Leader-Driven Collusion*

Juan Ortner

Boston University

August 19, 2026

Abstract

Cartels are often organized around one or more firms that act as leaders and coordinate collusion. Motivated by these instances, I study optimal repeated-game equilibria in a market with a cartel leader and several 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.

*Contact information: jortner@bu.edu.

For helpful comments, I thank Ben Brooks, Joe Harrington, Marc Rysman, Andy Skrzypacz, Takuo Sugaya, Alex Wolitzky and seminar participants at Boston University, Columbia University and the 2025 SITE workshop on Dynamic Games and Contracts.

1 Introduction

Cartels are often organized around one or more firms that act as leaders and coordinate 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).

Motivated by these instances of collusion, I study optimal repeated-game equilibria in a market with one firm, which I call the "cartel leader," and several competitors. I show that Pareto-optimal equilibria that put enough weight on the leader may involve an initial phase with gradually increasing prices. Moreover, play under Pareto-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 discount factor. One possible interpretation is that the cartel leader is the largest firm in the industry, and may face 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 concerns the equilibrium that maximizes the leader's payoff. As an initial step, I characterize a leader-optimal equilibrium under the constraint that, along the path of play, prices are constant at all periods. I show that this equilibrium features constant market 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

[1] The Online Appendix studies an extension in which the cartel leader has a cost advantage.

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.

The leader-optimal equilibrium is 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 concerns Pareto optimal equilibria that maximize the weighted sum of the leader's and competitors' profits. When the weight on the leader is sufficiently large, Pareto optimal equilibria retain the same key properties of the leader-optimal equilibrium: there is an initial phase with gradually increasing prices and with the leader capturing the entire market, and a second phase with stationary prices and stationary market shares. When the weight on the leader is small, Pareto optimal 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 stationary phase.

I 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 (demand-contingent) prices and market shares. In contrast

to the symmetric equilibrium in Rotemberg and Saloner (1986), the stationary phase of the leader-optimal equilibrium exhibits pro-cyclical variation in the market shares of the less patient firms: these 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 that 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, this approach could potentially be used more broadly to compute bounds on the set of equilibrium payoffs. In contrast to the recursive techniques standard in repeated games (Abreu et al., 1990), the Lagrangian approach allows one to compute Pareto optimal equilibria in any direction without the need to derive the entire frontier, making the analysis tractable and computationally less demanding.

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. Proofs are in the Appendix and the Online Appendix. The Online Appendix also studies extensions to settings with asymmetric costs, capacity constraints, and differentiated products.

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 by Obara and Zincenko (2017), who also study a repeated Bertrand oligopoly game with

[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), Ortner et al. (2024), Kawai et al. (2026), Sugaya and Wolitzky (2026).

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

Because my model allows the leader to be more patient than its competitors, 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 papers are Pavoni et al. (2018), Marcet and Marimon (2019) and Rahman (2025), who use duality techniques to study dynamic contracts and repeated games. Martimort (1999) 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).

2 Model

Stage game. Consider a market with firms $N = \{1, \ldots, n\}$ producing a homogeneous good and competing in prices. All firms share the same constant marginal cost c, which I normalize to $c = 0$. 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 $p_i < p_j$

[3] Harrington (1989) establishes the same result, restricting attention to stationary equilibria.

[4] Section 5 discusses Obara and Zincenko (2017) in more detail.

[5] Section 4 relates my results to those in Lehrer and Pauzner (1999). See also Dasgupta and Ghosh (2022), who introduce the notion of *self-accessibility* to study repeated games with unequal discounting.

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 at price $\min_k p_k \leq 1$, 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 $p_j = \min_k p_k$.[6] 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 homogeneous. 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 requires firms' average discount factor to be strictly larger than $(n-1)/n$. As shown by Obara and Zincenko (2017), no collusion is possible when 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 $\sigma_i : H \to \mathbb{R}_+ \times [0,1]$. For each strategy profile $\sigma = (\sigma_i)_{i \in N}$, 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 denote the set of pure strategy subgame perfect Nash equilibria (SPNE). The goal of the paper is to characterize Pareto-optimal equilibria in Σ. In particular, for each

[6] Athey and Bagwell (2001), Obara and Zincenko (2017) and Chassang and Ortner (2019) assume a similar endogenous tie-breaking technology.

$\alpha \in [0,1]$, I solve

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

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., the leader-optimal equilibrium. 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. Let $\mathcal{F} = ([0,1] \times \Delta^n)^{\mathbb{N}_0}$ be the set of feasible sequences of prices and market shares, with Δ^n the simplex in $\mathbb{R}^n$. Say that a sequence $(p, \mathbf{x}) \in \mathcal{F}$ is *implementable* if there exists $\sigma \in \Sigma$ 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 $(p, \mathbf{x}) \in \mathcal{F}$ is implementable if and only if*

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

In words, a sequence $(p, \mathbf{x})$ is implementable if and only if firms don't have an incentive to undercut the lowest price at any period. For each $\sigma \in \Sigma$, $h \in H$ and $i \in N$, let $\pi_i(\sigma|h)$ denote firm i's flow profits under σ at history h.

Lemma 2. *For every $\sigma \in \Sigma$, there exists $\widehat{\sigma} \in \Sigma$ 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., the leader-optimal equilibrium. This equilibrium can be motivated as arising when the leader proposes the collusive agreement, subject to the constraint that fringe firms' collusive profits are weakly larger than their Nash equilibrium profits (which are zero in the current setting). As I discuss in Section 4.3, several detected cartels started when the largest firms in the industry approached their competitors and offered to initiate collusion. Moreover, the dynamics of those cartels are broadly in line with the dynamics that arise under the leader-optimal equilibrium.

As an initial step towards characterizing the leader-optimal equilibrium, I characterize the leader-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.

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

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 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 $1-(n-1)(1-\delta)$. 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 $p \leq \sum_{s=t}^{\infty} \delta^{s-t} x_{i,s} p$, 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 $i \neq 1$, $1 = \sum_{s=t}^{\infty} \delta^{s-t} x_{i,s}$. Then, for all t and all $i \neq 1$,

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

where the second equality uses $\sum_{s=t+1}^{\infty} \delta^{s-(t+1)} x_{i,s} = 1$. Hence, $x_{i,t} = 1 - \delta$ for all $i \neq 1, t$.

Note that firm 1's IC constraints are satisfied with slack under the equilibrium in Proposition 1. Let $x_{1,t}^{\mathsf{st}} = 1 - (n-1)(1-\delta)$ and $p_t^{\mathsf{st}} = 1$ denote, respectively, firm 1's market share and the price under the equilibrium in Proposition 1. Then, for all t,

$$p_t^{\mathsf{st}} = 1 < \frac{1}{1-\delta_1}(1 - (n-1)(1-\delta)) = \sum_{s=t}^{\infty} \delta_1^{s-t} x_{1,s}^{\mathsf{st}} p_s^{\mathsf{st}},$$

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 $n > 2$, and that $t^* = n - 2$ when $\delta_1 = \delta$.[7]

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 $1 - (n-1)(1-\delta)$.

The equilibrium in Proposition 2 features two phases. In the second phase, which starts at t^*, the cartel arrangement is the same as the optimal stationary arrangement in Proposition

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

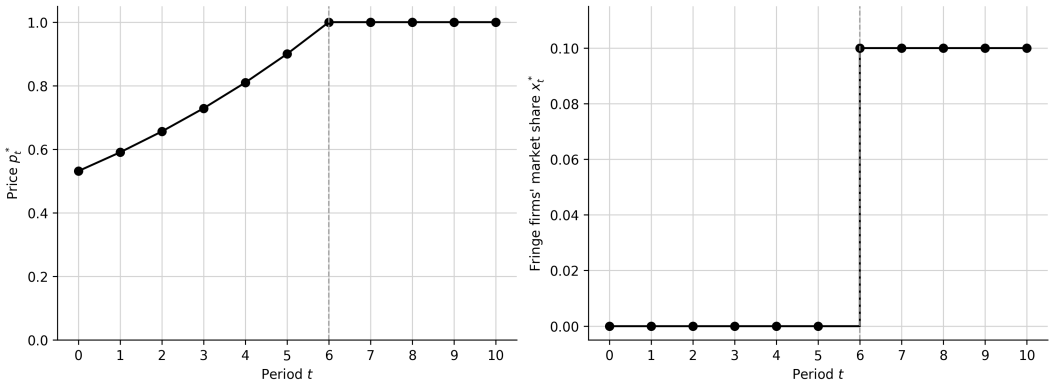

Figure 1: Prices and fringe firms' market shares under a leader-optimal equilibrium. Parameter values: $\delta_1 = 0.92, \delta = 0.9, n = 7$.

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}. Figure 1 illustrates the equilibrium in Proposition 2.

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 see why the leader-optimal equilibrium is non-stationary when $n > 2$, consider a strategy profile such that: (i) at $t = 0$, all firms choose price $p_0 = \delta$ and firm 1 captures the entire market; (ii) from $t = 1$ onwards, firms play according to the stationary equilibrium in Proposition 1. Note that no firm has incentives to deviate at $t = 0$: firm 1 doesn't gain by deviating since it captures the entire market, and firms $i \neq 1$ don't gain by deviating since their continuation payoff at $t = 0$ is $\sum_{t=1}^{\infty} \delta^t(1-\delta) = \delta = p_0$. Hence, this strategy profile is

an SPNE. Firm 1's profits under this equilibrium are $\delta + \delta_1 \overline{V}_1^{\mathsf{st}}$. Note that

$$\delta + \delta_1 \overline{V}_1^{\mathsf{st}} > \overline{V}_1^{\mathsf{st}} = 1 - (n-1)(1-\delta) + \delta_1 \overline{V}_1^{\mathsf{st}} \iff n > 2.$$

Hence, firm 1's optimal equilibrium must feature non-stationary prices when there are three or more firms in the market (i.e., when $t^* > 0$).

More generally, the benefit for firm 1 of marginally reducing the market share of fringe firms at period t is $\delta_1^t (n-1) p_t$. On the other hand, this reduction in market share tightens the IC constraint of fringe firms at all periods $s \leq t$, leading to lower prices. It can be shown (see the proof sketch in Section 4.4) that this cost is equal to $p_t \sum_{\tau=0}^{t} \delta^{t-\tau} \delta_1^\tau$. This explains why firm 1 serves the entire market at $t < t^*$ under the equilibrium in Proposition 2.

Relation to Lehrer and Pauzner (1999). Lehrer and Pauzner (1999) study repeated games with perfect monitoring and differential discounting. One of their key observations is that the Pareto frontier of the set of feasible and individually rational repeated-game payoffs is attained by streams of payoffs that satisfy the following monotonicity property: the flow payoff of the most patient player is increasing over time, and the flow payoff of the least patient player is decreasing over time. By standard folk theorem arguments, Pareto optimal SPNE exhibit the same monotonicity property as players become arbitrarily patient.

Consider next the Bertrand oligopoly game in this paper, and note that the set of feasible payoffs of the stage game is $\mathcal{V} = \{v \in \mathbb{R}_+^n : \sum_i v_i \leq 1\}$. Hence, for any $\alpha \in [1/2, 1]$, the stream $(v_t) \in \mathcal{V}^{\mathbb{N}_0}$ that maximizes $\alpha \sum_{t=0}^{\infty} \delta_1^t v_{1,t} + (1-\alpha) \sum_{i \neq 1} \sum_{t=0}^{\infty} \delta^t v_{i,t}$ has $v_{1,t} = 1$ and $v_{i,t} = 0$ for all t, $i \neq 1$. The stream of payoffs under the equilibrium in Proposition 2 converges to this optimal stream as $\delta_1, \delta \to 1$, consistent with Lehrer and Pauzner (1999).

4.3 Empirical Implications

Under the leader-optimal equilibrium in Proposition 2, prices start at a relatively low level and increase 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 gradual price increases 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 profits. Importantly, in all these cases, the cartel started operating after the largest firms in the industry approached their competitors to propose collusive agreements (Connor, 2007). Proposition 3 below shows that similar dynamics arise under Pareto optimal 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.

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. Hence, the declining market share of the largest firm, accompanied by rising prices, is also a potential marker of collusion.

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.

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 (the so-called "Gary dinners," named after Judge Elbert H. Gary, chairman of the board of US Steel). These meetings "were instituted first in stress of panic, but, their potency being demonstrated, they were afterwards called to control prices in periods of industrial calm."[8] 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 induce 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

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

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

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

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

[8] *United States v. United States Steel Corporation*, 251 U.S. 417 (1920).

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. Naturally, some constraints in (P) don't bind. In particular, one can expect (IC-leader), and the constraints $x_t \leq 1/(n-1)$ and $p_t \geq 0$, not to bind. Let $M > 0$ be large, and consider the following relaxed program:

$$\widehat{V}_1 \equiv \sup_{(x,p)\in([-M,M]^2)^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \delta_1^t (1-(n-1)x_t)p_t \tag{P'}$$

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

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

Relative to program (P), program (P') ignores firm 1's IC constraints and the constraints that $x_t \leq 1/(n-1)$ 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:

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

where $(\lambda,\mu,\gamma) = (\lambda_t,\mu_t,\gamma_t)_{t=0}^{\infty}$ are non-negative multipliers associated with the constraints in (P'). Let $\ell_+^1 = \{(a_t)_{t\geq 0} : a_t \geq 0 \text{ for all } t \text{ and } \sum_{t=0}^{\infty} a_t < \infty\}$. Then, by standard results in Lagrangian duality (e.g., Boyd and Vandenberghe, 2004),

$$\inf_{(\lambda,\mu,\gamma)\in(\ell_+^1)^3} \sup_{(x,p)\in([-M,M]^2)^{\mathbb{N}_0}} \mathcal{L}(x,p,\lambda,\mu,\gamma) \geq \sup_{(x,p)\in([-M,M]^2)^{\mathbb{N}_0}} \inf_{(\lambda,\mu,\gamma)\in(\ell_+^1)^3} \mathcal{L}(x,p,\lambda,\mu,\gamma) = \widehat{V}_1. \tag{3}$$

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

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

$$\begin{aligned} &\forall(\lambda,\mu,\gamma) \in (\ell_+^1)^3, \forall(x,p) \in ([-M,M]^2)^{\mathbb{N}_0}, \\ &\quad \mathcal{L}(x^*,p^*,\lambda,\mu,\gamma) \geq \mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*) \geq \mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*), \end{aligned} \tag{4}$$

then $(\lambda^*,\mu^*,\gamma^*)$ solves $\inf_{(\lambda,\mu,\gamma)} \sup_{(p,x)} \mathcal{L}(x,p,\lambda,\mu,\gamma)$, and (x^*,p^*) solves program (P').

To establish Proposition 2, I find multipliers $(\lambda^*,\mu^*,\gamma^*) \in (\ell_+^1)^3$ 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

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

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 $\delta_1^t(n-1)p_t$. Second, since it increases fringe firms' profits in period 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 $p_t \mu_t$.

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 $\mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*)$, multipliers $(\lambda^*,\mu^*,\gamma^*)$ are set

so that the terms in brackets in (5) are all zero. Hence, multipliers $(\lambda^*, \mu^*, \gamma^*)$ solve

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

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

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 $\tau \leq t$, 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, 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 study equilibria that maximize a weighted sum of the leader's profits and fringe firms' profits. By Lemma 2, without loss I can restrict attention to equilibria under which fringe firms obtain the same flow profits 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

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

Lemma A.1 in the Appendix shows that $t^*(\alpha) < \infty$ 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 with gradually increasing prices and with firm 1 capturing the entire market, and a second stationary phase with constant monopoly prices and with fringe firms capturing 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.

I now characterize the solution to (P_α) for $\alpha \in [0, 1/2)$. 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 $t_2^* < \infty$, and hence $t^*(\alpha) < \infty$.[9] Note further that $t^*(\alpha)$ is decreasing in α, and, when $\delta_1 > \delta$, $t^*(\alpha) = 0$ for $\alpha < 1/2$ close 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

[9] By Assumption 1, $\frac{1}{1-\delta_1}(1-(n-1)(1-\delta)) > 1$, and so t_2^* is well defined.

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 equals zero for α close to $1/2$ (provided $\delta_1 > \delta$).

To understand why equilibria that maximize the weighted sum of the leader's and the 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 $t^*(\alpha) = t_1^*(\alpha) < t_2^*$, 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 2 plots the Pareto frontier for parameter values $\delta_1 = 0.92$, $\delta = 0.9$ and $n = 7$, with fringe firms' profits on the horizontal 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 $1/2$. Note that, on the path of play under the equilibria in Propositions 3 and 4, firms' continuation payoffs move along the frontier towards the stationary payoffs marked with the black circle. The gray arrows in Figure 2 illustrate these equilibrium dynamics.

Relation to Obara and Zincenko (2017). Obara and Zincenko (2017) also study a repeated Bertrand oligopoly game with heterogeneous discount factors, and focus on equilibria with the property that prices are always equal to the monopoly price. They show that

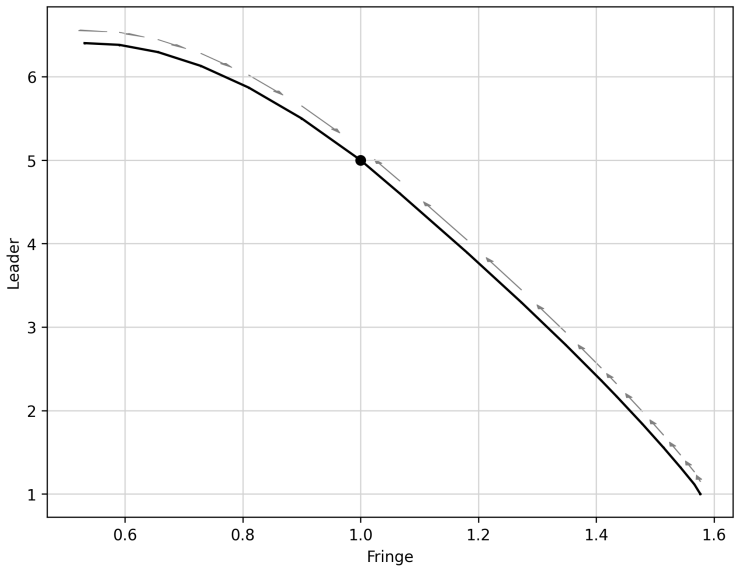

Figure 2: Pareto frontier, with fringe firms' payoffs on the horizontal axis and the leader's payoffs on the vertical axis. Parameter values: $\delta_1 = 0.92, \delta = 0.9, n = 7$. The black circle corresponds to the payoffs under the stationary equilibrium in Proposition 1.

optimal equilibria within this class have the same properties as the equilibria in Proposition 4: less patient firms enter the market weakly before 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 optimal for $\alpha < 1/2$, but not for $\alpha > 1/2$.

Obara and Zincenko (2017) 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 (though they don't characterize Pareto-optimal equilibria with this property). Second, they show that under any Pareto optimal equilibrium, play eventually converges to a stationary profile. Propositions 3 and 4 strengthen these results by characterizing Pareto optimal equilibria and Pareto optimal payoffs.

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 $\{r_1, \ldots, r_K\}$, with $0 < r_1 < r_2 < \ldots < r_K$. For $k = 1, \ldots, K$, 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 $h_t = (r_s, \mathbf{p}_s, \rho_s)_{s<t}$. A pure strategy for player i maps each pair (h_t, r_t) into an action $(p_{i,t}, \rho_{i,t}) \in \mathbb{R}_+ \times [0, 1]$.

For conciseness, I focus on leader-optimal equilibria. Let Σ denote the set of pure strategy SPNE of this game. The leader-optimal equilibrium payoff is $\overline{V}_1 \equiv \sup_{\sigma \in \Sigma} V_1(\sigma)$, 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.

Note that Lemma 2 extends to this environment. Hence, it is without loss to focus on equilibria under which all fringe firms earn the same flow profits at all periods. The problem

of finding the leader-optimal equilibrium can then be written as

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

where (x, p) is now 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, along the path of play 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

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

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 net deviation gain: at state 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 $k = 1, \ldots, K$. 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.[10] Hence, these market shares and prices can be sustained in a SPNE. 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 features pro-cyclical pricing and

[10]The proof of Proposition 5 in the Online Appendix shows this.

pro-cyclical market shares for fringe firms. Indeed, market share $x(r)$ is increasing in r: fringe firms serve a larger fraction of the market during booms, to deter them from defecting. The leader's market share, $1-(n-1)x(r)$, is decreasing in r.

Let V_f denote a 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, $\overline{r} = \sum_k q_k r_k$ and $\sum_k q_k = 1$.

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

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

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 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 the leader-optimal equilibrium in Proposition 2 extend to this setting with stochastic demand. In particular, the leader-optimal equilibrium has two phases: an initial phase with slowly increasing prices and with firm 1 capturing the entire market, and a second phase with prices at the monopoly price and with stationary state-dependent 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. Interestingly, this pattern aligns with the behavior of the US Steel Corporation in the early 1900s, whose market share fluctuated counter-cyclically (McCraw and Reinhardt, 1989).

7 Conclusion

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

I end the paper by discussing additional results and extensions that I cover in the Online Appendix.

Uniqueness. While Propositions 2-4 characterize Pareto-optimal equilibria for every Pareto weight $\alpha \in [0,1]$, in general there might be multiple equilibria attaining $\overline{V}(\alpha)$. There are at least two potential sources of multiplicity. First, when all firms share the same discount factor (i.e., $\delta_1 = \delta$), intertemporal changes in allocations are payoff neutral. Second, reallocations of market share among fringe firms can also generate multiplicity. However, Appendix B.1 shows that, when $\delta_1 > \delta$, the prices and allocations in Proposition 3 that attain $\overline{V}(\alpha)$ for $\alpha \in (1/2, 1]$ are (generically) unique.

Cost asymmetries. The model above allows firms to have different time preferences, but assumes that they all share the same marginal cost. Online Appendix B.2 allows for asymmetric costs. In particular, I consider a version of the model in which firm 1 has a cost advantage relative to fringe firms.[11] I show that, in this setting, firm 1's optimal equilibrium shares the same features as 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 of the initial phase can be positive even with two firms (i.e., even if $n = 2$).

[11] Mailath and Zemsky (1991) study optimal static collusion among firms with heterogeneous costs.

Capacity constraints. 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. One reason for this is that, in this model, firms are assumed to have enough capacity to serve the entire market. Online Appendix B.3 studies a version of the game in which firms face capacity constraints. I show that the leader-optimal equilibrium (within the class of equilibria in which firms charge a common price) also has an initial phase with gradually increasing prices, and a second stationary phase. Moreover, the leader sells at full capacity during the initial phase, and fringe firms serve the remaining fraction of the market.

This extension also allows us to better study the transition from competition to collusion. A focal Nash equilibrium of the stage game with capacity constraints is for firm 1 to sell at full capacity and for fringe firms to sell the rest (all at a price equal to marginal cost). Hence, under the leader-optimal equilibrium, firms initially freeze their market shares at their pre-collusion level, and fringe firms gain market share over time at the expense of the market leader.

Differentiated products. In the baseline model and the two extensions mentioned above, the evolution of market shares under the leader-optimal equilibrium is stark: the market shares of all firms change discretely at some date t^*. In Online Appendix B.4, I study a model with two firms selling differentiated products. For simplicity, I focus on collusive equilibria supported by the threat of Nash reversion. I show that prices and quantities can vary smoothly over time under Pareto-optimal equilibria that put sufficiently greater weight on one firm than on the other.

A Appendix

This appendix contains the proofs of Propositions 2, 3 and 4. Other proofs and extensions are in the Online Appendix.

Proof of Proposition 2. The Proposition follows from Proposition 3. ■

Before turning to the proof of Proposition 3, I establish that $t^*(\alpha) < \infty$ for all $\alpha \in [1/2, 1]$.

Lemma A.1. *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, $(n-1)\delta > n - 1 - \delta_1 > \delta_1(n-2)$, 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. In addition, 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):

$$\widehat{V}(\alpha) \equiv \sup_{(x,p) \in ([-M,M]^2)^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \alpha \delta_1^t (1 - (n-1)x_t) p_t + \sum_{t=0}^{\infty} (1-\alpha)(n-1)\delta^t x_t p_t \qquad (\mathrm{P}'_\alpha)$$

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

$$\forall t, \quad p_t \leq 1, p_t x_t \geq 0. \qquad (\text{Feasibility'})$$

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

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

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

$$\begin{aligned} &\forall(\lambda,\mu,\gamma) \in (\ell_+^1)^3, \forall (x,p) \in ([-M,M]^2)^{\mathbb{N}_0}, \\ &\quad \mathcal{L}(x^*,p^*,\lambda,\mu,\gamma) \geq \mathcal{L}(x^*,p^*,\lambda^*,\mu^*,\gamma^*) \geq \mathcal{L}(x,p,\lambda^*,\mu^*,\gamma^*), \end{aligned} \tag{9}$$

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 satisfies the remaining constraints in program (P_α), so it is also a solution to that program.

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

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

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

Lemma A.2. *(i)* $(\lambda^*,\mu^*,\gamma^*) \in (\ell_+^1)^3$*;*

(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 $t^*(\alpha)$. Note first that $\lambda^*, \mu^*, \gamma^*$ are all summable. Indeed, μ^* is summable since it takes non-zero values at finitely many periods; λ^* is summable since λ_t^* decays geometrically for all $t > t^*$; and γ^* is summable since $\gamma_t^* = \alpha\delta_1^t - \lambda_t^*$, and since λ_t^* is summable.

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 $t = t^*$,

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

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 $s = 0, \ldots, t$, 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 (9). I use the following Lemma.

Lemma A.3. *(x^*, p^*) satisfies the constraints in program (P'_α). Moreover, for all t,*

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

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

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

For all $t \geq t^*$,

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

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 $t \geq t^*$. □

Note that the complementary slackness conditions in Lemma A.3, and the fact that (x^*, p^*) satisfies all the constraints in (P'_α), imply that for all $(\lambda, \mu, \gamma) \in (\ell_+^1)^3$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (9) holds.

I now show that the second set of inequalities in (9) also holds. Note that $\sum_{t=0}^{\infty} \lambda_t \sum_{s=t}^{\infty} \delta^{s-t} x_s p_s = \sum_{t=0}^{\infty} x_t p_t \sum_{\tau=0}^{t} \delta^{t-\tau} \lambda_\tau$. Hence, 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.2(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 (9) also holds, 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

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

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

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

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):

$$\widehat{V}(\alpha) \equiv \sup_{(x,p)\in([-M,M]^2)^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \alpha\delta_1^t(1-(n-1)x_t)p_t + \sum_{t=0}^{\infty}(1-\alpha)(n-1)\delta^t x_t p_t \qquad (\mathrm{P}''_\alpha)$$

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

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

$$\forall t, \quad p_t x_t \leq p_t/(n-1), p_t \leq 1. \qquad (\text{Feasibility}'')$$

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

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

To prove the result, I construct multipliers $(\lambda^*, \mu^*, \gamma^*) \in (\ell^1_+)^4$ that, together with the allocations and prices (x^*, p^*) in the statement of the Proposition, satisfy the saddle-point condition (9). 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:

$$\begin{aligned}\mathcal{L}(x, p, \lambda, \mu, \gamma) = \sum_{t=0}^{\infty} p_t x_t \left[(n-1) \left(-\alpha \delta_1^t + (1-\alpha)\delta^t - \mu_t - \sum_{\tau=0}^{t} \delta_1^{t-\tau} \lambda_{1,\tau} \right) + \sum_{\tau=0}^{t} \delta^{t-\tau} \lambda_{f,\tau} \right] \\ + \sum_{t=0}^{\infty} p_t \left(\alpha \delta_1^t - \lambda_{f,t} - \gamma_t + \sum_{\tau=0}^{t-1} \delta_1^{t-\tau} \lambda_{1,\tau} + \mu_t \right) + \sum_{t=0}^{\infty} \gamma_t. \qquad (10)\end{aligned}$$

Recall the definitions of $t_1^*(\alpha)$ and t_2^* in the main text, and 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 $(\lambda^*, \mu^*, \gamma^*)$:

$$\begin{aligned}\lambda_{1,t}^* &= \mathbf{1}_{t=0} \left((1-\alpha) \left(\frac{\delta}{\delta_1} \right)^{t^*} - \alpha \right), \\ \mu_t^* &= \mathbf{1}_{t<t^*} (1-\alpha) \delta_1^t \left(\left(\frac{\delta}{\delta_1} \right)^{t} - \left(\frac{\delta}{\delta_1} \right)^{t^*} \right), \\ \gamma_t^* &= \begin{cases} \alpha + (1-\alpha) \left(1 - \left(\frac{\delta}{\delta_1} \right)^{t^*} \right) & \text{if } t = 0, \\ (1-\alpha)\delta^t & \text{if } t \in (0, t^*] \text{ and } t^* > 0, \\ (1-\alpha)\delta^{t^*} \delta_1^{t-t^*-1} ((n-1)\delta - (n-2)\delta_1) & \text{if } t > t^*, \end{cases} \\ \lambda_{f,t}^* &= \begin{cases} 0 & \text{if } t \leq t^*, \\ (n-1)(1-\alpha)\delta^{t^*} \delta_1^{t-t^*-1} (\delta_1 - \delta) & \text{if } t \geq t^* + 1. \end{cases}\end{aligned}$$

I use the following Lemma:

Lemma A.4. *(i)* $(\lambda^*, \mu^*, \gamma^*) \in (\ell^1_+)^4$*;*

(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). Note first that $\lambda^*, \mu^*, \gamma^*$ are all summable. Indeed, λ^*_1 and μ^* are summable since they take non-zero values at only finitely many periods; and λ^*_f and γ^* are summable since they decay geometrically for all $t > t^*$.

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, $\gamma^*_t \geq 0$ 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 for all $s = t^* + 1, \ldots, t-1$. 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

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

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 next $t \in (0, t^*)$, 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.5. *(x^*, p^*) satisfies the constraints in program (P''_α). Moreover, for all t,*

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

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 $t \le t^*$,

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

with equality at $t=0$ and strict inequality for all $t \in (0, t^*]$. For $t > t^*$,

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

where the strict inequality follows from Assumption 1.

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

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

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 $1-\delta < 1/(n-1)$. For $t > t^*$, we have that

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

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 hold 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. □

The complementary slackness conditions in Lemma A.5, and the fact that (x^*, p^*) satisfies all the constraints in (P''_α), imply that for all $(\lambda, \mu, \gamma) \in (\ell_+^1)^4$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (9) holds.

To see that the second set of inequalities in (9) also holds, note that equation (10) and Lemma A.4(ii)-(iii) imply 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, (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,

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

I use the following Lemma:

Lemma A.6. *(i)* $(\lambda^*, \mu^*, \gamma^*) \in (\ell_+^1)^4$*;*

(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). Note that the multipliers are summable: λ_1^* and μ^* are summable because they take non-zero values at only finitely many periods, and γ^*, λ_f^* are summable since they decay geometrically for all $t > t^*$. Clearly, $\lambda_{1,t}^* \geq 0$ for all t. That $\mu_t^* \geq 0$ for all t follows since $-\alpha\delta_1^t + (1-\alpha)\delta^t > 0$ for all $t \leq t_1^*$, and since I'm considering the case with $t_2^* > t_1^* = t^*$.

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

Next, I show that $\gamma_t^* \geq 0$ for all t. The inequality clearly holds for all $t \leq t^*$. Consider next $t = t^*+1 = t_1^*+1$ (since I'm in the case with $t^* = t_1^* < t_2^*$). Let $\beta \equiv \frac{\delta_1}{\delta} \geq 1$, and note that

$$(1-\alpha)\delta^{t^*} - \alpha\delta_1^{t^*} \geq 0 \iff \alpha\beta^{t^*} \leq 1-\alpha, \tag{11}$$

where the first inequality follows since $(1-\alpha)\delta^t - \alpha\delta_1^t \geq 0$ for all $t \leq t_1^*$ 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 (11)), 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 $t > t^* + 1$,

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

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 $\tau \leq t^*$ 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 $s = t^* + 1, \ldots, t-1$. 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 now 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 $t = t^* + 1$,

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

Finally, for $t > t^* + 1$,

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

□

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

$$\begin{aligned}
\lambda_{f,t}^* \left(\sum_{s=t}^{\infty} \delta^{s-t} x_s^* p_s^* - p_t^* \right) &= 0, \\
\lambda_{1,t}^* \left(\sum_{s=t}^{\infty} \delta_1^{s-t} (1-(n-1)x_s^*) p_s^* - p_t^* \right) &= 0, \\
\mu_t^* p_t^* (1-(n-1)x_t^*) &= 0, \\
\gamma_t^* (1 - p_t^*) &= 0.
\end{aligned}$$

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

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

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 $t \geq t^* + 1$,

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

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 $t \leq t^*$,

$$\sum_{s=t}^{\infty} \delta^{s-t} x_s^* p_s^* > \sum_{s=t}^{\infty} \delta^{s-t}(1-\delta) = 1 = p_t^*,$$

where the strict inequality 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 $1-\delta < 1/(n-1)$. For $t > t^*$,

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

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 $t \leq t^*$, and since fringe firms' IC constraints are satisfied with equality for all $t > t^*$. The second set of equalities also holds, since $\lambda_{1,t}^* = 0$ for all t. The third set of equalities holds since $x_t^* = 1/(n-1)$ for all $t \leq t^*$, and since $\mu_t^* = 0$ for all $t > t^*$. The last set of equalities holds since $p_t^* = 1$ for all t. □

The complementary slackness conditions in Lemma A.7, and the fact that (x^*, p^*) satisfies all the constraints in (P''_α), imply that for all $(\lambda, \mu, \gamma) \in (\ell_+^1)^4$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (9) holds.

To see that the second set of inequalities in (9) also holds, note that equation (10) and Lemma A.6(ii)-(iii) imply 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, (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 2 (i.e., $t_2^* > t_1^*$). ■

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 (2026): "The Value of Privacy in Cartels: An Analysis of the Inner Workings of a Bidding Ring," *Review of Economic Studies*, 93, 600–628.

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.

McCraw, T. K. and F. Reinhardt (1989): "Losing to win: US Steel's pricing, investment decisions, and market share, 1901–1938," *The Journal of Economic History*, 49, 593–619.

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.

Pavoni, N., C. Sleet, and M. Messner (2018): "The dual approach to recursive optimization: theory and examples," *Econometrica*, 86, 133–172.

Rahman, D. M. (2025): "Duality and Recursion in Dynamic Games," .

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.

——— (2026): "Collusion with Optimal Information Disclosure," *The Quarterly Journal of Economics*, 141, 2555–2595.

B Supplementary Appendix – for online publication

B.1 Uniqueness

As discussed in the Conclusion, there are at least two sources of multiplicity: (i) intertemporal changes in allocations under equal discounting; and (ii) reallocations of market share among fringe firms. To see how (i) can give rise to multiplicity, note that for $\alpha = 1/2$, a "bid-rotation" scheme in which firms take turns capturing the entire market at the monopoly price (and in which deviations are punished with Nash reversion) is a Pareto-optimal equilibrium whenever δ is high enough. To see how (ii) can give rise to multiplicity, note that the equilibria in Proposition 4 feature non-binding incentive constraints for fringe firms during the initial phase. Hence, for $\alpha < 1/2$, one can construct alternative equilibria attaining $\overline{V}(\alpha)$ by reallocating market share among fringe firms at periods $t < t^*(\alpha)$.

The next result shows that, for $\alpha > 1/2$ and $\delta_1 > \delta$, under a genericity condition all equilibria that attain $\overline{V}(\alpha)$ have on-path prices and market shares equal to those in Proposition 3. Recall that $t^*(\alpha) = \inf\left\{t \in \mathbb{N}_0 : \alpha \sum_{\tau=0}^{t} \delta_1^\tau \delta^{t-\tau} \geq \alpha\delta_1^t(n-1) - (1-\alpha)\delta^t(n-1)\right\}$ for $\alpha > 1/2$. Note that, generically,

$$\alpha \sum_{\tau=0}^{t^*(\alpha)} \delta_1^\tau \delta^{t^*(\alpha)-\tau} > \alpha\delta_1^{t^*(\alpha)}(n-1) - (1-\alpha)\delta^{t^*(\alpha)}(n-1). \tag{12}$$

As in the main paper, I maintain Assumption 1.

Proposition B.1. *Fix $\delta_1 > \delta$ and $\alpha \in (1/2, 1]$ such that (12) holds. Then, any equilibrium that attains $\overline{V}(\alpha)$ has the same on-path prices and market shares as those in Proposition 3.*

Proof of Proposition B.1. Fix $\delta_1 > \delta$ and $\alpha \in (1/2, 1]$ satisfying (12). For all bounded sequences (x_t, p_t), with $x_t = (x_{i,t})_{i\neq 1}$, define

$$J(x,p) \equiv \sum_{t=0}^{\infty} \left(\alpha\delta_1^t \left(1 - \sum_{i\neq 1} x_{i,t}\right) p_t + (1-\alpha)\delta^t \sum_{i\neq 1} x_{i,t} p_t \right).$$

The problem of finding a Pareto optimal equilibrium with weight α is

$$\overline{V}(\alpha) = \sup_{(x,p)} J(x,p) \qquad (\mathrm{P}_{gen})$$
$$\text{s.t. } \forall t, \forall i \neq 1, \quad p_t \leq \sum_{s=t}^{\infty} \delta^{s-t} x_{i,s} p_s, \quad p_t \leq \sum_{s=t}^{\infty} \delta_1^{s-t} \left(1 - \sum_{i \neq 1} x_{i,s}\right) p_s, \quad p_t \in [0,1], \quad x_t \in \Delta_{\leq}^{n-1},$$

where $\Delta_{\leq}^{n-1} = \left\{a \in \mathbb{R}_+^{n-1} : \sum_i a_i \leq 1\right\}$. Fix $M > 0$ large and consider the following relaxed program:

$$\widehat{V}(\alpha) = \sup_{(x,p) \in ([(-M,M)^n])^{\mathbb{N}_0}} J(x,p) \qquad (\mathrm{P}_{rel})$$
$$\text{s.t. } \forall t, \forall i \neq 1, \quad p_t \leq \sum_{s=t}^{\infty} \delta^{s-t} x_{i,s} p_s, \quad p_t \leq 1, \quad p_t x_{i,t} \geq 0.$$

The Lagrangian associated with (P_{rel}) is

$$\begin{aligned}\mathcal{L}(x,p,(\lambda_i)_{i\neq 1},(\mu_i)_{i\neq 1},\gamma) = J(x,p) &+ \sum_{t=0}^{\infty}\sum_{i\neq 1} \lambda_{i,t}\left(\sum_{s=t}^{\infty} \delta^{s-t} x_{i,s} p_s - p_t\right) \\ &+ \sum_{t=0}^{\infty}\sum_{i\neq 1} \mu_{i,t} p_t x_{i,t} + \sum_{t=0}^{\infty} \gamma_t (1 - p_t),\end{aligned}$$

where $(\lambda_i), (\mu_i) \in (\ell_+^1)^{n-1}$ and $\gamma \in \ell_+^1$. Note that the Lagrangian can be re-written as:

$$\begin{aligned}\mathcal{L}(x,p,(\lambda),(\mu),\gamma) = \sum_{t=0}^{\infty}\sum_{i\neq 1} p_t x_{i,t}&\left(-\alpha\delta_1^t + (1-\alpha)\delta^t + \sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_{i,\tau} + \mu_{i,t}\right) \\ &+ \sum_{t=0}^{\infty} p_t \left(\alpha\delta_1^t - \sum_{i\neq 1}\lambda_{i,t} - \gamma_t\right) + \sum_{t=0}^{\infty}\gamma_t.\end{aligned} \qquad (13)$$

Recall the multipliers $\lambda^*, \mu^*, \gamma^*$ in the proof of Proposition 3. For each $t, i \neq 1$, let $\lambda_{i,t}^* = \frac{1}{n-1}\lambda_t^*$ and $\mu_{i,t}^* = \frac{1}{n-1}\mu_t^*$. Lemma A.2 implies that all these multipliers are non-negative and summable. Lemma A.2 also implies that, for all $i \neq 1$ and all t, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_{i,\tau}^* = \alpha\delta_1^t - (1-\alpha)\delta^t - \mu_{i,t}^*$. Since $\gamma_t^* = \alpha\delta_1^t - \lambda_t^*$, it follows that, for all (x,p), $\mathcal{L}(x,p,(\lambda^*),(\mu^*),\gamma^*) = \sum_{t=0}^{\infty}\gamma_t^*$.

Let (x^*, p^*) be as in the statement of Proposition 3: i.e., $p_t^* = \min\{\delta^{t^*-t}, 1\}$ for all t; $x_{i,t}^* = (1-\delta)\mathbf{1}_{t \geq t^*}$ for all $t, i \neq 1$. By the same arguments as in Lemma A.3, (x^*, p^*) satisfies the complementary slackness conditions. Together with (13), it follows that $(x^*, p^*, (\lambda^*), (\mu^*), \gamma^*)$ is a saddle point of $\mathcal{L}(x, p, (\lambda), (\mu), \gamma)$, and so (x^*, p^*) solves (P_{rel}). Complementary slackness and (13) give us $J(x^*, p^*) = \mathcal{L}(x^*, p^*, (\lambda^*), (\mu^*), \gamma^*) = \sum_{t=0}^{\infty} \gamma_t^*$. Moreover, we know from Proposition 3 that (x^*, p^*) also solves (P_{gen}).

Let (x', p') be any solution to (P_{gen}), so that $J(x', p') = J(x^*, p^*)$. Then,

$$J(x', p') \leq \mathcal{L}(x', p', (\lambda^*), (\mu^*), \gamma^*) = \sum_{t=0}^{\infty} \gamma_t^* = J(x^*, p^*) = J(x', p'),$$

where the inequality follows since multipliers $(\lambda^*), (\mu^*), \gamma^*$ are all summable and non-negative and since (x', p') satisfies all the constraints in (P_{gen}), the first equality follows since $\sum_{t=0}^{\infty} \gamma_t^* = \mathcal{L}(x, p, (\lambda^*), (\mu^*), \gamma^*)$ for all bounded (x, p), the second equality follows from the arguments above, and the last equality follows since (x', p') and (x^*, p^*) are solutions to (P_{gen}). Hence, $J(x', p') = \mathcal{L}(x', p', (\lambda^*), (\mu^*), \gamma^*)$. Since all multipliers are summable and non-negative, this equality implies that $(x', p', (\lambda^*), (\mu^*), \gamma^*)$ must satisfy complementary slackness:

$$\forall i \neq 1, \forall t, \qquad \lambda_{i,t}^* \left(\sum_{s=t}^{\infty} \delta^{s-t} x_{i,s}' p_s' - p_t' \right) = 0, \qquad \mu_{i,t}^* p_t' x_{i,t}' = 0, \qquad \gamma_t^* (1 - p_t') = 0.$$

By the proof of Proposition 3, and using (12), we have

$$\lambda_{i,t}^* > 0 \text{ for all } i \neq 1, t, \qquad \mu_{i,t}^* > 0 \text{ for all } i \neq 1, t < t^*, \qquad \gamma_t^* > 0 \text{ for all } t \geq t^*.$$

$\mu_{i,t}^* > 0$ for all $t < t^*, i \neq 1$ implies $p_t' x_{i,t}' = 0$ for all $t < t^*, i \neq 1$. $\gamma_t^* > 0$ for all $t \geq t^*$ implies $p_t' = 1$ for all $t \geq t^*$. Since $\lambda_{i,t}^* > 0$ for all $t, i \neq 1$, the IC constraints of fringe firms must hold with equality for all t. Together with $p_t' = 1$ for all $t \geq t^*$, the same arguments as in Section 4.1 imply $x_{i,t}' = 1 - \delta$ for all $t \geq t^*, i \neq 1$. Lastly, binding IC constraints for all fringe firms at $t < t^*$ implies that, for all $t < t^*, i \neq 1$, $p_t' = \sum_{s=t}^{\infty} \delta^{s-t} p_s' x_{i,s}' = \delta^{t^*-t}$, where the second equality uses $p_s' x_{i,s}' = 1 - \delta$ for all $i \neq 1, s \geq t^*$ and $p_s' x_{i,s}' = 0$ for all $i \neq 1, s < t^*$.

Together with $p'_t x'_{i,t} = 0$ for all $t < t^*, i \neq 1$, this implies that $x'_{i,t} = 0$ for all $i \neq 1, t < t^*$. Hence, $(x', p') = (x^*, p^*)$. This completes the proof. ■

B.2 Asymmetric Costs

This appendix studies an extension of the baseline model in which firm 1 has a cost advantage relative to fringe firms. In particular, I consider the same model as in the main text, with the added feature that fringe firms' marginal cost is $c \in (0, 1)$, while firm 1's cost is zero. I assume that all firms have the same discount factor: $\delta_i = \delta$ for all $i \in N$. As in the main text, I maintain Assumption 1, which with equal discounting boils down to $\delta > (n-1)/n$.

Lemma 2 continues to hold in this environment. Hence, the problem of finding the leader-optimal equilibrium can be written as:

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

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

IC-leader, Feasibility

where, for any $y \in \mathbb{R}$, $(y)^+ = \max\{y, 0\}$. Note that the IC constraints of firm 1 remain unchanged relative to the baseline model: with endogenous tie-breaking, the stage game has a Nash equilibrium in which all firms charge price $p = 0$ and firm 1 captures the entire market and earns zero profits (i.e., firm 1 chooses $\rho_1 = 1$ and all firms $i \neq 1$ choose $\rho_i = 0$).

Define

$$t^*(c) \equiv \inf \left\{ t \in \mathbb{N}_0 : t \geq \frac{n-1}{1-c} - 1 \right\}.$$

Proposition B.2. *With asymmetric costs, there exists a leader-optimal equilibrium in which, at every period t along the path of play,*

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

(ii) if $t \geq t^*(c)$, 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 $1-(n-1)(1-\delta)$.

Proof of Proposition B.2. Let $M > 0$ be large, and consider the following auxiliary program

$$\tilde{V}_1 \equiv \sup_{(x,p)\in([-M,M]^2)^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \delta^t \left(c + (p_t - c)^+ - (n-1)x_t \frac{(p_t-c)^+}{1-c} \right) \tag{P'-c}$$

$$\text{s.t. } \forall t, \quad (p_t - c)^+ \leq \sum_{s=t}^{\infty} \delta^{s-t} x_s (p_s - c)^+ \tag{IC-fringe-cost'}$$

$$\forall t, \quad \frac{(p_t-c)^+}{1-c} \leq 1, (p_t-c)^+ x_t \geq 0 \tag{Feasibility-cost}$$

Note that (IC-fringe-cost) implies (IC-fringe-cost'), and (Feasibility) implies (Feasibility-cost). Moreover, for any x, p satisfying $x \geq 0$ and $p \in [0,1]$, $c + (p-c)^+ - (n-1)x\frac{(p-c)^+}{1-c} \geq (1-(n-1)x)p$. Hence, $\tilde{V}_1 \geq \overline{V}_1$. The Lagrangian associated with (P'-c) is

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

Let (p^*, x^*) be the prices and market shares in the statement of Proposition B.2, and let $(\lambda^*, \mu^*, \gamma^*)$ be given by

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

Note that $(\lambda^*, \mu^*, \gamma^*) \in (\ell_+^1)^3$. Indeed, the multipliers are all summable, since they either decay geometrically or take non-zero values at finitely many periods. The definition of $t^*(c)$ implies that, for all t, $\lambda_t^* \geq 0$ and $\mu_t^* \geq 0$. Moreover, $\gamma_t^* = 0$ for all $t < t^*(c)$, and

$\gamma_t^* = (1-c)\delta^t > 0$ for all $t > t^*(c)$. At $t = t^*(c)$,

$$\gamma_{t^*(c)}^* = (1-c)\delta^{t^*(c)}\left(t^*(c) + 1 - \frac{n-1}{1-c}\right) \geq 0.$$

I now show that $(\lambda^*, \mu^*, \gamma^*)$ and (p^*, x^*) are a saddle point of $\mathcal{L}$, which implies that (p^*, x^*) is a solution to (P'-c). I start with two Lemmas. For the rest of the proof, I use t^* instead of $t^*(c)$.

Lemma B.1. *For all $t \geq 0$, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau^* = \delta^t \frac{n-1}{1-c} - \mu_t^*$.*

Proof. For $t < t^*$,

$$\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau^* = \sum_{\tau=0}^{t} \delta^{t-\tau}\delta^\tau = \delta^t(t+1) = \delta^t \frac{n-1}{1-c} - \mu_t^*.$$

For $t = t^*$,

$$\sum_{\tau=0}^{t^*} \delta^{t^*-\tau}\lambda_\tau^* = \delta \sum_{\tau=0}^{t^*-1} \delta^{t^*-1-\tau}\lambda_\tau^* + \lambda_{t^*}^* = \delta^{t^*}\frac{n-1}{1-c} - \delta\mu_{t^*-1} + \lambda_{t^*}^* = \delta^{t^*}\frac{n-1}{1-c}.$$

Since $\mu_{t^*} = 0$, the equality also holds at t^*. For $t > t^*$

$$\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau^* = \delta \sum_{\tau=0}^{t-1} \delta^{t-1-\tau}\lambda_\tau^* + \lambda_t^* = \delta^t\frac{n-1}{1-c} + \lambda_t^* = \delta^t\frac{n-1}{1-c}.$$

Since $\mu_t = 0$ for all $t > t^*$, the equality also holds for all such t. □

Lemma B.2. *(x^*, p^*) satisfy all the constraints in program* (P'-c)*. Moreover, for all t,*

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

Proof. Clearly, (p^*, x^*) satisfy the feasibility constraints in (P'-c). Since $x_t^* = 0$ for all $t < t^*$ and $\mu_t^* = 0$ for all $t \geq t^*$, the second set of complementary slackness conditions also holds. The third set of complementary slackness conditions holds since $\gamma_t^* = 0$ for all $t < t^*$, and since $p_t^* = 1$ for all $t \geq t^*$.

Next, I show that the IC constraint of fringe firms binds at all t. For $t \geq t^*$,

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

For $t < t^*$,

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

Hence the IC constraints in (P'-c) and the first set of complementary slackness conditions hold. This completes the proof. □

Lemma B.2 implies that $\mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*) \leq \mathcal{L}(x^*, p^*, \lambda, \mu, \gamma)$ for all $(\lambda, \mu, \gamma) \in (\ell_+^1)^3$. Next, note that $\mathcal{L}$ can be re-written as

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

By Lemma B.1, $\mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*) = \frac{c}{1-\delta} + \sum_{t=0}^{\infty} \gamma_t^*$ for all (x, p). Hence, for all (x, p), $\mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*) \geq \mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*)$. Therefore, (x^*, p^*) solves (P'-c).

To finish the proof, I show that (x^*, p^*) also solves (P-c). Note first that, under (x^*, p^*), the objective in (P'-c) coincides with the objective in (P-c). Next, I show that (x^*, p^*) satisfies all the constraints in program (P-c). Clearly, $p_t^* \in [0, 1]$ for all t. Next, Assumption 1 implies $(1 - \delta) < 1/(n-1)$, and so $x_t^* \in [0, 1/(n-1)]$ for all t. Since $p_t^* > c$ for all t, and since (x^*, p^*) satisfies the fringe firms' IC constraints in (P'-c), (x^*, p^*) also satisfies the fringe firms' IC constraints in (P-c).

Lastly, I show that (x^*, p^*) satisfies firm 1's IC constraints in (P-c). For $t < t^*$, firm 1's IC constraint holds since $x_t^* = 0$ for all such t. For $t \geq t^*$,

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

where the strict inequality follows from Assumption 1. ■

B.3 Capacity Constraints

This appendix studies an extension of the baseline model in which firms have capacity constraints. Fix $(K_i)_{i \in N} \in [0,1)^n$. Firm i can only produce up to K_i units, at zero cost. I assume that, for all $i \neq 1$, $K_i = K_f \leq K_1$, and that, for all i, $\delta_i = \delta \in (0,1)$. I further assume efficient rationing: firms charging a lower price serve consumers first, up to their capacity. As in the baseline model, firms charging the same price can endogenously determine their market shares up to their capacities. I also make the following assumption:

Assumption 3. *For all i, $\sum_{j \neq i} K_j \geq 1$, and*

$$\frac{K_1 + (n-1)K_f - 1}{K_1 + (n-1)K_f} < \delta < \frac{(n-1)K_f + K_1 - 1}{(n-1)K_f}.$$

The first part of Assumption 3 says that, for each firm, the joint capacity of all of its competitors is enough to serve the entire market. This assumption implies that a firm that sets a higher price than all of its competitors earns zero profits. The second part of Assumption 3 says that the discount factor lies in an intermediate range. The lower bound on δ plays the same role as Assumption 1. To understand the upper bound on δ, consider a strategy profile in which, along the path of play, the price is always equal to 1 and firm 1 sells at full capacity. Note that, if sustainable, this would be firm 1's optimal equilibrium. The upper bound on the discount factor implies that such a strategy profile is not sustainable: fringe firms would have a strict incentive to defect and undercut the market price.

For simplicity, I focus attention on equilibria in which firms charge the same price along the path of play. Note that, under the common-price restriction, Lemma 2 continues to hold. Hence, the problem of finding the leader-optimal equilibrium (within the class of equilibria

with common prices) is:

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

$$\text{s.t. } \forall t, \quad K_f p_t \leq \sum_{s=t}^{\infty} \delta^{s-t} x_s p_s, \tag{IC-Fringe-cc}$$

$$\forall t, \quad K_1 p_t \leq \sum_{s=t}^{\infty} \delta^{s-t} \left(1 - (n-1)x_s\right) p_s \tag{IC-1-cc}$$

$$\forall t, \quad \frac{1-K_1}{n-1} \leq x_t \leq K_f, p_t \in [0,1] \tag{Feasibility-cc}$$

Define the sequence (a_t) such that, for all t,

$$a_t = \frac{(n-1)K_f a_{t-1} + (n-1)K_1}{(n-1)K_f + K_1 - 1}, \tag{14}$$

with $a_{-1} = 0$. The solution to this sequence is:

$$\forall t, \quad a_t = \frac{(n-1)K_1}{1-K_1} \left[\left(\frac{K_f(n-1)}{K_f(n-1) + K_1 - 1} \right)^{t+1} - 1 \right].$$

Define

$$t^*_{cc} \equiv \inf\{t \in \mathbb{N}_0 : a_t \geq n-1\},$$

and note that $K_1 < 1$ implies $t^*_{cc} < \infty$.

Proposition B.3. *With capacity constraints, there exists a leader-optimal equilibrium in which, at every period t along the path of play,*

(i) if $t < t^*_{cc}$, the price is equal to $\left(\frac{\delta K_f(n-1)}{(n-1)K_f + K_1 - 1}\right)^{t^*_{cc}-t}$, firm 1 sells at full capacity, and each fringe firm earns a market share of $(1-K_1)/(n-1)$;

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

Proof of Proposition B.3. Let $M > 0$ be large, and consider the following relaxation of

(P-cc)

$$\widehat{V}_1 \equiv \sup_{(p,x)\in([-M,M]^2)^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \delta^t \left(1-(n-1)x_t\right) p_t \qquad \text{(P-cc')}$$

$$\text{s.t. } \forall t, \quad K_f p_t \leq \sum_{s=t}^{\infty} \delta^{s-t} x_s p_s,$$

$$\forall t, \quad p_t x_t \geq p_t \frac{1-K_1}{n-1}, p_t \leq 1.$$

Since this is a relaxation of (P-cc), $\widehat{V}_1 \geq \overline{V}_1$. The Lagrangian associated with (P-cc') is

$$\mathcal{L}(p,x,\lambda,\mu,\gamma) = \sum_{t=0}^{\infty} \delta^t \left(1-(n-1)x_t\right) p_t + \sum_{t=0}^{\infty} \lambda_t \left(\sum_{s=t}^{\infty} \delta^{s-t} x_s p_s - K_f p_t \right)$$
$$+ \sum_{t=0}^{\infty} \mu_t p_t (x_t - (1-K_1)/(n-1)) + \sum_{t=0}^{\infty} \gamma_t (1-p_t).$$

For the rest of the proof, I write t^* instead of t^*_{cc}. Consider the following multipliers:

$$\lambda_t^* = \begin{cases} \delta^t (a_t - a_{t-1}) & \text{if } t < t^*, \\ \delta^{t^*} \left((n-1) - a_{t^*-1}\right) & \text{if } t = t^*, \\ 0 & \text{if } t > t^*, \end{cases} \quad \mu_t^* = \mathbf{1}_{t<t^*} \delta^t (n-1-a_t), \quad \gamma_t^* = \delta^t - K_f \lambda_t^* - \frac{1-K_1}{n-1} \mu_t^*.$$

I now show that the prices and quantities (p^*, x^*) in the statement of the proposition and the multipliers $(\lambda^*, \mu^*, \gamma^*)$ are a saddle-point of $\mathcal{L}$. I use the following Lemmas.

Lemma B.3. *$(\lambda^*, \mu^*, \gamma^*) \in (\ell^1_+)^3$. Moreover, for all t,*

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

Proof. I start by showing that $\lambda^*, \mu^*, \gamma^* \in \ell^1_+$. Note first that the multipliers are all summable, since they either decay exponentially or take non-zero values at finitely many

periods. By the definition of t^*, $\mu_t^* \geq 0$ for all t. Next, note that, for all $t < t^*$,

$$\lambda_t^* = \delta^t(a_t - a_{t-1}) = \delta^t \frac{(n-1)K_1}{K_f(n-1)+K_1-1}\left(\frac{K_f(n-1)}{K_f(n-1)+K_1-1}\right)^t > 0.$$

Moreover, $\lambda_{t^*}^* > 0$ by the definition of t^*, and $\lambda_t^* = 0$ for all $t > t^*$.

Next, I show that $\gamma^* \geq 0$. For $t < t^*$,

$$\gamma_t^* = \delta^t - K_f\lambda_t^* - \frac{1-K_1}{n-1}\mu_t^* = \frac{\delta^t}{n-1}\left((n-1)K_f a_{t-1} + (n-1)K_1 - a_t((n-1)K_f + K_1 - 1)\right) = 0,$$

where the second equality follows from the definition of λ_t^*, μ_t^*, and the last equality follows from (14). For $t = t^*$,

$$\gamma_{t^*}^* = \delta^{t^*}(1 - (n-1)K_f + a_{t^*-1}K_f) \geq 0,$$

where the equality follows from the definition of λ_t^*, μ_t^*, and the inequality follows since

$$n - 1 \leq a_{t^*} = \frac{(n-1)K_f a_{t^*-1} + (n-1)K_1}{(n-1)K_f + K_1 - 1}.$$

For $t > t^*$, $\gamma_t^* = \delta^t > 0$.

Next, I show that the equality in the statement of the Lemma holds. For all $t < t^*$,

$$\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau^* = \delta^t \sum_{\tau=0}^{t}(a_\tau - a_{\tau-1}) = \delta^t a_t,$$

where the last equality uses $a_{-1} = 0$. By the definition of μ_t^*, it follows that $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau^* = \delta^t(n-1) - \mu_t^*$ for all $t < t^*$. For $t = t^*$,

$$\sum_{\tau=0}^{t^*} \delta^{t^*-\tau}\lambda_\tau^* = \delta \sum_{\tau=0}^{t^*-1} \delta^{t^*-1-\tau}\lambda_\tau^* + \lambda_{t^*}^* = \delta^{t^*} a_{t^*-1} + \lambda_{t^*}^* = \delta^{t^*}(n-1) = \delta^{t^*}(n-1) - \mu_{t^*}^*.$$

For $t > t^*$,

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

□

Lemma B.4. (x^*, p^*) *satisfy all the constraints in* (P-cc'). *Moreover, for all* t,

$$\lambda_t^* \left(\sum_{s=t}^{\infty} \delta^{s-t} x_s^* p_s^* - K_f p_t^* \right) = 0, \quad \mu_t^* \left(p_t^* x_t^* - p_t^* \frac{1-K_1}{n-1} \right) = 0, \quad \gamma_t^* (1 - p_t^*) = 0.$$

Proof. The feasibility constraints in (P-cc') hold since $p_t^* \in (0,1]$ for all t and since $x_t^* = \frac{1-K_1}{n-1}$ for $t < t^*$ and $x_t^* = (1-\delta)K_f > \frac{1-K_1}{n-1}$ for $t \geq t^*$ (the inequality follows from Assumption 3).

Next, I show that the IC constraints of fringe firms hold with equality at all t. For $t \geq t^*$,

$$\sum_{s=t}^{\infty} \delta^{s-t} x_s^* p_s^* = K_f = p_t^* K_f.$$

The proof that the IC constraint holds with equality at all $t < t^*$ follows from backward induction. Note first that, for all $t < t^*$,

$$\frac{p_t^*}{p_{t+1}^*} = \frac{\delta K_f (n-1)}{(n-1)K_f + K_1 - 1}. \tag{15}$$

Suppose that the constraint holds with equality at $t+1 \leq t^*$; i.e., $K_f p_{t+1}^* = \sum_{s=t+1}^{\infty} \delta^{s-t-1} x_s^* p_s^*$. Then,

$$\sum_{s=t}^{\infty} \delta^{s-t} x_s^* p_s^* = \frac{1-K_1}{n-1} p_t^* + \delta K_f p_{t+1}^* = K_f p_t^*,$$

where the last equality follows from (15). Hence, since the IC constraint holds with equality at t^*, it also holds with equality at all $t < t^*$.

The first set of complementary slackness conditions holds since the IC constraints of fringe firms hold with equality at all t. The second set of complementary slackness conditions holds since $x_t^* = \frac{1-K_1}{n-1}$ for all $t < t^*$, and since $\mu_t^* = 0$ for all $t \geq t^*$. The third set of complementary slackness conditions holds since $\gamma_t^* = 0$ for all $t < t^*$ (see the proof of Lemma B.3) and since

$p_t^* = 1$ for all $t \geq t^*$. □

Lemma B.4 implies that $\mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*) \leq \mathcal{L}(x^*, p^*, \lambda, \mu, \gamma)$ for all $(\lambda, \mu, \gamma) \in (\ell_+^1)^3$. Next, note that $\mathcal{L}$ can be re-written as

$$\begin{aligned}\mathcal{L}(x, p, \lambda, \mu, \gamma) = &\sum_{t=0}^{\infty} p_t x_t \left(-\delta^t (n-1) + \sum_{\tau=0}^{t} \delta^{t-\tau} \lambda_\tau + \mu_t \right) \\ &+ \sum_{t=0}^{\infty} p_t \left(\delta^t - K_f \lambda_t - \gamma_t - \mu_t \frac{1 - K_1}{n-1} \right) + \sum_{t=0}^{\infty} \gamma_t.\end{aligned}$$

Lemma B.3 implies that, for all (x, p), $\mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*) = \sum_{t=0}^{\infty} \gamma_t^*$. Hence, for all (x, p), $\mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*) \geq \mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*)$. Therefore, (x^*, p^*) solves (P-cc').

Lastly, I show that (x^*, p^*) also satisfies the constraints in (P-cc) that I ignored, and so (x^*, p^*) also solves (P-cc). Note that, for all t, $p_t^* > 0$ and $x_t^* < 1/(n-1)$ (the last inequality follows from Assumption 3). The IC constraints of firm 1 hold for all $t < t^*$, since firm 1 sells at full capacity at all such t. For $t \geq t^*$,

$$\sum_{s=t}^{\infty} \delta^{s-t} (1 - (n-1) x_s^*) p_s^* = \frac{1}{1-\delta} (1 - (n-1)(1-\delta) K_f) > K_1 = p_t^* K_1,$$

where the strict inequality follows from Assumption 3. ■

B.4 Differentiated Products

This appendix studies a setting with two firms that sell differentiated products. Firms compete in prices and have a constant marginal cost equal to zero. Firm i's demand is

$$D_i(p_i, p_j) = 1 - b p_i + s p_j,$$

with $b > s > 0$. Firm i's flow profits are $\pi_i(p_i, p_j) = p_i D_i(p_i, p_j)$. Firms discount future profits with a common discount factor $\delta \in (0, 1)$.[12]

I restrict attention to equilibria in which deviations are punished via Nash reversion: at any off-path history, both firms choose the Nash equilibrium price $p^N = 1/(2b - s)$ and earn Nash profits $\pi^N = b/(2b - s)^2$. Define $V^N \equiv \pi^N/(1 - \delta)$.

Let $d_i(p) \equiv \max_{p_i'} \pi_i(p_i', p_j) + \delta V^N = (1 + sp_j)^2/(4b) + \delta V^N$ denote firm i's optimal deviation profits. Fix $\alpha \in [0, 1]$ and $P > 0$ large. The problem of finding the subgame perfect Nash equilibrium with Nash reversion that maximizes the weighted sum of payoffs is

$$V_\alpha \equiv \sup_{(p)\in([0,P]^2)^{\mathbb{N}_0}} \sum_{t=0}^{\infty} \delta^t \left(\pi_1(p_t) + \alpha\pi_2(p_t)\right) \tag{P-diff}$$

$$\text{s.t. } \forall t, \forall i, \quad \sum_{s=t}^{\infty} \delta^{s-t}\pi_i(p_s) \geq d_i(p_t). \tag{IC-diff}$$

Attaching normalized multiplier $\delta^t\lambda_{i,t}$ to firm i's date-t incentive constraint, the Lagrangian associated with (P-diff) is

$$\mathcal{L}(p, \lambda) = \sum_{t=0}^{\infty} \delta^t \left(\pi_1(p_t) + \alpha\pi_2(p_t)\right) + \sum_{i=1}^{2}\sum_{t=0}^{\infty} \delta^t\lambda_{i,t} \left(\sum_{s=t}^{\infty} \delta^{s-t}\pi_i(p_s) - d_i(p_t)\right)$$

The Lagrangian can be re-written as

$$\mathcal{L}(p, \lambda) = \sum_{t=0}^{\infty} \delta^t \left(\left(1 + \sum_{s=0}^{t} \lambda_{1,s}\right) \pi_1(p_t) + \left(\alpha + \sum_{s=0}^{t} \lambda_{2,s}\right) \pi_2(p_t) - \lambda_{1,t}d_1(p_t) - \lambda_{2,t}d_2(p_t)\right). \tag{16}$$

Note that the weight on each firm's profits increases with past multipliers. Intuitively, an increase in $\pi_i(p_t)$ relaxes firm i's incentive constraints at all $s \leq t$, and hence increases the Lagrangian by $\delta^t(\alpha_i + \sum_{s\leq t} \lambda_{i,s})$ (where $\alpha_1 = 1$ and $\alpha_2 = \alpha$). If firm 1's incentive constraints are slack at the equilibrium that solves (P-diff) (which would be true when δ is

[12]For simplicity, this linear demand system allows negative quantities at some price vectors. All equilibrium paths considered below, including those in the numerical calculations, have strictly positive quantities.

high enough), then $\lambda_{1,t} = 0$ for all t, the weight on firm 1 is 1 for all t, and the weight on firm 2 is $\alpha_t = \alpha + \sum_{s=0}^{t} \lambda_{2,s}$. The weight α_t solves $\alpha_t = \alpha_{t-1} + \lambda_{2,t}$, with $\alpha_{-1} = \alpha$.

For $\alpha \geq 0$ and $\lambda \geq 0$, let $p(\alpha, \lambda)$ be given by

$$p(\alpha, \lambda) \in \arg \max_{(p_1,p_2)\in[0,P]^2} \pi_1(p_1, p_2) + \alpha\pi_2(p_1, p_2) - \lambda d_2(p_1). \tag{17}$$

Note that prices $p(\alpha, \lambda)$ maximize the period-t flow profits of (16) when $\lambda_{1,t} = 0, \lambda_{2,t} = \lambda$ and the weight α_t on firm 2's profits is α. The objective in (17) is strictly concave when $4b^2\alpha > s^2(1+\alpha)^2$. For $\alpha \leq 1$, this holds when $\alpha > \underline{\alpha} \equiv \left(\frac{b-\sqrt{b^2-s^2}}{s}\right)^2$ (note that $b > s$ implies $\underline{\alpha} < 1$). For $\alpha \in (\underline{\alpha}, 1]$, when the maximizer is interior, the prices that solve (17) are

$$\begin{aligned} p_1(\alpha, \lambda) &= \frac{\alpha(2b + s(1+\alpha) - \lambda s)}{4b^2\alpha + \lambda s^2\alpha - s^2(1+\alpha)^2}, \\ p_2(\alpha, \lambda) &= \frac{(2b\alpha + s(1+\alpha) - \lambda s^2/(2b))}{4b^2\alpha + \lambda s^2\alpha - s^2(1+\alpha)^2}. \end{aligned}$$

For $i = 1, 2$, let $p_i(\alpha) = p_i(\alpha, 0)$. Define $\Pi_2(\alpha) \equiv \pi_2(p(\alpha))$ and $D_2(\alpha) \equiv d_2(p(\alpha))$ to be, respectively, firm 2's flow profits and deviation profits at prices $p(\alpha) = p(\alpha, 0)$. Lastly, define α^∞ to be a solution to

$$\Pi_2(\alpha^\infty) = (1-\delta)D_2(\alpha^\infty). \tag{18}$$

To see the intuition behind α^∞, consider a strategy profile such that firms charge prices $p(\alpha^\infty, 0)$ at every on-path history, and such that defections are punished by Nash reversion. By equation (18), this strategy profile satisfies firm 2's IC constraints with equality at every period. Moreover, for δ high enough, this strategy profile also satisfies firm 1's IC constraints, and hence is a subgame perfect Nash equilibrium.

Lemma B.5. *There exists $\underline{\delta}_1 < 1$ such that, for every $\delta \in (\underline{\delta}_1, 1)$, there is a solution $\alpha^\infty \in (\underline{\alpha}, 1)$ to (18).*

Proof. As an initial step, I show that there is a unique $\alpha^0 \in (\underline{\alpha}, 1)$ satisfying $\Pi_2(\alpha^0) - \pi^N = 0$. Moreover, α^0 also satisfies $D_2(\alpha^0) > V^N$. Consider the unconstrained maximizers given by (17), with $\lambda = 0$. As $\alpha \downarrow \underline{\alpha}$, $4b^2\alpha - s^2(1+\alpha)^2$ converges to zero from above, and so these

prices diverge to $+\infty$. Since $p_2(\alpha) - p_1(\alpha) = \frac{s(1-\alpha^2)}{4b^2\alpha - s^2(1+\alpha)^2} > 0$ for all $\alpha < 1$, and since $b > s$, firm 2's profit at the unconstrained maximizer goes to $-\infty$ as $\alpha \to \underline{\alpha}$. Therefore, there exists $\alpha^L \in (\underline{\alpha}, 1)$ such that $\Pi_2(\alpha^L) < \pi^N$. Let $P > 0$ be large enough so that the unconstrained maximizers $p_1(\alpha), p_2(\alpha)$ lie in $(0, P)$ for all $\alpha \in [\alpha^L, 1]$. At $\alpha = 1$, the prices $p(\alpha)$ are the joint monopoly prices, $p_1(1) = p_2(1) = \frac{1}{2(b-s)}$. Hence, $\Pi_2(1) > \pi^N$. By continuity, there exists $\alpha^0 \in (\alpha^L, 1)$ with $\Pi_2(\alpha^0) = \pi^N$.

Next, I show that there is a unique such α^0. Prices $p(\alpha)$ maximize $\pi_1(p) + \alpha\pi_2(p)$. Hence,

$$D_p\pi_1(p(\alpha)) + \alpha D_p\pi_2(p(\alpha)) = 0.$$

Differentiating with respect to α gives

$$\left(D^2_{pp}\pi_1(p(\alpha)) + \alpha D^2_{pp}\pi_2(p(\alpha))\right) p'(\alpha) + D_p\pi_2(p(\alpha)) = 0.$$

Since $\pi_1(p) + \alpha\pi_2(p)$ is strictly concave on $[\alpha^L, 1]$, the Hessian is negative definite. Hence

$$\begin{aligned}\Pi_2'(\alpha) &= D_p\pi_2(p(\alpha))p'(\alpha) \\ &= -D_p\pi_2(p(\alpha))^\top \left[D^2_{pp}\pi_1(p(\alpha)) + \alpha D^2_{pp}\pi_2(p(\alpha))\right]^{-1} D_p\pi_2(p(\alpha)) > 0.\end{aligned}$$

The strict inequality holds because $D_p\pi_2(p(\alpha)) \neq 0$; in particular its p_1 component is $sp_2(\alpha) > 0$. Therefore, there is a unique $\alpha^0 \in [\alpha^L, 1]$ with $\Pi_2(\alpha^0) = \pi^N$. Since $\Pi_2(\alpha)$ is weakly increasing in α,[13] it follows that $\Pi_2(\alpha) \leq \Pi_2(\alpha^L) < \pi^N$ for all $\alpha < \alpha^L$, and so there is a unique $\alpha^0 \in (0, 1)$ solving $\Pi_2(\alpha^0) = \pi^N$.

At α^0, firm 2's current flow profit equals its Nash profit. Since firm 2 is not charging a best response to $p_1(\alpha^0)$, $\max_{p_2} \pi_2(p_1(\alpha^0), p_2) > \pi_2(p(\alpha^0)) = \Pi_2(\alpha^0) = \pi^N$. Hence,

$$D_2(\alpha^0) = \max_{p_2} \pi_2(p_1(\alpha^0), p_2) + \delta V^N > \pi^N + \delta V^N = V^N.$$

[13] This follows since $\Pi_2(\alpha) = \pi_2(p(\alpha))$, with $p(\alpha) \in \arg\max_p \pi_1(p) + \alpha\pi_2(p)$.

Next, I show that there exists α^∞ solving (18). By the arguments above, at $\alpha = \alpha^0$,

$$\Pi_2(\alpha^0) - \pi^N = 0, \qquad D_2(\alpha^0) > V^N, \qquad \Pi_2'(\alpha^0) > 0.$$

Note that $(1-\delta)D_2(\alpha^0) = (1-\delta)\max_{p_2}\pi_2(p_1(\alpha^0), p_2) + \delta\pi^N \to \pi^N$ as $\delta \to 1$. Hence, at $\delta = 1$, α^0 solves (18). By the implicit-function theorem, for δ close enough to one, there is a unique solution α^∞ to (18) in an interval $[\alpha^0 - \gamma, \alpha^0 + \gamma]$ with $\gamma > 0$ satisfying

$$\alpha^\infty = \alpha^0 + (1-\delta)\frac{D_2(\alpha^0) - V^N}{\Pi_2'(\alpha^0)} + o(1-\delta).$$

Since $D_2(\alpha^0) > V^N$ and $\Pi_2'(\alpha^0) > 0$, $\alpha^\infty > \alpha^0 > \underline{\alpha}$ and $\alpha^\infty \to \alpha^0$ as $\delta \to 1$. □

Fix $\eta > 0$ small. Since $\Pi_2'(\alpha^0) > 0$ and $\Pi_2(\alpha^0) = \pi^N$ (see proof of Lemma B.5), and since $\Pi_2(\alpha)$ is monotone in α, we have $\Pi_2(\alpha) \geq \Pi_2(\alpha^0 + \eta) > \pi^N$ for all $\alpha \geq \alpha^0 + \eta$ and $\Pi_2(\alpha) \leq \Pi_2(\alpha^0 - \eta) < \pi^N$ for all $\alpha \leq \alpha^0 - \eta$. Let $\overline{d} \equiv \max_{p_j \in [0,P]} \max_{p_i \geq 0} \pi_i(p_i, p_j)$ be a player's maximal flow deviation profit, so that for all p, $d_2(p) \leq \overline{d} + \delta V^N$. Then, there exists $\underline{\delta}_2 < 1$ such that for all $\delta \in (\underline{\delta}_2, 1)$,

$$\begin{aligned} &\forall \alpha \geq \alpha^0 + \eta, \forall p \in [0,P]^2 \quad \Pi_2(\alpha) > (1-\delta)d_2(p), \\ &\forall \alpha \leq \alpha^0 - \eta, \forall p \in [0,P]^2 \quad \Pi_2(\alpha) < \delta\pi^N \leq (1-\delta)d_2(p), \end{aligned} \tag{19}$$

where the last inequality follows since, for all p, $(1-\delta)d_2(p) \geq \delta\pi^N$. Let $\underline{\delta}' = \max\{\underline{\delta}_1, \underline{\delta}_2\} < 1$, with $\underline{\delta}_1$ as defined in Lemma B.5.

Proposition B.4. *There exists $\underline{\delta} < 1$ such that, for every $\delta > \underline{\delta}$, every $\alpha \in [0, \alpha^\infty)$ and any sequences $(\alpha_t, \lambda_t)_{t=0}^\infty$ satisfying $(\lambda_t) \in \ell_+^1$, $\alpha_{-1} = \alpha$ and*

$$\forall t, \quad \alpha_t = \alpha_{t-1} + \lambda_t, \tag{20}$$

$$\forall t, \quad \pi_2\big(p(\alpha_t, \lambda_t)\big) + \delta d_2\big(p(\alpha_{t+1}, \lambda_{t+1})\big) = d_2\big(p(\alpha_t, \lambda_t)\big), \tag{21}$$

the price sequence $(p(\alpha_t, \lambda_t))_{t=0}^\infty$ solves program (P-diff) *for Pareto weight α. Moreover,*

$\lim_{t\to\infty} \alpha_t = \alpha^\infty$.

Proposition B.4 shows that, if we can find sequences $(\alpha_t, \lambda_t)_{t=0}^\infty$ satisfying certain conditions, then the corresponding sequence of prices $(p(\alpha_t, \lambda_t))$ is a solution to (P-diff). The proof of the proposition is below, but the intuition is as follows. Note first that condition (21) implies that prices $(p(\alpha_t, \lambda_t))$ satisfy firm 2's IC constraints with equality at every period. Moreover, for δ high enough (i.e., $\delta > \underline{\delta}$), firm 1's IC constraints are slack under $(p(\alpha_t, \lambda_t))$. Hence, multipliers $(\lambda_{1,t}^*, \lambda_{2,t}^*)$ with $\lambda_{1,t}^* = 0$ and $\lambda_{2,t}^* = \lambda_t$ for all t satisfy the complementary slackness conditions. Since prices $(p(\alpha_t, \lambda_t))$ solve (17) at each period, it follows that $(p(\alpha_t, \lambda_t))$ and $(\lambda_{1,t}^*, \lambda_{2,t}^*)$ are a saddle point of $\mathcal{L}$, and hence $(p(\alpha_t, \lambda_t))$ solves (P-diff).

Figure B.1 provides a numerical solution to (P-diff), based on Proposition B.4. Under the equilibrium that solves this program, the prices of both firms increase over time, firm 1's quantity falls, and firm 2's quantity increases. The figure also plots the evolution of the cumulative weight on firm 2's profits α_t, and the value of multipliers $\lambda_t = \alpha_t - \alpha_{t-1}$.

Proof of Proposition B.4. Let $\eta > 0$ be as defined before the statement of the proposition. Note that, for all $a \in [0, \alpha^0 + \eta]$ and all $l \in [0, a]$, $\max_{p\in[0,P]^2}(\pi_1(p) + a\pi_2(p) - ld_2(p)) > (1+a)\pi^N - lV^N$, where the inequality follows since the Nash price vector is not a maximizer, and since at the Nash price vector $d_2(p) = V^N$. Hence, there exists $\kappa > 0$ such that, for all $a \in [0, \alpha^0 + \eta]$ and all $l \in [0, a]$,

$$\max_{p\in[0,P]^2}(\pi_1(p) + a\pi_2(p) - ld_2(p)) > (1+a)\pi^N - lV^N + \kappa. \tag{22}$$

Let $\underline{\delta} \geq \underline{\delta}'$ be such that, for all $\delta \in (\underline{\delta}, 1)$,

$$\kappa > (1-\delta)(1 + \alpha^0 + \eta)(\overline{d} - \pi^N), \tag{23}$$

with $\underline{\delta}'$ and $\overline{d}$ as defined before the statement of the proposition.

Fix $\delta \in (\underline{\delta}, 1)$, and let (α_t, λ_t) be sequences satisfying the conditions in the statement of the proposition, and let $(p_t^*) = (p(\alpha_t, \lambda_t))$. Since $(\lambda_t) \in \ell_+^1$, there exists $\overline{\alpha} \geq \alpha$ such that

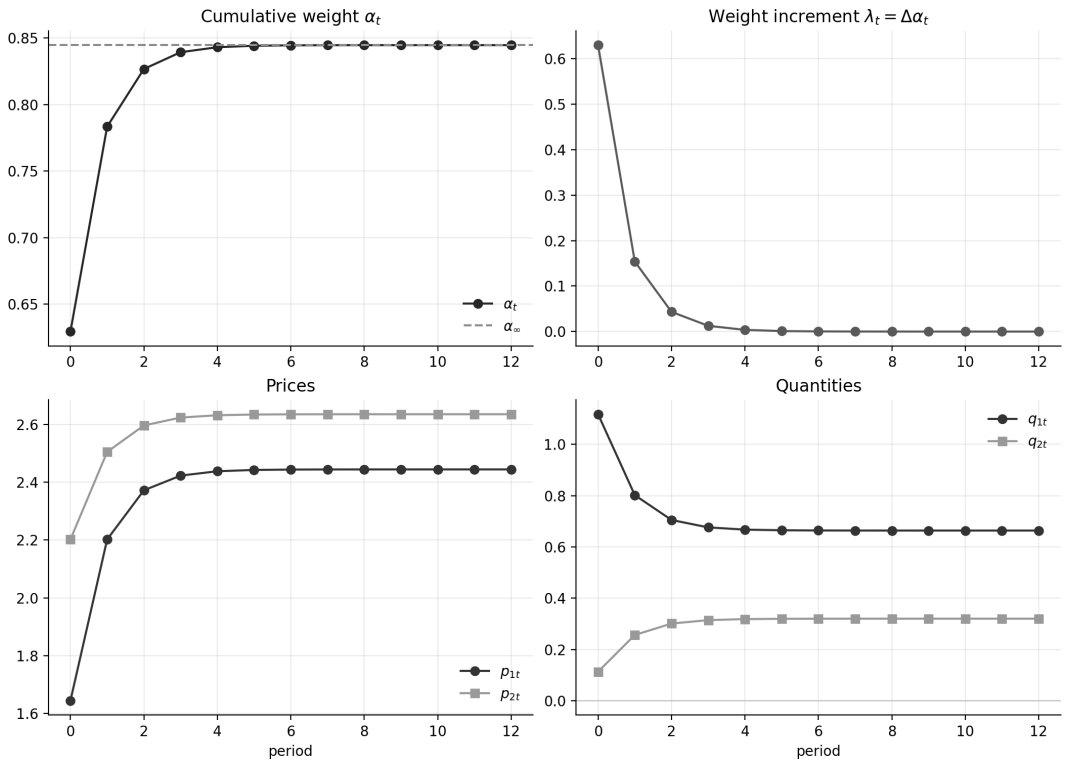

Figure B.1: Pareto-optimal equilibrium with differentiated products. Parameters: $b = 1, s = 0.8, \delta = 0.9, \alpha = 0$.

$\overline{\alpha} = \lim_{t\to\infty} \alpha_t = \alpha + \sum_{t=0}^{\infty} \lambda_t$.

Note that (21) implies that, for all t and $T > t$,

$$d_2(p_t^*) = \sum_{s=t}^{T} \delta^{s-t} \pi_2(p_s^*) + \delta^{T+1-t} d_2(p_{T+1}^*).$$

Since $\lim_{T\to\infty} \delta^{T+1-t} d_2(p_{T+1}^*) = 0$ (since prices are bounded), it follows that

$$\forall t, \quad \sum_{s=t}^{\infty} \delta^{s-t} \pi_2(p_s^*) = d_2(p_t^*). \tag{24}$$

In words, under $(p_t^*) = (p(\alpha_t, \lambda_t))$, firm 2's IC constraint holds with equality at every period.

Next, I show that $\overline{\alpha} = \lim_{t\to\infty} \alpha_t = \alpha^\infty$. As a first step, I show that $\overline{\alpha} \in (\alpha^0 - \eta, \alpha^0 + \eta)$, with $\eta > 0$ as defined before the statement of the proposition. Suppose by contradiction that $\overline{\alpha} \notin (\alpha^0 - \eta, \alpha^0 + \eta)$. Since $\alpha_t \to \overline{\alpha}$ and $\lambda_t \to 0$, the static objectives $\pi_1(p) + \alpha_t \pi_2(p) - \lambda_t d_2(p)$

converge uniformly on $[0,P]^2$ to $\pi_1(p)+\overline{\alpha}\pi_2(p)$. Therefore, every limit point q of (p_t^*) satisfies

$$q \in \arg\max_{p\in[0,P]^2}\left\{\pi_1(p)+\overline{\alpha}\pi_2(p)\right\}.$$

Since firm 2's profits at a solution to $\max_{p\in[0,P]^2}\left\{\pi_1(p)+\alpha\pi_2(p)\right\}$ are weakly increasing in α, $\overline{\alpha} \notin (\alpha^0-\eta,\alpha^0+\eta)$ implies that $\pi_2(q) \notin (\Pi_2(\alpha^0-\eta),\Pi_2(\alpha^0+\eta))$ for every such limit point q. Therefore, we have that

$$\liminf_{t\to\infty}\pi_2(p_t^*) \geq \Pi_2(\alpha^0+\eta) \text{ or } \limsup_{t\to\infty}\pi_2(p_t^*) \leq \Pi_2(\alpha^0-\eta).$$

Together with (24), this implies

$$\liminf_{t\to\infty}(1-\delta)d_2(p_t^*) \geq \Pi_2(\alpha^0+\eta) \text{ or } \limsup_{t\to\infty}(1-\delta)d_2(p_t^*) \leq \Pi_2(\alpha^0-\eta),$$

which contradicts (19). Hence, $\overline{\alpha} \in (\alpha^0-\eta,\alpha^0+\eta)$. The static maximizer $p(\alpha,0)$ solving (17) is unique on this interval, and so $p_t^* \to p(\overline{\alpha},0)$. Using this in (24) gives

$$D_2(\overline{\alpha}) = \lim_{t\to\infty} d_2(p_t^*) = \lim_{t\to\infty}\sum_{s=t}^{\infty}\delta^{s-t}\pi_2(p_s^*) = \frac{\Pi_2(\overline{\alpha})}{1-\delta}.$$

Thus, $\overline{\alpha}$ solves (18). Since (18) has a unique solution in $(\alpha^0-\eta,\alpha^0+\eta)$, $\overline{\alpha}=\alpha^\infty$.

I now show that $(p_t^*)=(p(\alpha_t,\lambda_t))$ solves (P-diff) with Pareto weight α. For each t, let $\lambda_{1,t}^*=0$ and $\lambda_{2,t}^*=\lambda_t$. To establish the result, I show that: (i) multipliers $(\lambda_{1,t}^*,\lambda_{2,t}^*)$ and prices $(p_t^*)=(p(\alpha_t,\lambda_t))$ are a saddle point of $\mathcal{L}$; and (ii) $(p_t^*)=(p(\alpha_t,\lambda_t))$ satisfies the IC constraints in Program (P-diff). This implies that $(p_t^*)=(p(\alpha_t,\lambda_t))$ is a solution to (P-diff).

Note first that, by definition, $p_t^*=p(\alpha_t,\lambda_t) \in \arg\max_{p\in[0,P]^2}\pi_1(p)+\alpha_t\pi_2(p)-\lambda_t d_2(p)$. Hence, by (16), $\mathcal{L}(p^*,\lambda_1^*,\lambda_2^*) \geq \mathcal{L}(p,\lambda_1^*,\lambda_2^*)$ for any other sequence $(p)\in([0,P]^2)^{\mathbb{N}_0}$.

The arguments above show that firm 2's IC constraint is satisfied with equality at every

period under (p_t^*). Hence, the following complementary slackness condition holds:

$$\forall t, \quad \lambda_{2,t}^* \left(\sum_{s=t}^{\infty} \delta^{s-t} \pi_2(p_s^*) - d_2(p_t^*) \right) = 0.$$

Next, note that by construction, for all t, $\alpha_t \in [0, \alpha^\infty]$ and $\lambda_t \in [0, \alpha^\infty]$. Using (22) at $(a, l) = (\alpha_t, \lambda_t)$ it follows that

$$\forall t, \quad \pi_1(p_t^*) + \alpha_t \pi_2(p_t^*) - \lambda_t d_2(p_t^*) > (1 + \alpha_t)\pi^N - \lambda_t V^N + \kappa. \tag{25}$$

Summing the right-hand side of (25) forward using discount factor δ yields

$$\begin{aligned} \sum_{s=t}^{\infty} \delta^{s-t}((1 + \alpha_s)\pi^N - \lambda_s V^N + \kappa) &= V^N \sum_{s=t}^{\infty} \delta^{s-t}((1 + \alpha_s)(1 - \delta) - (\alpha_s - \alpha_{s-1})) + \frac{\kappa}{1 - \delta} \\ &= V^N(1 + \alpha_{t-1}) + \frac{\kappa}{1 - \delta} \end{aligned}$$

Similarly, summing the left-hand side of (25) forward using discount factor δ yields

$$\begin{aligned} \sum_{s=t}^{\infty} \delta^{s-t}(\pi_1(p_s^*) + \alpha_s \pi_2(p_s^*) - \lambda_s d_2(p_s^*)) &= \sum_{s=t}^{\infty} \delta^{s-t} \pi_1(p_s^*) + \sum_{s=t}^{\infty} \delta^{s-t}(\alpha_s \pi_2(p_s^*) - (\alpha_s - \alpha_{s-1}) d_2(p_s^*)) \\ &= \sum_{s=t}^{\infty} \delta^{s-t} \pi_1(p_s^*) + \sum_{s=t}^{\infty} \delta^{s-t}(\alpha_{s-1} d_2(p_s^*) - \delta \alpha_s d_2(p_{s+1}^*)) \\ &= \sum_{s=t}^{\infty} \delta^{s-t} \pi_1(p_s^*) + \alpha_{t-1} d_2(p_t^*), \end{aligned}$$

where the second equality uses (21). Combining these two equalities with (25) yields

$$\forall t, \quad \sum_{s=t}^{\infty} \delta^{s-t} \pi_1(p_s^*) - d_1(p_t^*) > \frac{\kappa}{1 - \delta} + V^N(1 + \alpha_{t-1}) - \alpha_{t-1} d_2(p_t^*) - d_1(p_t^*).$$

By (23), the right-hand side is strictly positive for all $\delta > \underline{\delta}$, and so firm 1's IC constraints

are slack for all t. Hence, since $\lambda^*_{1,t} = 0$ for all t, it follows that

$$\forall t, \quad \lambda^*_{1,t} \left(\sum_{s=t}^{\infty} \delta^{s-t} \pi_1(p^*_s) - d_1(p^*_t) \right) = 0.$$

These complementary slackness conditions imply that $\mathcal{L}(p^*, \lambda^*_1, \lambda^*_2) \leq \mathcal{L}(p^*, \lambda_1, \lambda_2)$ for all $(\lambda_1, \lambda_2) \in (\ell^1_+)^2$. Hence, $(p^*_t) = (p(\alpha_t, \lambda_t))$ solves (P-diff). This completes the proof. ■

B.5 Proofs of Section 3

Proof of Lemma 1. Suppose $(p_t, \mathbf{x}_t) \in \mathcal{F}$ is implementable by some strategy profile $\sigma \in \Sigma$. Then, for every on-path 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 $\sigma \in \Sigma$, it must be that for all i and all on-path histories h_t

$$\sum_{s=t}^{\infty} \delta_i^{s-t} x_{i,s} p_s \geq p_t.$$

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

Next, let $(p_t, \mathbf{x}_t) \in \mathcal{F}$ 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 $\sigma \in \Sigma$. Hence, if $(p_t, \mathbf{x}_t)$ satisfies (IC), it is implementable by some strategy profile $\sigma \in \Sigma$. ■

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

$$\forall t, \forall i, \quad p_t \leq \sum_{s=t}^{\infty} \delta_i^{s-t} x_{i,s} p_s. \tag{26}$$

Summing (26) across all $i \neq 1$ and dividing by $n-1$ yields

$$\forall t, \quad p_t \leq \sum_{s=t}^{\infty} \delta^{s-t} \left(\frac{\sum_{i \neq 1} x_{i,s}}{n-1} \right) p_s \tag{27}$$

Consider the sequence of allocations $(\widehat{\mathbf{x}}_t)$ such that $\widehat{x}_{1,t} = x_{1,t}$ for all t, and, for all $i \neq 1$, $\widehat{x}_{i,t} = \frac{\sum_{j \neq 1} x_{j,t}}{n-1}$. By (27), $(p_t, \widehat{\mathbf{x}}_t)$ satisfies (IC). By Lemma 1, there exists an equilibrium $\widehat{\sigma} \in \Sigma$ that implements $(p_t, \widehat{\mathbf{x}}_t)$. By construction $V_1(\widehat{\sigma}) = V_1(\sigma)$, $\sum_{i \neq 1} V_i(\widehat{\sigma}) = \sum_{i \neq 1} V_i(\sigma)$ 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 I can choose $\widehat{\sigma}$ such that, off the equilibrium path, firms play the static Nash equilibrium. Hence, for all h and all $i, j \neq 1$, $\pi_i(\widehat{\sigma}|h) = \pi_j(\widehat{\sigma}|h)$. ■

B.6 Proofs of Propositions 1 and 5

Proof of Proposition 1. I first show that any equilibrium $\sigma \in \Sigma^{\mathsf{st}}$ 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 $\sigma \in \Sigma^{\mathsf{st}}$ 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 $(\mathbf{x}_t)_{t=0}^{\infty}$ denote the market shares along the path of play under σ, with $\mathbf{x}_t = (x_{i,t})_{i \in N}$. Then, by Lemma 1 and $p_t = p \in (0, 1]$ for all t, it must be that

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

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 the problem of finding the equilibrium in Σ^{st} that maximizes player 1's profits as

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

$$\text{s.t. (IC-}\mathsf{st}\text{)}$$

Fix $M > 0$ large. 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:

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

$$\text{s.t. } \forall t, \quad 1 \leq \sum_{s=t}^{\infty} \delta^{s-t} x_s \qquad \text{(IC-}\mathsf{st}\text{-fringe)}$$

The Lagrangian associated with this program is

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

where $\lambda = (\lambda_t)_{t=0}^{\infty}$ is a sequence of non-negative multipliers associated with the constraints in (P'_{st}). To establish the Proposition, I find multipliers $\lambda^* \in \ell_+^1$ that, together with the allocations x^* in the statement of the Proposition, satisfy the saddle-point condition:

$$\forall \lambda \in \ell_+^1, \forall x \in [-M, M]^{\mathbb{N}_0}, \qquad \mathcal{L}(x^*, \lambda) \geq \mathcal{L}(x^*, \lambda^*) \geq \mathcal{L}(x, \lambda^*), \qquad (28)$$

This implies that x^* solves (P'_{st}). 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, and $\sum_{t=0}^{\infty} \lambda_t^* < \infty$. Hence, $\lambda^* \in \ell_+^1$. I use the following Lemma.

Lemma B.6. *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 $t = 0$, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau^* = \lambda_0^* = (n-1)\delta_1^0$. Next, let $t \geq 1$ and suppose $\sum_{\tau=0}^{s} \delta^{s-\tau}\lambda_\tau^* = (n-1)\delta_1^s$ for $s = 0, \ldots, t-1$. 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 $t \geq 0$, $\sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau^* = (n-1)\delta_1^t$. □

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 (28). Note first that, for all t,

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

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

$$\forall t, \quad \lambda_t^* \left(\sum_{s=t}^{\infty} \delta^{s-t} x_s^* - 1 \right) = 0.$$

This implies that, for all $\lambda \in \ell_+^1$, $\mathcal{L}(x^*, \lambda) = \mathcal{L}(x^*, \lambda^*)$, so the first inequalities in (28) hold.

Note next that $\sum_{t=0}^{\infty} \lambda_t \sum_{s=t}^{\infty} \delta^{s-t} x_s = \sum_{t=0}^{\infty} x_t \sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau$. Hence, the Lagrangian $\mathcal{L}(x, \lambda^*)$ 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 B.6. 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,

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

where the strict inequality follows from Assumption 1. ■

Proof of Proposition 5. Fix $M > 0$ large. Let $\mathbb{F}$ denote the filtration generated by (r_t) and define $\ell_{\mathbb{F}}^{\infty}(M) \equiv \{(a_t)_{t\geq 0} : (a_t) \text{ is adapted to } \mathbb{F} \text{ and } a_t \in [-M, M] \forall t\}$. Consider the following relaxation of Program (P_s):

$$\begin{aligned} \widehat{V}_1 \equiv \sup_{(x,p)\in(\ell_{\mathbb{F}}^{\infty}(M))^2} & \mathbb{E}_0\left[\sum_{t=0}^{\infty} \delta_1^t(1-(n-1)x_t)p_t\right] \\ \text{s.t. } \forall t, \quad & p_t \leq \mathbb{E}_t\left[\sum_{s=t}^{\infty} \delta^{s-t} x_s p_s\right] \\ \forall t, \quad & p_t \leq r_t, p_t x_t \geq 0. \end{aligned} \tag{P'_s}$$

The Lagrangian associated with this program is

$$\begin{aligned} \mathcal{L}(x,p,\lambda,\mu,\gamma) &= \mathbb{E}_0\left[\sum_{t=0}^{\infty} \delta_1^t(1-(n-1)x_t)p_t + \sum_{t=0}^{\infty} \mathbb{E}_t\left[\lambda_t\left(\sum_{s=t}^{\infty} \delta^{s-t}x_s p_s - p_t\right)\right]\right. \\ &\qquad \left. + \sum_{t=0}^{\infty} \mathbb{E}_t\left[\mu_t p_t x_t\right] + \sum_{t=0}^{\infty} \mathbb{E}_t\left[\gamma_t(r_t - p_t)\right]\right] \\ &= \mathbb{E}_0\left[\sum_{t=0}^{\infty} p_t x_t\left(-(n-1)\delta_1^t + \mu_t + \sum_{\tau=0}^{t} \delta^{t-\tau}\lambda_\tau\right) + \sum_{t=0}^{\infty} p_t\left(\delta_1^t - \lambda_t - \gamma_t\right) + \sum_{t=0}^{\infty} r_t\gamma_t\right], \end{aligned} \tag{29}$$

where the equality uses the Law of Iterated Expectations and $\sum_{t=0}^{\infty}\lambda_t\sum_{s=t}^{\infty}\delta^{s-t}x_s p_s = \sum_{t=0}^{\infty} x_t p_t \sum_{\tau=0}^{t}\delta^{t-\tau}\lambda_\tau$. I note that (λ, μ, γ) are now stochastic processes adapted to (r_t);

i.e., $\lambda, \mu, \gamma \in \ell^1_{\mathbb{F},+} \equiv \{a = (a_t)_{t\geq 0} : (a_t) \text{ is adapted to } \mathbb{F}, a_t \geq 0 \text{ a.s., and } \sum_{t=0}^{\infty} \mathbb{E}[a_t] < \infty\}$.

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,

$$\begin{aligned}
&\forall (\lambda, \mu, \gamma) \in (\ell^1_{\mathbb{F},+})^3, \forall (x, p) \in (\ell^\infty_{\mathbb{F}}(M))^2 \\
&\quad \mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*) \geq \mathcal{L}(x, p, \lambda^*, \mu^*, \gamma^*). \qquad (30)
\end{aligned}$$

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.2.

Lemma B.7. *(i)* $(\lambda^*, \mu^*, \gamma^*) \in (\ell^1_{\mathbb{F},+})^3$*;*

(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 $p^*_t = \delta^{t^*-t}\overline{r}$ 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 B.8. *(x^*, p^*) satisfies the constraints in program (P'_s). Moreover, for all t,*

$$\mathbb{E}_t\left[\lambda_t^*\left(\sum_{s=t}^{\infty}\delta^{s-t}x_s^*p_s^* - p_t^*\right)\right] = 0, \qquad \mathbb{E}_t\left[\mu_t^*p_t^*x_t^*\right] = 0, \qquad \mathbb{E}_t[\gamma_t^*(r_t - p_t^*)] = 0.$$

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 $t < t^*$,

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

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, and since λ_t^* is deterministic, 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 holds since $\gamma_t^* = 0$ for all $t < t^*$, and since $p_t^* = r_t$ for all $t \geq t^*$. □

The complementary slackness conditions in Lemma B.8, and the fact that (x^*, p^*) satisfies all the constraints in (P'_s), imply that for all $\lambda, \mu, \gamma \in \ell^1_{\mathbb{F},+}$, $\mathcal{L}(x^*, p^*, \lambda, \mu, \gamma) \geq \mathcal{L}(x^*, p^*, \lambda^*, \mu^*, \gamma^*)$. Hence, the first set of inequalities in (30) holds.

Next, note that equation (29) and Lemma B.7 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 (30) 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

$$p_t^* < \mathbb{E}_t \left[\sum_{s=t}^{\infty} \delta_1^{s-t} (1 - (n-1) x_s^*) p_s^* \right],$$

where the inequality uses $x_s^* = 0$ for all $s < t^*$ and $x_s^* = x(r_s) < 1/(n-1)$ for all $s \geq t^*$. For $t \geq t^*$, $x_t^* = x(r_t) = \frac{r_t - \delta \overline{r}}{r_t}$ and $p_t^* = r_t$, and so

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

where the strict inequality follows since

$$\begin{aligned} \frac{\delta_1}{1-\delta_1} \sum_k q_k r_k \left(1 - (n-1) \frac{r_k - \delta \overline{r}}{r_k} \right) = \frac{\delta_1}{1-\delta_1} \overline{r} ((n-1)\delta - (n-2)) > (n-1)(r_K - \delta \overline{r}) \\ \iff \delta_1 ((n-1) r_K - (n-2) \overline{r}) > (n-1)(r_K - \delta \overline{r}), \end{aligned}$$

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