)$. The Lagrangian associated to (P''_α) is

$$\begin{aligned}
\mathcal{L}(x,p,\lambda,\mu,\gamma) = &\sum_{t=0}^{\infty} \alpha\delta_1^t \left(1-(n-1)x_t\right) p_t + \sum_{t=0}^{\infty}(1-\alpha)(n-1)\delta^t x_t p_t \\
&+ \sum_{t=0}^{\infty} \lambda_{f,t} \left(\sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s - (1-x_t)p_t \right) \\
&+ \sum_{t=0}^{\infty} \lambda_{1,t} \left(\sum_{s=t+1}^{\infty} \delta_1^{s-t}(1-(n-1)x_s)p_s - (n-1)x_t p_t \right) \\
&+ \sum_{t=0}^{\infty} \mu_t p_t(1-(n-1)x_t) + \sum_{t=0}^{\infty} \gamma_t(1-p_t),
\end{aligned}$$

To prove the result, I construct multipliers $(\lambda^*, \mu^*, \gamma^*) \ge 0$ that, together with the allocations and prices (x^*, p^*) in the statement of the Proposition, satisfy the saddle-point condition (12). As explained in Section 4.4, if (x^*, p^*) satisfies the constraints in (P''_α), this implies that (x^*, p^*) solves the relaxed program (P''_α). In addition, if (x^*, p^*) also satisfies the remaining constraints in (P_α), then it also solves (P_α).

Note that the Lagrangian can be re-written as follows:

$$\mathcal{L}(x,p,\lambda,\mu,\gamma) = \sum_{t=0}^{\infty} p_t x_t \left[(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \mu_t - \sum_{\tau=0}^{t}\delta_1^{t-\tau}\lambda_{1,\tau}\right) + \sum_{\tau=0}^{t}\delta^{t-\tau}\lambda_{f,\tau}\right]$$
$$+ \sum_{t=0}^{\infty} p_t \left(\alpha\delta_1^t - \lambda_{f,t} - \gamma_t + \sum_{\tau=0}^{t-1}\delta_1^{t-\tau}\lambda_{1,\tau} + \mu_t\right) + \sum_{t=0}^{\infty}\gamma_t. \tag{13}$$

Recall the definitions of $t_1^*(\alpha)$ and t_2^* in the main text. Recall further that, for $\alpha \in [0,1/2)$, $t^*(\alpha) = \min\{t_1^*(\alpha), t_2^*\}$. There are two cases to consider: (1) $t_2^* \leq t_1^*(\alpha)$, so that $t^*(\alpha) = t_2^*$, and (2) $t_2^* > t_1^*(\alpha)$, so that $t^*(\alpha) = t_1^*(\alpha)$. I start with case (1). To economize on notation, from now on I write t_1^* and t^* instead of $t_1^*(\alpha)$ and $t^*(\alpha)$.

Case 1: $t^* = t_2^* \leq t_1^*$. Consider the following multipliers $(\lambda^*, \mu^*, \gamma^*)$:

$$\lambda_{1,t}^* = \mathbf{1}_{t=0}\left((1-\alpha)\left(\frac{\delta}{\delta_1}\right)^{t^*} - \alpha\right),$$

$$\mu_t^* = \mathbf{1}_{t<t^*}(1-\alpha)\delta_1^t\left(\left(\frac{\delta}{\delta_1}\right)^{t} - \left(\frac{\delta}{\delta_1}\right)^{t^*}\right),$$

$$\gamma_t^* = \begin{cases} \alpha + (1-\alpha)\left(1-\left(\frac{\delta}{\delta_1}\right)^{t^*}\right) & \text{if } t=0, \\ (1-\alpha)\delta^t & \text{if } t \in (0,t^*] \text{ and } t^*>0, \\ (1-\alpha)\delta^{t^*}\delta_1^{t-t^*-1}((n-1)\delta-(n-2)\delta_1) & \text{if } t>t^*, \end{cases}$$

$$\lambda_{f,t}^* = \begin{cases} 0 & \text{if } t \leq t^*, \\ (n-1)(1-\alpha)\delta^{t^*}\delta_1^{t-t^*-1}(\delta_1-\delta) & \text{if } t \geq t^*+1. \end{cases}$$

I use the following Lemma:

Lemma A.5. *(i)* $(\lambda^*, \mu^*, \gamma^*) \geq 0$*;*

(ii) for all t, $(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \sum_{\tau=0}^{t}\delta_1^{t-\tau}\lambda_{1,\tau}^* - \mu_t^*\right) + \sum_{\tau=0}^{t}\lambda_{f,\tau}^*\delta^{t-\tau} = 0$*;*

(iii) for all t, $\alpha\delta_1^t - \lambda_{f,t}^* + \sum_{\tau=0}^{t-1}\delta_1^{t-\tau}\lambda_{1,\tau}^* + \mu_t^* - \gamma_t^* = 0$.

Proof. I start with part (i). Clearly, $\lambda^*_{1,t} \geq 0$ for all $t > 0$. For $t = 0$, $\lambda^*_{1,0} = \frac{1}{\delta_1^{t^*}}\left((1-\alpha)\delta^{t^*} - \alpha\delta_1^{t^*}\right)$, which is strictly larger than zero since $0 \leq t^* \leq t_1^*$, and since $(1-\alpha)\delta^t - \alpha\delta_1^t > 0$ for all $t \leq t_1^*$.

Next, note that $\delta_1 \geq \delta$ implies $\left(\frac{\delta}{\delta_1}\right)^t \geq \left(\frac{\delta}{\delta_1}\right)^{t^*}$ for all $t < t^*$. Hence, $\mu^*_t \geq 0$ for all t. Clearly, $\gamma^*_t > 0$ for all $t \leq t^*$. Moreover, since Assumption 1 implies $(n-1)\delta > (n-2)\delta_1$, $\gamma^*_t > 0$ for all $t > t^*$. Lastly, $\delta_1 \geq \delta$ implies $\lambda^*_{f,t} \geq 0$ for all t.

I now turn to part (ii). Consider first $t < t^*$. Note that, for all such t, $\sum_{\tau=0}^{t} \lambda^*_{f,\tau}\delta^{t-\tau} = 0$, and $\sum_{\tau=0}^{t} \delta_1^{t-\tau}\lambda^*_{1,\tau} = \delta_1^t \lambda^*_{1,0} = \delta_1^t\left((1-\alpha)\left(\frac{\delta}{\delta_1}\right)^{t^*} - \alpha\right)$. Hence, for all $t < t^*$,

$$
\begin{aligned}
&(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \sum_{\tau=0}^{t}\delta_1^{t-\tau}\lambda^*_{1,\tau} - \mu^*_t\right) + \sum_{\tau=0}^{t}\lambda^*_{f,\tau}\delta^{t-\tau} \\
=&(n-1)\left((1-\alpha)\delta^t - \delta_1^t\left((1-\alpha)\left(\frac{\delta}{\delta_1}\right)^{t^*}\right) - \mu^*_t\right) = 0.
\end{aligned}
$$

For $t = t^*$, we have $\sum_{\tau=0}^{t^*}\delta_1^{t-\tau}\lambda^*_{1,\tau} = \delta_1^{t^*}\lambda^*_{1,0} = (1-\alpha)\delta^{t^*} - \alpha\delta_1^{t^*}$, $\mu^*_{t^*} = 0$, and $\sum_{\tau=0}^{t^*}\delta^{t-\tau}\lambda^*_{f,\tau} = 0$. Hence, for $t = t^*$,

$$
(n-1)\left(-\alpha\delta_1^{t^*} + (1-\alpha)\delta^{t^*} - \sum_{\tau=0}^{t^*}\delta_1^{t^*-\tau}\lambda^*_{1,\tau} - \mu^*_{t^*}\right) + \sum_{\tau=0}^{t^*}\lambda^*_{f,\tau}\delta^{t^*-\tau} = 0
$$

Lastly, I show that the equality holds for all $t > t^*$. To show this, I first show that, for all $t > t^*$, $\sum_{\tau=0}^{t}\delta^{t-\tau}\lambda^*_{f,\tau} = (n-1)(1-\alpha)\delta^{t^*}(\delta_1^{t-t^*} - \delta^{t-t^*})$. To see why, note that, for $t = t^* + 1$,

$$
\sum_{\tau=0}^{t^*+1}\delta^{t^*+1-\tau}\lambda^*_{f,\tau} = \lambda^*_{f,t^*+1} = (n-1)(1-\alpha)\delta^{t^*}(\delta_1 - \delta),
$$

so the equality holds for $t = t^* + 1$. Towards an induction, suppose that the equality holds to all $s = t^* + 1, ..., t-1$. Note then that

$$\begin{aligned}\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda^*_{f,\tau} = \lambda^*_{f,t} + \delta \sum_{\tau=0}^{t-1} \delta^{t-1-\tau}\lambda^*_{f,\tau} \\ &= (n-1)(1-\alpha)\delta^{t^*}\delta_1^{t-t^*-1}(\delta_1 - \delta) + \delta(n-1)(1-\alpha)\delta^{t^*}(\delta_1^{t-1-t^*} - \delta^{t-1-t^*}) \\ &= (n-1)(1-\alpha)\delta^{t^*}\left(\delta_1^{t-t^*} - \delta^{t-t^*}\right),\end{aligned}$$

where the second equality uses the definition of $\lambda^*_{f,t}$ and the induction hypothesis. Hence, for all $t > t^*$, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda^*_{f,\tau} = (n-1)(1-\alpha)\delta^{t^*}(\delta_1^{t-t^*} - \delta^{t-t^*})$. Using this equality, together with $\sum_{\tau=0}^{t} \delta_1^{t-\tau}\lambda^*_{1,\tau} = \delta_1^t\lambda_{1,0} = (1-\alpha)\delta_1^{t-t^*}\delta^{t^*} - \alpha\delta_1^t$ and $\mu^*_t = 0$ for all $t > t^*$, we have

$$\begin{aligned}&(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \sum_{\tau=0}^{t}\delta_1^{t-\tau}\lambda^*_{1,\tau}\right) + \sum_{\tau=0}^{t}\lambda^*_\tau\delta^{t-\tau} \\ =&(n-1)\left((1-\alpha)\delta^t - (1-\alpha)\delta^{t^*}\delta_1^{t-t^*}\right) + (n-1)(1-\alpha)\delta^{t^*}(\delta_1^{t-t^*} - \delta^{t-t^*}) = 0.\end{aligned}$$

Lastly, I prove part (iii). Consider first $t = 0$, and note that

$$\begin{aligned}&\alpha\delta_1^t - \lambda^*_{f,t} + \sum_{\tau=0}^{t-1}\delta_1^{t-\tau}\lambda^*_{1,\tau} + \mu^*_t - \gamma^*_t \\ =&\alpha + (1-\alpha)\left(1 - \left(\frac{\delta}{\delta_1}\right)^{t^*}\right) - \alpha - (1-\alpha)\left(1 - \left(\frac{\delta}{\delta_1}\right)^{t^*}\right) = 0.\end{aligned}$$

Consider first $t \in (0, t^*]$, and note that

$$\begin{aligned}&\alpha\delta_1^t - \lambda^*_{f,t} + \sum_{\tau=0}^{t-1}\delta_1^{t-\tau}\lambda^*_{1,\tau} + \mu^*_t - \gamma^*_t \\ =&\alpha\delta_1^t + \delta_1^t\left((1-\alpha)\left(\frac{\delta}{\delta_1}\right)^{t^*} - \alpha\right) + (1-\alpha)\delta_1^t\left(\left(\frac{\delta}{\delta_1}\right)^{t} - \left(\frac{\delta}{\delta_1}\right)^{t^*}\right) - (1-\alpha)\delta^t = 0.\end{aligned}$$

Consider next $t = t^*$, and note that

$$\begin{aligned}
&\alpha\delta_1^{t^*} - \lambda^*_{f,t^*} + \sum_{\tau=0}^{t^*-1} \delta_1^{t^*-\tau}\lambda^*_{1,\tau} + \mu^*_{t^*} - \gamma^*_{t^*} \\
=&\alpha\delta_1^{t^*} + \delta_1^{t^*}\left((1-\alpha)\left(\frac{\delta}{\delta_1}\right)^{t^*} - \alpha\right) - (1-\alpha)\delta^{t^*} = 0.
\end{aligned}$$

Lastly, consider $t > t^*$, and note that

$$\begin{aligned}
&\alpha\delta_1^{t} - \lambda^*_{f,t} + \sum_{\tau=0}^{t-1} \delta_1^{t-\tau}\lambda^*_{1,\tau} + \mu^*_{t} - \gamma^*_{t} \\
=&\alpha\delta_1^{t} - (n-1)(1-\alpha)\delta^{t^*}\delta_1^{t-t^*-1}(\delta_1-\delta) + \delta_1^t\left((1-\alpha)\left(\frac{\delta}{\delta_1}\right)^{t^*} - \alpha\right) \\
&- (1-\alpha)\delta^{t^*}\delta_1^{t-t^*-1}((n-1)\delta - (n-2)\delta_1) = 0.
\end{aligned}$$

□

Let (x^*, p^*) be the allocations and prices in the statement of the Proposition: i.e., $p^*_t = 1$ for all t, $x^*_t = 1/(n-1)$ for $t < t^*$, and $x_t = 1-\delta$ for $t > t^*$. For $t = t^*$, let $x^*_t = x$ be such that

$$\delta_1^{t^*}\left((1-(n-1)x) + \frac{\delta_1}{1-\delta_1}(1-(n-1)(1-\delta))\right) = 1.$$

Note that $x \in (1-\delta, 1/(n-1))$. To see why, note that for $x \geq 1/(n-1)$

$$\delta_1^{t^*}\left((1-(n-1)x) + \frac{\delta_1}{1-\delta_1}(1-(n-1)(1-\delta))\right) \leq \frac{\delta_1^{t^*+1}}{1-\delta_1}(1-(n-1)(1-\delta))) < 1,$$

where the strict inequality follows from the definition of t_2^* and from $t_2^* = t^*$ (since I'm considering the case with $t_2^* \leq t_1^*$). On the other hand, for $x < 1-\delta$

$$\delta_1^{t^*}\left((1-(n-1)x) + \frac{\delta_1}{1-\delta_1}(1-(n-1)(1-\delta))\right) > \frac{\delta_1^{t^*}}{1-\delta_1}(1-(n-1)(1-\delta))) \geq 1,$$

where the last inequality again follows from the definition of t_2^* and from $t_2^* = t^*$. Hence, it must be that $x \in (1-\delta, 1/(n-1))$.

I use the following Lemma, which shows that (x^*, p^*) satisfies all the constraints in program (P''_α), and that (x^*, p^*) and $(\lambda^*, \mu^*, \gamma^*)$ satisfy complementary slackness conditions.

Lemma A.6. *(x^*, p^*) satisfies the constraints in program* (P''_α)*. Moreover, for all t,*

$$\lambda^*_{f,t}\left(\sum_{s=t+1}^{\infty} \delta^{s-t} x^*_s p^*_s - (1-x^*_t)p^*_t\right) = 0,$$

$$\lambda^*_{1,t}\left(\sum_{s=t+1}^{\infty} \delta_1^{s-t}(1-(n-1)x^*_s)p^*_s - (n-1)x^*_t p^*_t\right) = 0,$$

$$\mu^*_t p^*_t (1-(n-1)x^*_t) = 0,$$

$$\gamma^*_t(1-p_t) = 0.$$

Proof. Note first that, under (x^*, p^*), firm 1's IC constraint binds at $t = 0$, and it is slack at any $t > 0$. Indeed, for all $t < t^*$,

$$\sum_{s=t+1}^{\infty} \delta_1^{s-t}(1-(n-1)x^*_s)p^*_s = \delta_1^{t^*-t}\left(1-(n-1)x + \frac{\delta_1}{1-\delta_1}(1-(n-1)(1-\delta))\right) \geq 1 = (n-1)x^*_t p^*_t,$$

with equality at $t = 0$ and strict inequality for all $t \in (0, t^* - 1)$. For $t \geq t^*$,

$$\sum_{s=t+1}^{\infty} \delta_1^{s-t}(1-(n-1)x^*_s)p^*_s = \frac{\delta_1}{1-\delta_1}(1-(1-\delta)(n-1)) > (n-1)(1-\delta) = (n-1)x^*_t p^*_t,$$

where the strict inequality follows from Assumption 1.

Next, note that the constraint $p^*_t x^*_t = x^*_t \leq 1/(n-1) = p^*_t/(n-1)$ is satisfied for all t (since, by Assumption 1, $1 - \delta < 1/(n-1)$), with equality for all $t < t^*$. The constraint that $p^*_t \leq 1$ is satisfied with equality at all t. Lastly, the IC constraints of fringe firms are satisfied for all $t \geq 0$, and bind for all $t > t^*$. Indeed, for all $t \leq t^*$,

$$\sum_{s=t+1}^{\infty} \delta^{s-t} x^*_s p^*_s \geq \sum_{s=t+1}^{\infty} \delta^{s-t}(1-\delta) = \delta \geq (1-x^*_t)p^*_t$$

where the first inequality follows since $x_s^* = 1/(n-1)$ for all $s < t^*$, $x_{t^*}^* = x \geq 1-\delta$, and $x_s^* = 1-\delta$ for all $s > t^*$ and since Assumption 1 implies $1-\delta < 1/(n-1)$, and the last inequality follows since $x_s^* \geq 1-\delta$ for all $s \leq t^*$. For $t > t^*$, we have that

$$\sum_{s=t+1}^{\infty} \delta^{s-t} x_s^* p_s^* = \frac{\delta}{1-\delta}(1-\delta) = \delta = (1-x_t^*)p_t^*.$$

Hence, (x^*, p^*) satisfies the constraints in program (P'_α).

Next, note that the first equalities hold for all $t \leq t^*$, since $\lambda_{f,t}^* = 0$ for all such t, and also holds for all $t > t^*$ since fringe firms' incentive constraints bind for all such t under (x^*, p^*).

Consider next the second set of equalities. Clearly, the equalities hold for all $t > 0$, since $\lambda_{1,t}^* = 0$ for all such t. For $t = 0$ the equality also holds, since firm 1's incentive constraint binds at $t = 0$.

The third set of equalities holds since $x_t^* = 1/(n-1)$ for all $t < t^*$, and since $\mu_t^* = 0$ for all $t \geq t^*$. Lastly, the fourth set of equalities holds since $p_t^* = 1$ for all t. $\square$

The complementary slackness conditions in Lemma A.6, and the fact that (x^*, p^*) satisfies all the constraints in (P''_α), implies that for all $(\lambda, \mu, \gamma) \geq 0$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (12) hold.

To see that the second set of inequalities in (12) also hold, note that equation (13) and Lemma A.5(ii)-(iii) imply that, for all (x, p), $\mathcal{L}(x^*, p^*, \lambda^*, \gamma^*, \mu^*) = \mathcal{L}(x, p, \lambda^*, \gamma^*, \mu^*) = \sum_{t=0}^{\infty} \gamma_t^*$. Hence, (x^*, p^*) solves (P''_α). Finally, it is clear that (x^*, p^*) also satisfies the remaining constraints (i.e., $p_t^* \geq 0$ for all t and $x_t^* \geq 0$ for all t), and so (x^*, p^*) also solves (P_α). This completes the proof of Case 1 (i.e., $t_2^* \leq t_1^*$).

Case 2: $t_2^* > t_1^* = t^*$**.** Consider the following multipliers $(\lambda^*, \mu^*, \gamma^*)$. For all t,

$$
\begin{aligned}
\lambda_{1,t}^* &= 0, \\
\mu_t^* &= \mathbf{1}_{t \le t^*} \left(-\alpha\delta_1^t + (1-\alpha)\delta^t\right), \\
\gamma_t^* &= \begin{cases} (1-\alpha)\delta^t & \text{if } t \le t^*, \\ \alpha\delta_1^t - (n-1)(\alpha\delta_1^t - (1-\alpha)\delta^t) & \text{if } t = t^*+1, \\ \alpha\delta_1^t - (n-1)\alpha\delta_1^{t-1}(\delta_1 - \delta) & \text{if } t > t^*+1, \end{cases} \\
\lambda_{f,t}^* &= \begin{cases} 0 & \text{if } t \le t^*, \\ (n-1)(\alpha\delta_1^t - (1-\alpha)\delta^t) & \text{if } t = t^*+1 \\ (n-1)\alpha\delta_1^{t-1}(\delta_1 - \delta) & \text{if } t > t^*+1. \end{cases}
\end{aligned}
$$

I use the following Lemma:

Lemma A.7. *(i)* $(\lambda^*, \mu^*, \gamma^*) \ge 0$*;*

(ii) for all t, $(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \sum_{\tau=0}^t \delta_1^{t-\tau}\lambda_{1,\tau}^* - \mu_t^*\right) + \sum_{\tau=0}^t \lambda_{f,\tau}^*\delta^{t-\tau} = 0$*;*

(iii) for all t, $\alpha\delta_1^t - \lambda_{f,t}^* + \sum_{\tau=0}^{t-1} \delta_1^{t-\tau}\lambda_{1,\tau}^* + \mu_t^* - \gamma_t^* = 0$.

Proof. Part (i). Clearly, $\lambda_{1,t}^* \ge 0$ for all t. That $\mu_t^* \ge 0$ for all t follows since $-\alpha\delta_1^t + (1-\alpha)\delta^t > 0$ for all $t \le t_1^*$, and since we are considering the case with $t_2^* > t_1^* = t_1^*$.

Next, note that $\lambda_{f,t}^* \ge 0$ for all $t \ne t^*+1$. Moreover, since $-\alpha\delta_1^t + (1-\alpha)\delta^t \le 0$ for all $t > t_1^* = t^*$, it follows that $\lambda_{f,t^*+1}^* \ge 0$.

Next, I show that $\gamma_t^* \ge 0$ for all t. The inequality clearly holds for all $t \le t^*$. Consider next $t = t^*+1 = t_1^*+1$ (since we are in the case with $t^* = t_1^* < t_2^*$). Let $\beta \equiv \frac{\delta_1}{\delta} \ge 1$, and note that

$$(1-\alpha)\delta^{t^*} - \alpha\delta_1^{t^*} \ge 0 \Longleftrightarrow \alpha\beta^{t^*} \le 1-\alpha, \tag{14}$$

where the first inequality follows since $(1-\alpha)\delta^t - \alpha\delta_1^t \ge 0$ for all $t \le t_1^*$ and since $t^* = t_1^*$.

Then,

$$\begin{aligned}\gamma^*_{t^*+1} &= (1-\alpha)(n-1)\delta^{t^*+1} - \alpha(n-2)\delta_1^{t^*+1} \\ &= \delta^{t^*+1}((1-\alpha)(n-1) - \alpha(n-2)\beta^{t^*+1}) \\ &\geq \delta^{t^*+1}((1-\alpha)(n-1) - (1-\alpha)(n-2)\beta) > 0,\end{aligned}$$

where the weak inequality follows since $\alpha\beta^{t^*} \leq 1-\alpha$ (by inequality (14)), and the strict inequality follows since Assumption 1 implies $(n-1)\delta - (n-2)\delta_1 > 0$, and so $(n-1)-(n-2)\beta > 0$. For $t > t^*$,

$$\gamma^*_t = \alpha\delta_1^{t-1}((n-1)\delta - (n-2)\delta_1) > 0,$$

where the strict inequality follows from Assumption 1.

I now turn to part (ii). Consider first $t \leq t^*$. Since $\lambda^*_{1,\tau} = \lambda^*_{f,\tau} = 0$ for all $\tau \leq t^*$, it follows that, for all $t \leq t^*$,

$$\begin{aligned}&(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \sum_{\tau=0}^{t}\delta_1^{t-\tau}\lambda^*_{1,\tau} - \mu^*_t\right) + \sum_{\tau=0}^{t}\lambda^*_{f,\tau}\delta^{t-\tau} \\ =&(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \mu^*_t\right) = 0,\end{aligned}$$

where I used $\mu^*_t = -\alpha\delta_1^t + (1-\alpha)\delta^t$ for all $t \leq t^*$.

Consider next $t > t^*$. Note first that, for all $t > t^*$, $\sum_{\tau=0}^{t}\lambda^*_{f,\tau}\delta^{t-\tau} = (n-1)(\alpha\delta_1^t - (1-\alpha)\delta^t)$. To see why, note that the equality holds for $t = t^*+1$, since $\lambda^*_{f,\tau} = 0$ for all $t \leq t^*$ and since $\lambda^*_{f,t^*+1} = (n-1)(\alpha\delta_1^{t^*+1} - (1-\alpha)\delta^{t^*+1})$. Towards an induction, suppose that the equality holds for all $s = t^*+1, ..., t-1$. Then,

$$\begin{aligned}\sum_{\tau=0}^{t}\lambda^*_{f,\tau}\delta^{t-\tau} &= \lambda^*_{f,t} + \delta\sum_{\tau=0}^{t-1}\lambda^*_{f,\tau}\delta^{t-1-\tau} \\ &= (n-1)\alpha\delta_1^{t-1}(\delta_1 - \delta) + \delta(n-1)(\alpha\delta_1^{t-1} - (1-\alpha)\delta^{t-1}) \\ &= (n-1)(\alpha\delta_1^t - (1-\alpha)\delta^t).\end{aligned}$$

Hence, for all $t > t^*$, $\sum_{\tau=0}^{t} \lambda^*_{f,\tau}\delta^{t-\tau} = (n-1)(\alpha\delta_1^t - (1-\alpha)\delta^t)$. Then, for all $t > t^*$,

$$\begin{aligned}
&(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t - \sum_{\tau=0}^{t} \delta_1^{t-\tau}\lambda^*_{1,\tau} - \mu^*_t\right) + \sum_{\tau=0}^{t} \lambda^*_{f,\tau}\delta^{t-\tau} \\
=&(n-1)\left(-\alpha\delta_1^t + (1-\alpha)\delta^t\right) + \sum_{\tau=0}^{t} \lambda^*_{f,\tau}\delta^{t-\tau} = 0.
\end{aligned}$$

I turn prove part (iii). For all $t \leq t^*$,

$$\begin{aligned}
&\alpha\delta_1^t - \lambda^*_{f,t} + \sum_{\tau=0}^{t-1} \delta_1^{t-\tau}\lambda^*_{1,\tau} + \mu^*_t - \gamma^*_t \\
=&\alpha\delta_1^t - \alpha\delta_1^t + (1-\alpha)\delta^t - (1-\alpha)\delta^t = 0.
\end{aligned}$$

For $t = t^* + 1$,

$$\begin{aligned}
&\alpha\delta_1^t - \lambda^*_{f,t} + \sum_{\tau=0}^{t-1} \delta_1^{t-\tau}\lambda^*_{1,\tau} + \mu^*_t - \gamma^*_t \\
=&\alpha\delta_1^t - (n-1)(\alpha\delta_1^t - (1-\alpha)\delta_2^t) - (\alpha\delta_1^t - (n-1)(\alpha\delta_1^t - (1-\alpha)\delta^t)) = 0.
\end{aligned}$$

Finally, for $t > t^*$,

$$\begin{aligned}
&\alpha\delta_1^t - \lambda^*_{f,t} + \sum_{\tau=0}^{t-1} \delta_1^{t-\tau}\lambda^*_{1,\tau} + \mu^*_t - \gamma^*_t \\
=&\alpha\delta_1^t - (n-1)\alpha\delta_1^{t-1}(\delta_1 - \delta) - (\alpha\delta_1^t - (n-1)\alpha\delta_1^{t-1}(\delta_1 - \delta)) = 0.
\end{aligned}$$

□

Let (x^*, p^*) be the allocations and prices in the statement of the Proposition: i.e., $p^*_t = 1$ for all t, $x^*_t = 1/(n-1)$ for all $t \leq t^*$, and $x^*_t = (1-\delta)$ for all $t > t^*$. I use the following Lemma, which shows that (x^*, p^*) satisfies all the constraints in program (P''_α), and that (x^*, p^*) and $(\lambda^*, \mu^*, \gamma^*)$ satisfy complementary slackness conditions.

Lemma A.8. *(x^*, p^*) satisfies the constraints in program (P''_α). Moreover, for all t,*

$$\lambda^*_{f,t}\left(\sum_{s=t+1}^{\infty}\delta^{s-t}x^*_s p^*_s-(1-x^*_t)p^*_t\right)=0,$$
$$\lambda^*_{1,t}\left(\sum_{s=t+1}^{\infty}\delta_1^{s-t}(1-(n-1)x^*_s)p^*_s-(n-1)x^*_t p^*_t\right)=0,$$
$$\mu^*_t p^*_t(1-(n-1)x^*_t)=0,$$
$$\gamma^*_t(1-p_t)=0.$$

Proof. Note first that, under (x^*, p^*), firm 1's IC constraints are satisfied for all $t\geq 0$. Indeed, for all $t\leq t^*$,

$$\sum_{s=t+1}^{\infty}\delta_1^{s-t}(1-(n-1)x^*_s)p^*_s=\frac{\delta_1^{t^*+1-t}}{1-\delta_1}(1-(n-1)(1-\delta))\geq 1=(n-1)x^*_t p^*_t,$$

where the inequality follows since $\frac{\delta_1^t}{1-\delta_1}(1-(n-1)(1-\delta)))\geq 1$ for all $t\leq t^*_2$, and since $t^*_2\geq t^*_1+1=t^*+1$. For $t\geq t^*+1$,

$$\sum_{s=t+1}^{\infty}\delta_1^{s-t}(1-(n-1)x^*_s)p^*_s=\frac{\delta_1}{1-\delta_1}(1-(1-\delta)(n-1))>(n-1)(1-\delta)=(n-1)x^*_t p^*_t,$$

where the strict inequality follows from Assumption 1.

Next, note that the constraint $p^*_t x^*_t\leq p^*_t/(n-1)$ is satisfied for all t (since, by Assumption 1, $1-\delta<1/(n-1)$), with equality for all $t\leq t^*$. The constraint that $p^*_t\leq 1$ is satisfied with equality at all t. Lastly, the IC constraints of fringe firms are satisfied with slack for all $t\leq t^*$, and are satisfied with equality for all $t>t^*$. Indeed, for all $t\leq t^*$,

$$\sum_{s=t+1}^{\infty}\delta^{s-t}x^*_s p^*_s\geq\sum_{s=t+1}^{\infty}\delta^{s-t}(1-\delta)=\delta>(1-x^*_t)p^*_t,$$

where the first equality follows since $x^*_s=1/(n-1)$ for all $s\leq t^*$ and $x^*_s=1-\delta$ for all $s>t^*$, and since Assumption 1 implies $1-\delta<1/(n-1)$, and the last inequality follows since $x_t^* = 1/(n-1) > 1-\delta$ for all $t \le t^*$. For $t > t^*$,

$$\sum_{s=t+1}^{\infty} \delta^{s-t} x_s^* p_s^* = \frac{\delta}{1-\delta}(1-\delta) = \delta = (1-x_t^*)p_t^*.$$

Hence, (x^*, p^*) satisfies the constraints in program (P''_α).

Next, note that (x^*, p^*) and $(\lambda^*, \mu^*, \gamma^*)$ satisfy the first set of equalities in the Lemma, since $\lambda_{f,t}^* = 0$ for all $t \le t^*$, and since fringe firms' IC constraints are satisfied with equality for all $t > t^*$. The second set of equalities also hold, since $\lambda_{1,t}^* = 0$ for all t. The third set of equalities hold since $x_t^* = 1/(n-1)$ for all $t \le t^*$, and since $\mu_t^* = 0$ for all $t > t^*$. The last set of equalities hold since $p_t^* = 1$ for all t. □

The complementary slackness conditions in Lemma A.8, and the fact that (x^*, p^*) satisfies all the constraints in (P''_α), implies that for all $(\lambda, \mu, \gamma) \ge 0$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \ge \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (12) hold.

To see that the second set of inequalities in (12) also hold, note that equation (13) and Lemma A.7(ii)-(iii) imply that, for all (x, p), $\mathcal{L}(x^*, p^*, \lambda^*, \gamma^*, \mu^*) = \mathcal{L}(x, p, \lambda^*, \gamma^*, \mu^*) = \sum_{t=0}^{\infty} \gamma_t^*$. Hence, (x^*, p^*) solves (P''_α). Finally, it is clear that (x^*, p^*) also satisfies the remaining constraints (i.e., $p_t^* \ge 0$ for all t and $x_t^* \ge 0$ for all t), and so (x^*, p^*) also solves (P_α). This completes the proof of Case 2 (i.e., $t_2^* > t_1^*$). ■

A.3 Proofs of Section 6

Proof of Proposition 5. Consider the following relaxation of Program (P_s):

$$\overline{V}_1 \equiv \sup_{(x,p)} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t) p_t \right] \tag{P'_s}$$
$$\text{s.t. } \forall t, \quad (1-x_t)p_t \le \mathbb{E}_t \left[\sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s \right]$$
$$\forall t, \quad p_t \le r_t, p_t x_t \ge 0.$$

The Lagrangian associated with this program is

$$\begin{aligned}\mathcal{L}(x,p,\lambda,\mu,\gamma) = \mathbb{E}_0 \Bigg[& \sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t)p_t + \sum_{t=0}^{\infty} \mathbb{E}_t \left[\lambda_t \left(\sum_{s=t+1}^{\infty} \delta^{s-t} x_s p_s - (1-x_t)p_t \right) \right] \\ & + \sum_{t=0}^{\infty} \mathbb{E}_t \left[\mu_t p_t x_t \right] + \sum_{t=0}^{\infty} \mathbb{E}_t \left[\gamma_t (r_t - p_t) \right] \Bigg] \\ = \mathbb{E}_0 & \left[\sum_{t=0}^{\infty} p_t x_t \left(-(n-1)\delta_1^t + \mu_t + \sum_{\tau=0}^{t} \delta^{t-\tau} \lambda_\tau \right) + \sum_{t=0}^{\infty} p_t \left(\delta_1^t - \lambda_t - \gamma_t \right) + \sum_{t=0}^{\infty} r_t \gamma_t \right], \end{aligned} \tag{15}$$

where the equality follows from applying the Law of Iterated Expectations and combining terms. I note that (x,p) and (λ,μ,γ) are now stochastic processes adapted to (r_t).

To establish the proposition, I construct multipliers $(\lambda^*,\mu^*,\gamma^*)$ that, together with the prices and allocations (p^*,x^*) in the statement of the proposition, satisfy the saddle-point condition,

$$\begin{aligned} & \forall(\lambda,\mu,\gamma) \geq 0, \forall(x,p) \\ & \quad \mathcal{L}(x^*,p^*,\lambda,\mu,\gamma) \geq \mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*) \geq \mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*). \end{aligned} \tag{16}$$

Consider the following (deterministic) multipliers $(\lambda^*,\mu^*,\gamma^*)$:

$$\begin{aligned} \lambda_t^* &= \begin{cases} \delta_1^t & \text{if } t < t^* \\ \delta_1^t (n-1) - \sum_{\tau=0}^{t-1} \delta_1^\tau \delta^{t-\tau} & \text{if } t = t^* \\ \delta_1^{t-1}(n-1)(\delta_1 - \delta) & \text{if } t > t^*, \end{cases} \\ \mu_t^* &= \mathbf{1}_{t<t^*} \left(\delta_1^t (n-1) - \sum_{\tau=0}^{t} \delta_1^\tau \delta^{t-\tau} \right) \\ \gamma_t^* &= \delta_1^t - \lambda_t^*. \end{aligned}$$

Note that these are the multipliers in the proof of Proposition 3 when $\alpha = 1$. Hence, the following Lemma follows immediately from Lemma A.3.

Lemma A.9. *(i)* $(\lambda^*, \mu^*, \gamma^*) \geq 0$*;*

(ii) For all $t \geq 0$*,* $\sum_{\tau=0}^{t} \lambda_\tau^* \delta^{t-\tau} = \delta_1^t(n-1) - \mu_t^*$*.*

Let (x^*, p^*) be the allocations and prices in the statement of the Proposition: i.e., $x_t^* = 0$ and $p_t^* = \delta^{t^*-t}$ for all $t < t^*$, and $x_t^* = x(r_t)$ and $p_t^* = r_t$ for all $t \geq t^*$. I use the following Lemma, which shows that (x^*, p^*) satisfies all the constraints in program (P'_s), and that (x^*, p^*) and $(\lambda^*, \mu^*, \gamma^*)$ satisfy complementary slackness conditions,

Lemma A.10. (x^*, p^*) *satisfies the constraints in program* (P'_s)*. Moreover, for all* t*,*

$$\mathbb{E}_0 \left[\lambda_t^* \left(\sum_{s=t+1}^{\infty} \delta^{s-t} x_s^* p_s^* - (1 - x_t^*) p_t^* \right) \right] = 0,$$

$$\mathbb{E}_0 \left[\mu_t^* p_t^* x_t^* \right] = 0,$$

$$\mathbb{E}_0 [\gamma_t^* (r_t - p_t)] = 0.$$

Proof. Note first that, under (x^*, p^*), incentive constraints for fringe firms bind for all t. Indeed, for all $t \geq t^*$, fringe firms' IC constraints bind by construction (see equation (8)). For all $t < t^*$,

$$\mathbb{E}_t \left[\sum_{s=t+1}^{\infty} \delta^{s-t} x_s^* p_s^* \right] = \delta^{t^*-t} \frac{1}{1-\delta} \sum q_k r_k x(r_k) = \delta^{t^*-t} \overline{r} = p_t^* = p_t^*(1 - x_t^*),$$

where I used $x(r_k) = \frac{r_k - \delta \overline{r}}{r_k}$ for all k and $\overline{r} = \sum_k q_k r_k$. Moreover, the non-negativity constraint on $p_t^* x_t^* \geq 0$ holds with equality for $t < t^*$, and is slack for $t \geq t^*$. Finally, the constraint on prices $p_t^* \leq r_t$ holds with equality for $t \geq t^*$, and is slack for $t < t^*$ (indeed, for all $t < t^*$, $p_t^* = \delta^{t^*-t} \overline{r} \leq \delta \overline{r} < r_1$, where the last inequality follows from Assumption 2).

Since fringe firms' IC constraints bind at all t, the first set of equalities holds for all t. The second set of equalities holds for all t since $x_t^* = 0$ for all $t < t^*$, and since $\mu_t^* = 0$ for all $t \geq t^*$. The third set of equalities hold since $\gamma_t^* = 0$ for all $t < t^*$, and since $p_t^* = r_t$ for all $t \geq t^*$. $\square$

The complementary slackness conditions in Lemma A.10, and the fact that (x^*, p^*) satisfies all the constraints in (P'_s), implies that for all $(\lambda, \mu, \gamma) \geq 0$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (16) hold.

Next, note that equation (15) and Lemma A.9 imply that, for all (x, p), $\mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*) = \mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*)$. Hence, the second set of inequalities in (16) holds, and so (x^*, p^*) solves (P'_s).

Lastly, I show that (x^*, p^*) also satisfies the constraints in program (P_s) that I ignored. Clearly, $p^*_t \geq 0$ for all t. Moreover, note that $x^*_t = 0$ for all $t < t^*$, and $x^*_t = x(r_t)$ for all $t \geq t^*$. By Assumption 2, $(n-1)x(r) < 1$ for all r, and so, for all t, $x^*_t < 1/(n-1)$. To see that (x^*, p^*) also satisfies the incentive compatibility constraints for firm 1, consider $t < t^*$, and note that

$$(n-1)x^*_t p^*_t = 0 < \mathbb{E}_t \left[\sum_{s=t+1}^{\infty} \delta_1^{s-t} (1 - (n-1)x^*_s) p^*_s \right],$$

where the equality uses $x^*_s = 0$ for all $s < t^*$ and $x^*_s = x(r_s) < 1/(n-1)$ for all $s \geq t^*$. For $t \geq t^*$, $x^*_t = x(r_t) = \frac{r_t - \delta \bar{r}}{r_t}$ and $p^*_t = r_t$, and so

$$(n-1)x^*_t p^*_t = (n-1)(r_t - \delta \bar{r}) < \frac{\delta_1}{1-\delta_1} \sum_k q_k r_k \left(1 - (n-1) \frac{r_k - \delta \bar{r}}{r_k} \right) = \mathbb{E}_t \left[\sum_{s=t+1}^{\infty} \delta_1^{s-t} (1 - (n-1)x_s) p_s \right],$$

where the strict inequality follows since

$$\frac{\delta_1}{1-\delta_1} \sum_k q_k r_k \left(1 - (n-1) \frac{r_k - \delta \bar{r}}{r_k} \right) = \frac{\delta_1}{1-\delta_1} \bar{r} ((n-1)\delta - (n-2)) > (n-1)(r_K - \delta \bar{r}),$$

where the inequality follows from Assumption 2(ii). Hence, (x^*, p^*) also solves (P_s). ■

References

Abreu, D. (1986): "Extremal equilibria of oligopolistic supergames," *Journal of Economic Theory*, 39, 191–225.

Abreu, D., B. Brooks, and Y. Sannikov (2020): "Algorithms for stochastic games with perfect monitoring," *Econometrica*, 88, 1661–1695.

Abreu, D., D. Pearce, and E. Stacchetti (1986): "Optimal cartel equilibria with imperfect monitoring," *Journal of Economic Theory*, 39, 251–269.

——— (1990): "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica: Journal of the Econometric Society*, 1041–1063.

Abreu, D. and Y. Sannikov (2014): "An algorithm for two-player repeated games with perfect monitoring," *Theoretical Economics*, 9, 313–338.

Athey, S. and K. Bagwell (2001): "Optimal collusion with private information," *RAND Journal of Economics*, 32, 428–465.

Athey, S., K. Bagwell, and C. Sanchirico (2004): "Collusion and price rigidity," *The Review of Economic Studies*, 71, 317–349.

Bernheim, B. D. and E. Madsen (2017): "Price cutting and business stealing in imperfect cartels," *American Economic Review*, 107, 387–424.

Boyd, S. and L. Vandenberghe (2004): *Convex optimization*, Cambridge university press.

Chassang, S. and J. Ortner (2019): "Collusion in auctions with constrained bids: Theory and evidence from public procurement," *Journal of Political Economy*, 127, 2269–2300.

Chen, B. and S. Takahashi (2012): "A folk theorem for repeated games with unequal discounting," *Games and Economic Behavior*, 76, 571–581.

Connor, J. M. (2007): *Global price fixing*, vol. 2, Springer.

Dasgupta, A. and S. Ghosh (2022): "Self-accessibility and repeated games with asymmetric discounting," *Journal of Economic Theory*, 200, 105312.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The folk theorem with imperfect public information," *Econometrica: Journal of the Econometric Society*, 997–1039.

Harrington, J. E. (1989): "Collusion among asymmetric firms: The case of different discount factors," *International Journal of Industrial Organization*, 7, 289–307.

——— (2004): "Cartel pricing dynamics in the presence of an antitrust authority," *RAND Journal of Economics*, 651–673.

——— (2005): "Optimal cartel pricing in the presence of an antitrust authority," *International Economic Review*, 46, 145–169.

——— (2008): "Detecting cartels," *Handbook of antitrust economics*, 213, 245.

Harrington, J. E. and A. Skrzypacz (2011): "Private monitoring and communication in cartels: Explaining recent collusive practices," *The American Economic Review*, 101, 2425–2449.

Judd, K. L., S. Yeltekin, and J. Conklin (2003): "Computing supergame equilibria," *Econometrica*, 71, 1239–1254.

Kawai, K., J. Nakabayashi, and J. Ortner (2022): "The Value of Privacy in Cartels: An Analysis of the Inner Workings of a Bidding Ring," .

Lehrer, E. and A. Pauzner (1999): "Repeated games with differential time preferences," *Econometrica*, 67, 393–412.

Mailath, G. J. and P. Zemsky (1991): "Collusion in second price auctions with heterogeneous bidders," *Games and Economic Behavior*, 3, 467–486.

Marcet, A. and R. Marimon (2019): "Recursive contracts," *Econometrica*, 87, 1589–1631.

Marshall, R. C. and L. M. Marx (2012): *The economics of collusion: Cartels and bidding rings*, Mit Press.

Martimort, D. (1999): "The life cycle of regulatory agencies: Dynamic capture and transaction costs," *The Review of Economic Studies*, 66, 929–947.

Obara, I. and F. Zincenko (2017): "Collusion and heterogeneity of firms," *The RAND Journal of Economics*, 48, 230–249.

Ortner, J., T. Sugaya, and A. Wolitzky (2024): "Mediated Collusion," *Journal of Political Economy*, 132, 1247–1289.

Ray, D. (2002): "The time structure of self-enforcing agreements," *Econometrica*, 70, 547–582.

Rotemberg, J. J. and G. Saloner (1986): "A supergame-theoretic model of price wars during booms," *The American Economic Review*, 390–407.

Skrzypacz, A. and H. Hopenhayn (2004): "Tacit collusion in repeated auctions," *Journal of Economic Theory*, 114, 153–169.

Sugaya, T. (2015): "Characterizing the limit set of perfect and public equilibrium payoffs with unequal discounting," *Theoretical Economics*, 10, 691–717.

Sugaya, T. and A. Wolitzky (2018): "Maintaining privacy in cartels," *Journal of Political Economy*, 126, 2569–2607.

——— (2024): "Collusion with Optimal Information Disclosure," .

United States v. United States Steel Corporation (1920): *251 U.S. 417*.

Watkins, M. W. (1961): "Electrical equipment antitrust cases: Their implications for government and for business," *The University of Chicago Law Review*, 29, 97–110.