No-regret Learning in Price Competitions under Consumer Reference Effects

Negin Golrezaei
Sloan School of Management, Massachusetts Institute of Technology golrezaei@mit.edu,
Patrick Jaillet
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology jaillet@mit.edu,
Jason Cheuk Nam Liang
Operations Research Center, Massachusetts Institute of Technology jcnliang@mit.edu,

We study long-run market stability for repeated price competitions between two firms, where consumer demand depends on firms’ posted prices and consumers’ price expectations called reference prices. Consumers’ reference prices vary over time according to a memory-based dynamic, which is a weighted average of all historical prices. We focus on the setting where firms are not aware of demand functions and how reference prices are formed but have access to an oracle that provides a measure of consumers’ responsiveness to the current posted prices. We show that if the firms run no-regret algorithms, in particular, online mirror descent (OMD), with decreasing step sizes, the market stabilizes in the sense that firms’ prices and reference prices converge to a stable Nash Equilibrium (SNE). Interestingly, we also show that there exist constant step sizes under which the market stabilizes. We further characterize the rate of convergence to the SNE for both decreasing and constant OMD step sizes.

Key words: price competition, consumer reference effect, no-regret learning, convergence in games

1. Introduction

In markets with repeated consumer-seller interactions, consumers develop price expectations (or reference prices) based on past observed prices. Such price memories would influence consumers’ willingness-to-pay and hence their purchasing decisions, eventually impacting the overall aggregate market demand. Due to such memory dependent reference price effects, developing pricing strategies is challenging because firms may not necessarily know how consumers form and adjust price expectations. The complexity of pricing is further increased with competition, as competitors’ pricing decisions impact not only a firm’s immediate demand but also consumers’ reference prices. Such challenges in pricing under competition and reference price effects make market stability particularly attractive to firms: under stable markets, long-term organizational planning and business strategy development
can be conducted more effectively (see Caves and Porter (1978)). Inspired by this, in this paper, we study the impact of consumer reference prices on the long-term stability of competitive markets.

We examine a simplified market scenario where two firms sequentially set prices to sell goods over an infinite time horizon, and demand of each firm’s goods are influenced by both firms’ current prices and the current consumers’ reference price, which is a weighted average of all past price trajectories. Also, the repeated price competitions occur in an opaque environment, where firms are not aware of any demand or reference price characteristics, and only have access to an oracle that returns consumers’ responsiveness to posted prices\(^1\). In such a market scenario, we consider that both firms run a general online mirror descent (OMD) algorithm\(^2\). Despite its simplicity, an OMD algorithm has been theoretically shown to have good performance guarantees in both purely stochastic and adversarial environments (see for example Bubeck and Slivkins (2012), Zimmert and Seldin (2018)), and hence would be a plausible option for firms in this opaque environment of interest.

Our goal is to investigate whether firms’ prices and consumer reference prices eventually stabilize in the long-run if firms run OMD. The notion of stability that we consider is represented by the convergence of firms’ price profiles and reference prices such that there is no incentive for firms to deviate, eliminating the possibility for long-run price cycles and fluctuations. Similar notions of stability under dynamic competition has been studied under various equilibria frameworks, and most relevant to this work are Markov perfect equilibrium and stationary equilibrium (see for example Hopenhayn (1992), Escobar (2007), Doraszelski and Satterthwaite (2010), Weintraub et al. (2011), Adlakha et al. (2015)). Nevertheless, these frameworks assume firms have complete information or optimize pricing decisions according to some prior on their competitors (or their aggregate) and the market. In contrast, our work focuses on competition in an opaque environment where firms post prices using no-regret learning algorithms like OMD. Here, we point out that our objective is not to present dynamic pricing policies that maximize firms’ cumulative revenue. Instead, we seek to shed

\(^1\) We consider a linear demand model, but firms are not aware of the functional form of the demand.

\(^2\) OMD algorithms are closely related to the regularized learning paradigm, which includes algorithms such as follow the regularized leader (FTRL), EXP3, Hedge, etc (see Hoi et al. (2018) for a comprehensive survey).
light on whether simple pricing policies like OMD that do not require a large amount of information eventually achieve market stability. Our contributions are summarized as follows.

- We characterize stability for dynamic competitive markets under consumers’ reference price effects by defining the notion of *Stable Nash Equilibrium (SNE)*. We theoretically demonstrate its existence and shed light on its structural properties; see Theorem 3.1.

- We transform the two-firm game with a dynamic state (reference price) that varies in time according to firms’ posted prices to a three-firm game without a state. The added virtual firm, which is referred to as nature, runs OMD with a constant step size (i.e. has a fast learning rate), and models how reference prices are affected by firms’ past pricing decisions.

- We show that prices and reference prices converge to an SNE and achieve stable markets when the two (real) firms adopt decreasing step sizes that go to zero at a moderate rate; see Theorem 5.1 for details. We further show that with decreasing step sizes, the market stabilizes at a linear rate. We highlight that obtaining these convergence results is challenging because in our three-firm game, there is a firm (nature) who adopts a constant step size and learns at a fast rate. Our results show that despite the need to deal with such an inflexible virtual firm, the real firms can stabilize the market by adopting decreasing step sizes. In fact, the existence of the inflexible virtual firm in our game does not allow us to use the results in the literature on multi-agent online learning, where multiple interacting agents make sequential decisions via running the OMD algorithm to maximize individual rewards (see Mertikopoulos and Staudigl (2017), Bravo et al. (2018), Mertikopoulos and Zhou (2019)). In the multi-agent online learning literature, the games are stateless and agents are required to take decreasing step sizes, i.e. unlike our setting, there is no inflexible agent.

- Interestingly, we also show that there exist constant step sizes under which markets will converge to an SNE at much faster rates compared to adopting decreasing step sizes. Additionally, we show through an example that not every constant step size results in a stable market. Roughly speaking, if the firms’ constant step size is compatible with nature’s constant step size, the market stabilizes at a faster rate compared to decreasing step sizes; see Corollary 5.3.1 and Theorem 5.4 for details.
Related Work

Reference Price Effects and Monopolist Pricing. Consumer reference effects have been validated empirically in many works including Tversky and Kahneman (1979, 1992), Kalyanaraman and Winer (1995), Baron et al. (2019). This motivated a wide range of research including Kopalle et al. (1996), Fibich et al. (2003), Popescu and Wu (2007), Ahn et al. (2007), Nasiry and Popescu (2011) that studies optimal dynamic monopolistic pricing under different demand and reference price update models, where the single firm has complete information on consumer demand as well as how reference prices update. There are also very recent works that address the dynamic pricing problem with consumer reference effects under uncertain demand. Baron et al. (2019) utilizes real retail data and concludes the inclusion of exposure effects to sales or number of consumers\(^3\) when considering reference price formations leads to more accurate forecasts in demand, and proposes a pricing policy using dynamic programming. den Boer and Keskin (2019) couples the problem of monopolistic dynamic pricing with reference effects and online demand learning. In our work, similar to Baron et al. (2019), den Boer and Keskin (2019), firms do not know the demand functions and how the reference prices are formed. But, while in Baron et al. (2019), den Boer and Keskin (2019), the form of demand model is known to the firm (monopolist) that aims to estimate model parameters, our work assumes competition between firms that do not know the form of demand and hence run OMD algorithms to increase revenue. Additionally, the algorithms proposed in Baron et al. (2019) and den Boer and Keskin (2019) aim to increase revenue from the firm’s perspective, while our work focuses on analyzing market stability for long-run competitions under reference effects.

Pricing in Competitive Markets without Reference Effects. A large stream of work studies static price competitions and characterizes structural properties of corresponding equilibria (for example, see Bernstein and Federgruen (2004), Gallego et al. (2006), Aksoy-Pierson et al. (2013)). Other works such as Adida and Perakis (2010), Levin et al. (2009), Gallego and Hu (2014) study

\(^3\) Exposure effects in reference price formation refer to considering reference prices as a weighted average of all historical prices, where weights depend on factors such as sales or number of consumers.
oligopolistic dynamic pricing under various inventory, market, or product characteristics. Nevertheless, these two lines of works are oblivious to consumer reference effects. In this work, we jointly tackle the dynamic pricing problems in competitive markets with reference price effects when the firms lack the knowledge of demand functions and reference price dynamics.

**Pricing in Competitive Markets with Reference Effects.** Similar to our work, the works of Coulter and Krishnamoorthy (2014) and Federgruen and Lu (2016) also consider price competitions under reference effects. Coulter and Krishnamoorthy (2014) considers a similar linear demand model and an identical reference price update dynamic, but the work only provides theoretical analysis on the two-firm, two-period price competition setting, for which they characterize the unique sub-game perfect Nash Equilibrium. On the other hand, Federgruen and Lu (2016) studies multiple-firm single-period price competition equipped with different reference price effects in consumers’ demand (e.g. the reference price is specified by the lowest posted price). Additionally, both of these works study the complete information setting. In contrast to these two papers, our work studies price competitions over an infinite time horizon where reference prices adjust over time, and provides theoretical guarantees for the convergence of pricing strategies under the partial information setting. Finally, our work is the first study that provides theoretical analyses on long-term market stability of repeated price competitions in the presence of consumer reference effects.

**Convergence in Games with Descent Methods.** In addition to Mertikopoulos and Staudigl (2017), Bravo et al. (2018), Mertikopoulos and Zhou (2019) that we discussed earlier, here we also review related literature that study convergence in games where multiple agents adopt descent methods. Rosen (1965) studies finding a Nash Equilibrium of concave games via having each agent run projected gradient descent under complete information, i.e., agents know each others’ payoff functions and decision constraints. Nedic and Ozdaglar (2009) studies a distributed network optimization problem to optimize a sum of convex objective functions corresponding to multiple agents. Our paper distinguishes itself from this line of work from two aspects: unlike the two aforementioned works, (i) our model involves a varying underlying state (i.e., reference prices) dependent on all
agents’ historical decisions, and can be modeled as a sequence of decisions made by an inflexible virtual agent that adopts descent methods with a constant step size; (ii) the agents (i.e., firms) in our model do not have any information on one another’s revenue function or how reference prices update. Finally, Balseiro and Gur (2019) considers multiple budget-constrained bidders participating in repeated second price auctions by adopting so-called adaptive pacing strategies, which is equivalent to the subgradient descent method. In their setting, the subgradient for each bidder’s objective is a function of all bidders’ decisions as well as its budget rate (i.e. total fixed budget divided by a given time horizon), which can be thought of as an underlying model state that remains constant over time. In contrast, in our setting, the gradient oracle each firm receives is not only a function of all firms’ decisions, but also of the reference price which varies over time according to firms’ past decisions, making our analysis more challenging.

2. Preliminaries

Consumer Demand and Reference Price Update Dynamics. We study a dynamic system where two firms simultaneously set prices in each period over an infinite time horizon to sell goods to consumers whose willingness-to-pay is affected by their price expectations, referred to as reference prices. We assume that the number of consumers is large so that demand for each firm is governed by the aggregate behavior of all consumers. Specifically, the demand of firm $i \in \{1, 2\}$ in time period $t$ with posted prices $p_t = (p_{1,t}, p_{2,t})$ and consumers’ reference price $r_t$ is given by

$$d_i(p_{i,t}, p_{-i,t}, r_t) = \alpha_i - \beta_i p_{i,t} + \delta_i p_{-i,t} + \gamma_i r_t,$$  

where $p_{i,t}$ is the price of firm $i$ and $p_{-i,t}$ is the price of the other firm. To simplify notation, we may denote $d_i(p_{i,t}, p_{-i,t}, r_t)$ with $d_i(p_t, r_t)$. We assume prices $p_{i,t}$ and reference prices $r_t$ are bounded, i.e., for $i \in \{1, 2\}$, $p_{i,t}, r_t \in \mathcal{P} = [p, \bar{p}]$ for some $0 < p < \bar{p} < \infty$, and $d_i(p_{i,t}, p_{-i,t}, r) \geq 0$ for any $p_{i,t}, p_{-i,t}, r \in \mathcal{P}$.

The boundedness of prices corresponds to real-world price floors or price caps and is not unnatural. In

Note to run OMD algorithms in Balseiro and Gur (2019), agents need to know the length of the time horizon. Such knowledge is not required in our setting.
Equation (1), \( \alpha_i, \delta_i, \gamma_i > 0 \) and \( \beta_i \geq m (\delta_i + \gamma_i) \), where \( m > 0 \). Later in this section, we will provide an interpretation for these parameters that characterize our linear demand model. We note that linear demand models, which are widely used in the literature (see Huang et al. (2013) for a comprehensive survey), can be viewed as a first-order approximation to more complex models.

After firms post prices, reference prices update according to the following dynamics:

\[
    r_{t+1} = ar_t + (1 - a)(\theta_1 p_{1,t} + \theta_2 p_{2,t})
\]

(2)

where \( \theta_1, \theta_2, a \in (0, 1) \) and \( \theta_1 + \theta_2 = 1 \). Here, \( \theta_i \), which is independent of prices, represents how visible firm \( i \) is to consumers: the larger the \( \theta_i \), the more visible firm \( i \) is, and the more it influences consumers’ price expectations. The reference price update dynamics can be viewed as a memory-based process that characterizes how consumers adjust price expectations for goods over time as they observe new prices. Reference prices are formed by a weighted average of historical prices, where more recent prices are assigned larger weights. The specific exponential weighting scheme adopted in this paper has been empirically validated in behavioral economics (see, for example, Winer (1986), Sorger (1988), Greenleaf (1995)). The parameter \( a \) in the reference price update model characterizes to what extent consumers’ reference price depends on past prices: As \( a \) increases, the reference prices depend less on recently observed prices. Empirical estimates of \( a \) typically range from 0.47 to 0.925 (see Greenleaf (1995), Briesch et al. (1997)) depending on the type of goods sold.

We now provide an economic interpretation for our linear demand model by rearranging terms:

\[
d_i(p_{i,t}, p_{-i,t}, r_t) = \alpha_i - (\beta_i - \gamma_i) p_{i,t} + \delta_i p_{-i,t} + \gamma_i (r_t - p_{i,t}).
\]

(3)

When the posted price is greater than the reference price, i.e., \( p_{i,t} > r_t \), the value \( p_{i,t} - r_t \) can be viewed as the consumers’ perceived price surcharge w.r.t. the reference price, and when \( p_{i,t} < r_t \), the value \( r_t - p_{i,t} \) is consumers’ perceived price discount. Observe that in this rearrangement, demand increases when consumers’ perceived price discount \( (r_t - p_{i,t})\mathbb{I}\{r_t > p_{i,t}\}\) increases, and decreases as price surcharge \( (p_{i,t} - r_t)\mathbb{I}\{p_{i,t} > r_t\} \) increases, which is a conventional representation of how reference prices affect consumer decisions in related literature, see, for example, Popescu and Wu (2007),
Furthermore, the coefficients $\beta_i - \gamma_i, \delta_i,$ and $\gamma_i$ measure the demand sensitivity of firm $i$ to its own prices $p_{i,t}$, its competitor’s prices $p_{-i,t}$, and price surcharge/discount respectively. With these interpretations, parameter $m > 0$ in the condition of $\beta_i \geq m(\delta_i + \gamma_i)$ can be viewed as a sensitivity margin that represents to what extent demand is more sensitive to a firm’s own prices relative to competitor’s prices and surcharge/discount. Take for example the case where $m = 1$: we have $\beta_i - \gamma_i > \delta_i$, which means demand is more affected by a firm’s own prices relative to its competitor’s prices (see Equation (3)). Additionally, for $m = 2$, we have both $\beta_i - \gamma_i > \delta_i$ and $\beta_i - \gamma_i > \gamma_i$, which represents the fact that demand is most affected by a firm’s own prices relative to competitor’s prices and surcharge/discounts. In this work, unless stated otherwise, we consider $m \geq 2$. In particular, we only use this condition for showing our convergence results in Section 5.

**Market Stability.** In this work, our goal is to present simple pricing policies for the firms that stabilize the market even when firms do not have complete information on market conditions. Define $\pi_i(p, r) := p_i \cdot d_i(p, r)$ as the single-period firm $i$’s revenue when prices are $p$ and the reference price is $r$. We say the market is stable at point $(p^*, r^*)$ if the following two conditions hold:

1. **Best-response Conditions.** for $i \in \{1, 2\}$, we have $\pi_i(p^*_i, p^*_{-i}, r^*) \geq \pi_i(p, p^*_{-i}, r^*)$; that is, firm $i$ cannot increase its revenue by posting another price $p \neq p^*_i$ when the other firm posts a price of $p^*_{-i}$ and the reference price is $r^*$.
2. **Stability Condition.** $r^* = \theta_1 p^*_1 + \theta_2 p^*_2$; that is, the reference price does not change if the firm $i \in \{1, 2\}$ keeps posting price $p^*_i$; see Equation (2).

Throughout the paper, we may refer to a point $(p^*, r^*)$ that satisfies the aforementioned conditions as a Stable Nash Equilibrium (SNE).

**Firms’ Information Structure.** We present pricing policies under a partial information setting. In this setting, a firm $i$ does not know $d_i$, $d_{-i}$, reference price update dynamics, and does not observe any of historical competing prices nor the current reference price. To be more specific, in

---

5 The dependency of demand on price surcharges and discounts are of the same order $\gamma_i$, which corresponds to so-called risk-neutral consumers. Related literature have also studied asymmetric demand dependencies on surcharges and discounts; see Popescu and Wu (2007), Nasiry and Popescu (2011), Hu et al. (2016).
this setting, firms do not know the specific form of the demand functions and reference update dynamics, which in our case are linear. Nevertheless, we assume that after firms post prices $p_t$ under reference price $r_t$, they can access a first-order oracle that outputs $\partial \pi_i(p_t, r_t)/\partial p_i$, which intuitively represents consumers’ responsiveness to a firm’s prices under current market conditions.\(^6\) We note that the partial information setting models real-world opaque environments where firms do not possess information of the market or its competitors. In this setting, firms set prices simultaneously, so a firm does not observe competitor’s prices in the current period before setting its own price.

3. Existence and Structural Properties of SNE

In this section, we show that an SNE exists. Recall that for any SNE, each firm best responds to its competitor as well as consumers’ reference price with no incentive for unilateral deviation. Let $\psi_i(p_{-i}, r) = \arg \max_{p \in P} \pi_i(p, p_{-i}, r), i \in \{1, 2\}$ be firm $i$’s best-response to the reference price $r$ and the price of the other firm $p_{-i}$. Further, for any reference price $r$, define set $B(r)$ as follows

$$B(r) = \{p : p_i = \psi_i(p_{-i}, r), i = 1, 2\}.$$  

(4)

As we will show in Theorem 3.1 below, $B(r)$ is non-empty and when it is not a singleton, it is an ordered set with total ordering.\(^8\) To show the existence of an SNE, we consider a simple pricing strategy that works as follows: in each period, firms set the largest best response profiles $p_t$ w.r.t. reference price $r_t$, i.e., $p_t = \max B(r_t)$ (because $B(\cdot)$ is an ordered set, $\max B(r_t)$ is well-defined). We show that for any initial reference price $r_1 \in P$, $(p_t, r_t)$ converges monotonically to an SNE. Of course, this pricing strategy is only possible under the complete information setting, where each firm knows its own demand function $d_i$, its competitor’s demand function $d_{-i}$, and the current reference

\(^6\) We note that such information can be obtained by a slight perturbation of the posted price. Furthermore, the assumption of having access to the first-order oracle is very common in the literature; see, for example, a comprehensive introduction to convex optimization in Nesterov (2013).

\(^7\) Here, the revenue function $\pi_i$ is quadratic, so $\arg \max_{p \in P} \pi_i$ is a singleton.

\(^8\) A set $A \subset \mathbb{R}^d$ is an ordered set with total ordering if for any $x, y \in A$, either $x \leq y$ or $y \leq x$ where the relationship $\leq$ and $\geq$ between two vectors is component-wise.
price. That is, the described pricing strategy cannot be implemented in our partial information setting. Nevertheless, the convergence under this policy confirms the existence of an SNE.

**Theorem 3.1 (Existence of an SNE)** Let $\mathcal{B}(r)$, defined in Equation (4), be the set of best-response profiles w.r.t. reference price $r$. Then, for a fixed reference price $r \in \mathcal{P}$, $\mathcal{B}(r)$ is non-empty, and when $\mathcal{B}(r)$ is not a singleton, it is an ordered set with total ordering. Furthermore, assume that in each period $t$, firms set the largest best response prices $p_t$ w.r.t. reference price $r_t$, i.e., $p_t = \max \mathcal{B}(r_t)$. Then, for any initial reference price $r_1 \in \mathcal{P}$, $(p_t, r_t)$ converges monotonically to an SNE.

The proof regarding the structural properties of the set of best response profiles $\mathcal{B}(r)$ is inspired by that of Tarski’s fixed point theorem (e.g., see Echenique et al. (2005)). The proof of the second half regarding the convergence of the pricing policy builds on that of Theorem 6 in Milgrom and Roberts (1990) (which shows monotonocity of pure-strategy Nash Equilibrium for parameterized games). Detailed proofs can be found in Section 6.1. Theorem 3.1 illustrates structural properties of SNEs: since $\mathcal{B}($·$)$ is an ordered set with total ordering, if there are multiple SNE’s, any two SNE’s $(p^*_a, r^*_a)$ and $(p^*_b, r^*_b)$ must either satisfy $p^*_a \geq p^*_b$ or $p^*_a \leq p^*_b$ under component-wise comparisons.

Due to the decision set boundaries, there may exist multiple SNE’s. To simplify our analyses, in the rest of the paper we assume there exists an SNE in the interior of the action set $\mathcal{P}$. Under this assumption, Lemma 3.2 (proof provided in Appendix A) shows that the interior SNE is unique.

**Assumption 1** There exists an SNE $(p^*, r^*)$ such that $(p^*, r^*) \in (\bar{p}, \bar{p})^3$.

**Lemma 3.2 (Uniqueness of SNE)** Under Assumption 1, there is a unique SNE $(p^*, r^*) \in (\bar{p}, \bar{p})^3$.

4. **No-regret Pricing Policies under Partial Information Setting**

Recall that under partial information, firms are unaware of the consumer demand function (they do not know the demand function is linear), reference prices, and reference price update dynamics.

---

9 This is a very mild assumption because in practice firms can enlarge the set of feasible prices to ensure an SNE lies in the interior.
Hence, a natural approach for firms to increase revenue is to employ so-called no-regret online learning algorithms that adjusts prices in a dynamic fashion. We study the regime in which firms adopt the general OMD algorithm. We start by the following standard definition.

**Definition 4.1 (Strong convexity)** Let $C \subset \mathbb{R}$ be a convex set. A function $R : C \to \mathbb{R}$ is said to be $\sigma$-strongly convex if for any $x, y \in C$, we have $R(x) - R(y) \geq \frac{dR(y)}{dy}(x - y) + \frac{\sigma^2}{2}(y - x)^2$.

In the OMD algorithm, each firm $i$ chooses a continuously differentiable and strongly convex regularizer $R_i : \mathbb{R} \to \mathbb{R}$ associated with strong-convexity parameter $\sigma_i$, a sequence of step sizes $\{\epsilon_{i,t}\}_t$, and, for our convenience, minimizes the cost function (i.e. inverse of revenue) $\tilde{\pi}_i := -\pi_i$, which is convex in $p_i$. Furthermore, each firm $i$ maintains a proxy variable $y_{i,t} \in \mathbb{R}$ over time, and in each period $t$, conducts pricing according to the following three steps:

1. Project the proxy variable $y_{i,t}$ back to the decision interval $\mathcal{P} = [\underline{p}, \bar{p}]$: $p_{i,t} = \Pi_{\mathcal{P}}(y_{i,t})$, where

   $\Pi_{\mathcal{P}} : \mathbb{R} \to \mathcal{P}$ is the projection operator such that $\Pi_{\mathcal{P}}(z) = z\{z \in \mathcal{P}\} + \underline{p}\{z < \underline{p}\} + \bar{p}\{z > \bar{p}\}$.

2. Access the first-order oracle $g_{i,t} := g_i(p_t, r_t)$ defined by $g_i : \mathcal{P}^3 \to \mathbb{R}$, where $g_i(p, r) = \partial \tilde{\pi}_i(p, r) / \partial p_i = 2\beta_i p_i - (\alpha_i + \delta_i p_i - \gamma_i r)$. We note that first-order feedback is very common in the optimization and learning literature as discussed in Section 2. Here, we point out that after a firm posts prices, it only obtains $g_{i,t}$ and does not necessarily observe the prices of its competitor nor the reference price.\(^\text{10}\)

3. Update proxy variable $y_{i,t+1}$ such that $R'_i(y_{i,t+1}) = R'_i(p_{i,t}) - \epsilon_{i,t} g_{i,t}$\(^\text{11}\) where $R'_i(q) := \frac{dR_i(y)}{dy} \bigg|_{y=q}$.

\(^{10}\) Firms do not know the linear form of demand, and hence cannot learn parameters and then best respond given parameter estimates.

\(^{11}\) $y_{i,t+1}$ exists when $R_i$ is continuously differentiable and convex, see Section 3.3 of Boyd et al. (2004) or Section 5.2 of Bubeck (2011).
We summarize the two-firm OMD pricing scheme in Algorithm 1.

**Algorithm 1** 2-firm OMD pricing under reference price updates

**Input:** \{\(R_i, \{\epsilon_{i,t}\}_t\), \(y_{i,1} = \arg\min_{y \in P} R_i(y)\) for \(i = 1, 2\).

1: for \(t = 1, 2, \ldots\) do
2: \hspace{1em} for \(i = 1, 2\) do
3: \hspace{2em} Set price: \(p_{i,t} = \Pi_P(y_{i,t})\).
4: \hspace{2em} Access gradient \(g_{i,t} = g_i(p_t, r_t)\).
5: \hspace{2em} Update proxy variable: \(R'_{i}(y_{i,t+1}) = R'(p_{i,t}) - \epsilon_{i,t}g_{i,t}\).
6: \hspace{1em} end for
7: end for
8: Reference price update (unobservable): \(r_{t+1} = ar_t + (1-a)(\theta_1p_{1,t} + \theta_2p_{2,t})\)

One can think of this sequential price competition with reference prices as a state-based dynamic game model where the reference price plays the role of an underlying state: each player (i.e., firm) has a continuous action space \(P\) and payoff function \(\widetilde{\pi}_i\) that depends on all players’ actions as well as an underlying state variable \(r_t\) that undergoes deterministic transitions. However, the view that we will adopt in the rest of the paper perceives reference prices \(r_t\) as price decisions \(p_{n,t} = r_t\) posted by a virtual firm which we refer to as nature and denote it by \(n\). This is possible if, for any \(\widetilde{\pi}_i, R_i, \{\epsilon_{i,t}\}_t\) \((i = 1, 2)\), we are able to construct a universal nature cost function \(\widetilde{\pi}_n(p_1, p_2, p_n)\), strongly convex regularizer \(R_n : \mathbb{R} \to \mathbb{R}\), and step size sequence \(\{\epsilon_{n,t}\}_t\), such that when firms 1, 2 and nature independently run the OMD algorithm with their respective regularizers and step sizes (as summarized in Algorithm 2), the resulting price profiles \(\{p_{1,t}, p_{2,t}, p_{n,t}\}_t\) recover the respective prices \(\{p_t, r_t\}_t\) of Algorithm 1. Here, note that \(g_{n,t} = g_n(p_{1,t}, p_{2,t}, p_{n,t}) = \partial \widetilde{\pi}_n(p_{1,t}, p_{2,t}, p_{n,t})/\partial p_{n,t}\).

The following Proposition 4.1 formalizes this view and shows that such \(\widetilde{\pi}_n, R_n,\) and \(\epsilon_{n,t}\) indeed exist. The proof of this lemma is provided in Appendix B, and we will refer to the dynamic game characterized in Algorithm 2 as the *induced 3-firm dynamic game*. 

**Algorithm 2** Induced 3-firm OMD pricing with no reference price

**Input:** \{\(R_i, \{\epsilon_{i,t}\}_t\) i=1,2,\(n\), \(y_{n,1} = r_1\), \(y_{i,1} = \arg\min_{y \in P} R_i(y)\) for \(i = 1, 2\).

1: for \(t = 1, 2, \ldots\) do
2: \hspace{1em} for \(i = 1, 2, n\) do
3: \hspace{2em} Set price: \(p_{i,t} = \Pi_P(y_{i,t})\).
4: \hspace{2em} Access gradient \(g_{i,t} = g_i(p_t, r_t)\).
5: \hspace{2em} Update proxy variable: \(R'_{i}(y_{i,t+1}) = R'(p_{i,t}) - \epsilon_{i,t}g_{i,t}\).
6: \hspace{1em} end for
7: end for
Proposition 4.1 (Induced 3-firm dynamic game) Fix any $\tilde{\pi}_i, R_i, \{\epsilon_{i,t}\}_t, i = 1, 2$, and initial reference price $r_1$. If nature (called firm $n$) is associated with cost function $\tilde{\pi}_n(p, r) = \frac{1}{2}r^2 - (\theta_1 p_1 + \theta_2 p_2) r$, and chooses regularizer $R_n(r) = \frac{1}{2} r^2$ and step size $\epsilon_{n,t} = 1 - a$, for any $t \geq 1$, then the price profiles $\{p_{1,t}, p_{2,t}, p_{n,t}\}_{t \geq 1}$ resulting from the game in Algorithm 2 recovers the induced price and reference price trajectory $\{p_t, r_t\}_{t \geq 1}$ of Algorithm 1.

We note that the choices for nature’s cost function $\tilde{\pi}_n$, regularizer $R_n$ and step sizes $\{\epsilon_{n,t}\}_t$ may not be unique, and in Proposition 4.1, we simply choose the most straightforward feasible candidate. Nonetheless, by this lemma, the nature takes constant step sizes $1 - a$, which implies that we have an inflexible (virtual) firm whose learning rate is always very fast. The existence of such a firm prohibits the direct application of methodologies introduced in multi-agent online learning to show convergence results as they require decreasing step sizes (e.g. Mertikopoulos and Zhou (2019)); see Section 5 for further discussions. By viewing reference prices as prices posted by nature, the induced 3-firm game is also associated with the static game that involves 3 players $i = 1, 2, n$ with respective payoffs $\{\tilde{\pi}_i\}_{i=1,2,n}$ and common action set $\mathcal{P}$. It turns out that the pure strategy Nash Equilibrium (PSNE) of this static game is unique and is identical to the SNE of Lemma 3.2.

Proposition 4.2 (PSNE of induced 3-firm static game) Consider the static game with players $i = 1, 2$ and nature $n$, who aims to minimize respective costs $\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_n$ with identical action set $\mathcal{P} = [\underline{p}, \bar{p}]$. Under Assumption 7, this game admits a unique PSNE $(p^*, r^*)$, i.e., $\tilde{\pi}_i(p_i^*, p_{-i}^*) \geq \tilde{\pi}_i(p_i, p_{-i}^*)$ for $\forall p_i \in \mathcal{P}$ and $i = 1, 2, n$. Furthermore, this PSNE is identical to the interior SNE of Lemma 3.2.

5. Convergence Results

The key challenge in showing convergence for the induced 3-firm OMD game play in Algorithm 2 lies in the fact that the step size sequence for nature is the constant $1 - a$, unlike previously studied multi-agent learning settings where step size sequences for every agent is typically decreasing in the time period $t$ (see for example Bravo et al. (2018), Tampubolon and Boche (2019), Mertikopoulos and Zhou (2019)). This highlights the fundamental issue in our problem of interest: will convergence
still occur if one of the players takes a constant (fixed) step size? In Section 5.1 we show that prices and reference prices converge to the unique interior SNE when the two firms adopt decreasing step sizes and characterize the corresponding convergence rate. In Section 5.2 we show that there exist constant step sizes for the two firms with which prices convergence to the SNE at faster rates compared to decreasing step sizes.

5.1. Decreasing Step Sizes

The first key result in this section is the following theorem, which states that if the two firms run the OMD algorithm with decreasing step sizes that do not go to zero too fast, then convergence to the SNE is guaranteed. We highlight that the firms’ step sizes are not required to be the same.

**Theorem 5.1 (Convergence under Decreasing Step Sizes)** If firm $i = 1, 2$ adopts regularizer $R_i$ that is $\sigma_i$-strongly convex and continuously differentiable, then $\{p_t, r_t\}_t$ converges when sequence of $\{\epsilon_{i,t}\}_t$ is nonincreasing with $\lim_{t \to \infty} \epsilon_{i,t} = 0$. Furthermore, if $\lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t} = \infty$ and $\lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t}^2 < \infty$, then $\{p_t, r_t\}_t$ converges to the unique interior SNE $(p^*, r^*)$.

The first part of Theorem 5.1 shows that prices converge when the firms’ step sizes go to zero eventually. This is an interesting result because in the induced 3-firm dynamic game presented in Algorithm 2, nature adopts a constant step size and learns quickly, while the two other firms are learning slowly through decreasing step sizes. In other words, convergence occurs even when nature reacts too fast to changes via adopting a constant step size. However, at the convergence point, firms may have the incentive to deviate, leading to an unstable market. The second part of the theorem addresses this concern and shows under sufficient step size conditions $\lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t} = \infty$ and $\lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t}^2 < \infty$, the market stabilizes as prices converge to the SNE. In fact, these conditions admit a large range of step sizes, e.g. $\epsilon_{i,t} = \Theta(1/t^\eta)$ for $\eta \in (\frac{1}{2}, 1]$. The proof of this theorem is provided in Section 6.2. Here, we provide some examples to solidify the aforementioned ideas.

An example is the extreme case where firms 1 and 2 adopt step sizes $\epsilon_{i,t} = 0$. This obviously guaranties convergence because prices are fixed at the initial prices, which are likely not the SNE, encouraging the firms to unilaterally deviate.
Example 1 (Decreasing Step Sizes) Consider the following demand and reference update model parameters: $\alpha = (5, 6)$, $\beta = (2, 3)$, $\delta = (0.4, 0.7)$, $\gamma = (0.1, 0.5)$, $\theta_1 = 0.8$, $a = 0.4$, $P = [1, 2]$, and initial prices $(p_1, r_1) = (1, 1, 1.5)$. These parameters admit the unique SNE given by $(p^*, r^*) = (1.41, 1.28, 1.39)$. We assume both firms adopt regularizer $R(x) = \frac{1}{2} x^2$, and consider two different decreasing step size sequences:

- With $\epsilon_{i,t} = 0.1/t^2$, the price profile eventually converges to the point $(\tilde{p}, \tilde{r}) = (1.21, 1.18, 1.20)$ which is not the SNE (see Figure 1a) and firms are incentivized to deviate, e.g., the best response for firm 1 w.r.t. $\tilde{p}_2 = 1.18$ and $\tilde{r} = 1.20$ is $1.40 \neq \tilde{p}_1$. Hence, under this step size sequence, firms may go through different epochs in the long run, in which firms converge in an epoch, and may decide to deviate and start over.

- With $\epsilon_{i,t} = 1/t$, we have $\lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t} = \infty$ and $\lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t}^2 < \infty$. Thus, per Theorem 5.1, prices and reference prices converge to the unique SNE; see Figure 1b. Moreover, we observe that (i) convergence occurs very quickly (for $t \geq 20$), and (ii) prices do not converge monotonically. The latter is in contrast with the pricing policy presented in Theorem 3.1.

In Example 1 we observe fast convergence to the SNE when firms choose decreasing step sizes. Inspired by this, we also characterize convergence rates for such step sizes:

**Theorem 5.2 (Convergence Rate under Decreasing Step Sizes)** For any sensitivity margin $m \geq 2$, if both firms adopt regularizer $R_i(z) = z^2$, there exists step sizes $\epsilon_{i,t} = \Theta(1/t)$ and an absolute constant $c$, which depends on $a$ and $\max\{\theta_1, \theta_2\}$, such that $\|p^* - p_t\|^2 \leq c/t$ for any $t \in \mathbb{N}^+$. The proof of this theorem constructs a sufficiently large absolute constant $c$ and shows $\|p^* - p_t\|^2 \leq c/t$ via induction. The main procedure involves bounding $\|p^* - p_{t+1}\|^2$ with $\|p^* - p_t\|^2$ and $|r_t - r^*|$, and developing a tight bound for $\sum_{t=1}^{T-1} \|p^* - p_t\|^2$. Bounding $\sum_{t=1}^{T-1} \|p^* - p_t\|^2$ helps us bound $|r_t - r^*|$ because the deviations of prices w.r.t. the interior SNE will cumulatively propagate into $|r_t - r^*|$ due

---

Here, we choose $\epsilon_{i,t} = 0.1/t^2$ because the gap between the convergence point and the SNE is more visible. For the more natural choice $\epsilon_{i,t} = 1/t^2$, we obtain similar results.
to reference price update dynamics. The detailed proof is provided in Section 6.3. We also remark that the condition \( m \geq 2 \) is a rather practical regime because this condition, as discussed in Section 2, implies a firm’s demand is more sensitive to its own prices compared to competitor’s prices and surcharge (or discounts) relative to reference prices. Finally, we remark that the constant \( c \) scales reasonably w.r.t. \( a \) and \( \max\{\theta_1, \theta_2\} \) as long as they are bounded away from 1; see Figure 2b in Appendix C.4 for an illustration for \( a \in [0.1, 0.9] \) and \( \max\{\theta_1, \theta_2\} \in \{0.5, 0.6, \ldots, 0.9\} \).

![Figure 1](image)

**Figure 1** Illustration of price and reference price trajectories in Examples 1 and 2 under different step size sequences. The y-axis represents price levels, as the x-axis denotes time.

5.2. Constant Step Sizes

We start by revisiting Example 1 and adopt constant step sizes.

**Example 2 (Constant Step Sizes)** Consider the same demand and reference update model parameters in Example 1.

- With \( \epsilon_{i,t} = 1 - a \), Figure 1c shows price profiles do not converge and oscillate in the long-run.
- With \( \epsilon_{i,t} = (1 - a)/\beta_i \), Figure 1d shows price profiles converge to the SNE at a faster rate compared to decreasing step sizes in Figure 1b.

Given this example, the first main result for this section is the following theorem (see proof in Appendix C) which shows that under some conditions, there exists constant step size proportional to \( \frac{1-a}{\beta_i} \) under which pricing profiles and reference price convergence to the unique interior SNE.
Theorem 5.3 (Sufficient Conditions for Convergence under Constant Step Sizes)

Suppose firm $i$ adopts regularizer $R_i$ that is $\sigma_i$-strongly convex and continuously differentiable. For strong-convexity parameters $\sigma_1, \sigma_2$ and sensitivity margin $m$, let $S_{i,m} := \{ z > 0 : f_{i,m}(z) < 0 \}$, where

$$f_{i,m}(z) = \begin{cases} 
4\sigma_i \left( \frac{2\sigma_i}{m^2} \right) z^2 - \left( (2 - \frac{1}{2m}) \sigma_i - \frac{\sigma_i}{2m} \right) z + \frac{3}{4} & i = 1, 2 \\
\frac{2}{m^2} (\sigma_1 + \sigma_2) z^2 + \frac{1}{2m} (\sigma_1 + \sigma_2) z - \frac{1}{4} & i = n
\end{cases}.$$  \hspace{1cm} (5)

Then, if $\cap_{i=1,2,n} S_{i,m} \neq \emptyset$, the step size sequence $\epsilon_{i,t} = s \sigma_i \left( \frac{1-a}{\beta_i} \right)$ for any $s \in \cap_{i=1,2,n} S_{i,m}$ guarantees $\{p_t, r_t\}_t$ converges to the unique interior SNE $(p^*, r^*)$.

This theorem indicates that under some conditions on $m, \sigma_1,$ and $\sigma_2$, there exist constant step sizes with which convergence to the unique interior SNE is guaranteed. The desired step size is proportional to $\frac{1-a}{\beta_i}$. This, roughly speaking, implies that prices converge to the SNE if firms adjust prices at a pace similar to that of nature. Recall that $1-a$ can be considered as the step size of nature, and by demand model in Equation (1), $\beta_i$ is firm $i$’s price sensitivity parameter. The conditions on $m, \sigma_1,$ and $\sigma_2$ in Theorem 5.3 are, in fact, quite mild: the following Corollary 5.3.1 provides an example where for any $m > 2$, we can find sufficiently large $\sigma_1 = \sigma_2$ that guarantees convergence to an SNE.

Corollary 5.3.1 (Convergence under Constant Step Sizes) For any sensitivity margin $m > 2$, assume both firms adopt continuously differentiable regularizer $R_i$ that is $\sigma$-strongly convex where $\sigma > \sigma_0$ and $\sigma_0 := \max \left\{ \frac{6(2m^2+1)}{(2m-1)^2}, \frac{(2m^2+7)^2}{8m^4-36m+8} \right\}$. Then there exists constant $s$ dependent on $m$ and $\sigma$ so if firm $i \in \{1, 2\}$ adopts step size $\epsilon_{i,t} = s \sigma_i \left( \frac{1-a}{\beta_i} \right)$, $\{p_t, r_t\}_t$ converges to the unique interior SNE.

This corollary provides sufficient conditions for the existence of constant step sizes that guarantee convergence to the SNE for any sensitivity margin $m > 2$. In fact, for suitable $m$, we can possibly find relatively small $\sigma_1$ and $\sigma_2$ so that the conditions are satisfied (e.g., $\sigma_1 = \sigma_2 = 4$ for $m = 5$). Note that $\sigma_0 = \Theta(m)$ for large $m$, which means firms generally need to take larger strong-convexity parameters as $m$ increases. (See Figure 2a in Appendix C.4 for illustration of $\sigma_0$ as a function of $m$.) Having
everything else fixed, the larger $\sigma$, the slower prices move.\footnote{This is so because for large $m$, a firm’s demand is very sensitive to its own prices, encouraging the firm to adjust prices slowly via large $\sigma$.} Moreover, we also characterize the convergence rate when firms adopt suitable constant step sizes via the following Theorem 5.4 and highlight that such fast learning rates give us much faster convergence to the SNE, compared to slow learning rates from decreasing step sizes.

**Theorem 5.4 (Convergence Rate for Constant Step Sizes)** For any sensitivity margin $m > 2$, assume that both firms use quadratic regularizer $R_i(z) = \frac{\sigma z^2}{2}$ for any $\sigma > \sigma_0$, where $\sigma_0$ is defined in Corollary 5.3.1. Then, there exists constant $s > 0$, dependent on $m$ and $\sigma$, such that if firm $i = 1, 2$ adopts step size $\epsilon_{i,t} = \sigma (1 - a) \beta_i$ for $t \in \mathbb{N}^+$, we have $\|p^* - p_i\|^2 \leq \frac{1 + 2\sigma}{\sigma} \left( \bar{p} - p \right)^2 \left( 1 + a \right)^t$.

### 5.3. Comparison with Multi-agent Online Learning

Our proof techniques for Theorems 5.1, 5.2, 5.3, and 5.4 are not standard, and here we highlight key differences compared to the established results in multi-agent online learning for dynamic concave games (e.g. Bravo et al. (2018), Tampubolon and Boche (2019), Mertikopoulos and Zhou (2019)). These works on dynamic games rely on a global variational stability assumption on the PSNE for the corresponding static game. In our setting, the second-order test in Mertikopoulos and Zhou (2019) shows that the PSNE in our induced 3-firm static game (see Lemma 4.2) indeed satisfies the global variational stability condition. That is, $\langle g(p), p^* - p \rangle \leq 0$ for $\forall p \in \mathcal{P}^3$, where we slightly abuse the notation and write $p = (p_1, p_2, p_n)$, and $g(p) = (g_1(p), g_2(p), g_n(p))$. If one enforces $\epsilon_{i,t} = \epsilon_t$ for $i = 1, 2, n$, showing the convergence results in this line of work boils down to verifying:

\[
\sum_{i=1,2,n} D_i(p^*_i, p_{i,t+1}) \overset{(a)}{\leq} \sum_{i=1,2,n} D_i(p^*_i, p_{i,t}) + \epsilon_t \langle g(p_t), p^* - p_t \rangle + c_1^2 c_2 \overset{(b)}{<} \sum_{i=1,2,n} D_i(p^*_i, p_{i,t}),
\]

where $c_2$ can be viewed as some absolute constant, and $D_i$ is Bregman divergence w.r.t. strongly convex regularizer $R_i$ (see Definition 6.2 in Section 6 for definition of Bregman divergence). At a high level, this shows that the distance between $p^*_i$ and $p_{i,t}$ shrinks over time and hence implies convergence.
to the SNE. The inequality (a) follows from classical mirror descent proofs; and inequality (b) utilizes
the variational stability condition by choosing small enough $\epsilon_t$ (typically in the order of $\Theta(1/t)$).
However, this procedure will not be applicable in our setting as reference prices can be viewed as
prices posted by nature (see Proposition 4.1), which is an inflexible firm with constant step size
sequence $1 - a$. In other words, we cannot enforce $\epsilon_{i,t} = \epsilon_t$ to be decreasing in $t$ for $i = 1, 2, n$.

6. Proofs for Key Theorems

In this section, we will provide proofs for some of our main results.

Definition 6.1 (Best response mapping) We define the best-response mapping as
$\psi: P^3 \rightarrow P^2$ such that $\psi(p, r) = (\psi_1(p_2, r), \psi_2(p_1, r))$. Then, we can rewrite the set of best-response profiles w.r.t.
reference price $r$, defined in Equation (4), as $B(r) = \{p \in P^2 : p = \psi(p, r)\}$. Note that for any SNE $(p^*, r^*)$, we must have $p^* \in B(r^*)$, and $p^*$ is a fixed point of the mapping $\psi(\cdot, r^*)$.

Definition 6.2 (Bregman Divergence) The Bregman divergence $D: C \times C \rightarrow \mathbb{R}^+$ associated with
convex set $C \subset \mathbb{R}$, and convex and continuously differentiable function $R: C \rightarrow \mathbb{R}$ is defined as
$D(x, y) = R(x) - R(y) - R'(y)(x - y) \geq 0$, where the inequality follows from convexity of $R$. Furthermore, if $R$ is $\sigma$-strongly convex, then $D(x, y) \geq \frac{\sigma^2}{2}(x - y)^2$.

We denote $D_i$ as the Bregman divergence associated with regularizer $R_i$ used by firm $i = 1, 2,$ and
$D_n$ as the Bregman divergence associated with regularizer $R_n$ used by nature.

Definition 6.3 Let $g^*_i$ be the partial derivative of the cost function $\tilde{\pi}_i$ w.r.t. $p_i$ evaluated at the
interior SNE $(p^*, r^*)$, i.e. for $i = 1, 2, n$, $g^*_i = \frac{\partial \tilde{\pi}_i(p_1, p_2, p_n)}{\partial p_i} \bigg|_{p_1=p_1^*, p_2=p_2^*, p_n=r^*}$.

6.1. Proof of Theorem 3.1

(i) By first order conditions, we know that $\arg \max_{p \in \mathbb{R}} \pi_i(p, p_{-i}, r) = \frac{\alpha_i + \delta_i p_{-i} + \gamma_i r}{2\beta_i}$. Hence, due to
boundary constraints on the decision set $P$ and the revenue function being quadratic, we have

$$\psi_i(p_{-i}, r) = \arg \max_{p \in P} \pi_i(p, p_{-i}, r) = \Pi_P \left( \frac{\alpha_i + \delta_i p_{-i} + \gamma_i r}{2\beta_i} \right)$$.
where $\Pi_p : \mathbb{R} \to \mathcal{P}$ is the projection operator such that $\Pi_p(z) = z \mathbb{I}\{ z \in \mathcal{P} \} + \frac{1}{2} \mathbb{I}\{ z \in \mathcal{P} \} + \frac{1}{2} \mathbb{I}\{ z > \bar{p} \}$. Hence, $\psi_t(p_{-i}, r)$ is a nondecreasing function in $p_{-i}$ and $r$, which further implies $\psi(p, r)$ is nondecreasing in $p$ and $r$. Again, recall for any $x, y$, the relationships $x \leq y$ and $y \leq x$ are component-wise comparisons.

We now follow a similar proof to that of Tarski’s fixed point theorem: consider the set $\mathcal{B}_+(r) = \{ p \in \mathcal{P}^2 : p \leq \psi(p, r) \}$. It is apparent that this set is nonempty because $(\bar{p}, \bar{p}) \in \mathcal{B}_+(r)$. Fix any $p \in \mathcal{B}_+(r)$. Then, we have $p \leq \psi(p, r)$ which further implies $\psi(p, r) \leq \psi'(p, r)$ since $\psi(p, r)$ is nondecreasing in $p$. Hence $\psi(p, r) \in \mathcal{B}_+(r)$. By taking $U(r) = \sup \mathcal{B}_+(r)$ (this is possible since all $p \in \mathcal{B}_+(r)$ are bounded), we have $p \leq U(r)$ so $p \leq \psi(p, r) \leq \psi(U(r), r)$. This further implies $U(r) \leq \psi(U(r), r)$ because $U(r)$ is the least upper bound of $\mathcal{B}_+(r)$, and thus $U(r) \in \mathcal{B}_+(r)$. This allows us to conclude $\psi(U(r), r) \leq U(r)$ and hence $U(r) = \psi(U(r), r)$, which means $U(r) = \sup \mathcal{B}_+(r)$ is a fixed point of the mapping $\psi(\cdot, r)$. Thus, $U(r)$ belongs in the set of best-response profiles $\mathcal{B}(r)$, confirming $\mathcal{B}(r)$ is not empty.

Next, we show that $\mathcal{B}(r)$ is an ordered set with total ordering if it is not a singleton. To do so, consider any $p, q \in \mathcal{B}(r)$ and without loss of generality assume $p_1 > q_1$. Since $p_1 = \psi_1(p_2, r)$ and $q_1 = \psi_1(q_2, r)$, by monotonicity of $\psi_1(\cdot, r)$ we have $p_2 > q_2$. Thus, $p > q$ and $\mathcal{B}(r)$ is an ordered set with total ordering.

(ii) In the proof of (i), we showed that $U(r) = \sup \{ p \in \mathcal{P}^2 : p \leq \psi(p, r) \}$ is a fixed point of the best-response mapping $\psi(\cdot, r)$ for any $r$ which allows us to conclude $U(r)$ is the largest best-response profile, i.e., $U(r) = \max \mathcal{B}(r)$, and hence $p_t = U(r_t)$. Furthermore, since $\psi(p, r)$ is increasing in $r$, we know that $U(\cdot) = \sup \{ p \in \mathcal{P}^2 : p \leq \psi(p, \cdot) \}$ is also an increasing function. In the following, we will argue that the reference prices $r_t$ is monotonically increasing or decreasing, which implies $p_t = U(r_t)$ is also monotonic, and hence converges since prices and reference prices are bounded.

We write $U(r) = (U_1(r), U_2(r))$. At $t = 1$, if $\theta_1 p_{1,1} + \theta_2 p_{2,1} = \theta_1 U_1(r_1) + \theta_2 U_2(r_1) \geq r_1$, then the reference price at $t = 2$ satisfies $r_2 = a r_1 + (1 - a) \theta_1 p_{1,1} + \theta_2 p_{2,1} \geq r_1$. By the monotonicity of $U(\cdot)$, we have $p_{i,2} = U_i(r_2) \geq U_i(r_1) = p_{i,1}$ for $i = 1, 2$. Thus,

$$r_3 = a r_2 + (1 - a) \theta_1 p_{1,2} + \theta_2 p_{2,2} \geq a r_1 + (1 - a) \theta_1 p_{1,1} + \theta_2 p_{2,1} = r_2.$$
A simple induction argument thus shows \( \{r_t\} \) is a nondecreasing sequence. Since \( r_t \leq \bar{p} \) for any \( t \in \mathbb{N} \), we know that \( \{r_t\} \) converges to some number \( r_+ \in [\bar{p}, \bar{p}] \) when \( \theta_1 p_{1,1} + \theta_2 p_{2,1} \geq r_1 \). Furthermore, we observe that \( \lim_{t \to \infty} \psi(U(r_t), r_t) = \psi(U(r_+), r_+) \) by the definition of \( \psi \). Also, from (i) we have \( \psi(U(r_t), r_t) = U(r_t) \) and \( \psi(U(r_+), r_+) = U(r_+) \) because \( U(r) \) is a fixed point of \( \psi(\cdot, r) \) for any \( r \). Hence, \( \lim_{t \to \infty} U(r_t) = U(r_+) \), which implies \( \{p_t = U(r_t)\} \) converges to \( U(r_+) \). Note that convergence is monotonic because \( U(\cdot) \) is nondecreasing. Therefore,

\[
\theta_1 U_1(r_+) + \theta_2 U_2(r_+) = \lim_{t \to \infty} \theta_1 U_1(r_t) + \theta_2 U_2(r_t) = \lim_{t \to \infty} r_{t+1} = r_+,
\]

which implies \((U(r_+), r_+)\) is an SNE. We can thus conclude that if \( \theta_1 p_{1,1} + \theta_2 p_{2,1} = \theta_1 U_1(r_1) + \theta_2 U_2(r_1) \geq r_1 \), firms’ prices and reference prices converge monotonically to an SNE \((U(r_+), r_+)\).

Following a symmetric argument, if \( \theta_1 p_{1,1} + \theta_2 p_{2,1} < r_1 \), we can show that \( \{r_t\} \) is a nonincreasing sequence. Since \( r_t \geq p \) for any \( t \in \mathbb{N} \), we know that \( \{r_t\} \) converges to some number \( r_- \in [\bar{p}, \bar{p}] \). Similar to the previous arguments, we can conclude that prices and reference prices converge monotonically to an SNE \((U(r_-), r_-)\). □

### 6.2. Proof for Theorem 5.1

The proof of this theorem is divided into two parts. In the first part, we show that the price profiles \((p_t, r_t)\) converge as \( t \to \infty \) under the condition \( \lim_{t \to \infty} \epsilon_{i,t} = 0, i \in \{1, 2\} \). In the second part, under the additional conditions \( \lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t} = \infty \) and \( \lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t}^2 < \infty \), we show that the price profiles converge to the unique interior SNE.

**First part: Convergence of prices and reference prices.** Recall that \( g_{i,t} = g_i(p_{i,t}, r_t) = 2 \beta_i p_{i,t} - (\alpha_i + \delta_i p_{-i,t} + \gamma_i r_t) \), and \( p_{i,t} \) and \( p_{-i,t} \) are both bounded. Hence, because for \( i = 1, 2 \), \( \lim_{t \to \infty} \epsilon_{i,t} = 0 \) and \( \{\epsilon_{i,t}\} \) is nonincreasing, we have for any small \( \epsilon > 0 \) there exist \( t_\epsilon \in \mathbb{N} \) such that

\[
|\epsilon_{i,t} g_{i,t}| \leq \frac{\epsilon^2}{6} \text{ for all } t > t_\epsilon.
\]

Our goal is to show that for \( t \geq t_\epsilon \), \(|p_{i,t+1} - p_{i,t}| \) is small, and hence \( p \) converges. For \( t > t_\epsilon \),

\[
|p_{i,t+1} - p_{i,t}| \leq |p_{i,t+1} - y_{i,t+1}| + |y_{i,t+1} - p_{i,t}| \overset{(a)}{=} |p_{i,t+1} - y_{i,t+1}| + \frac{\epsilon}{6}.
\]

(6)
To see why inequality (a) holds recall that \( y_{i,t+1} \) is the proxy variable in Step 5 of Algorithm 1 such that \( R_i'(y_{i,t+1}) - R_i'(p_{i,t}) = \epsilon_{i,t} g_{i,t} \). Hence,

\[
\sigma_i^2 |y_{i,t+1} - p_{i,t}| \leq |R_i'(y_{i,t+1}) - R_i'(p_{i,t})| = |\epsilon_{i,t} g_{i,t}| \leq \frac{\sigma_i^2 \epsilon}{6}, \quad t > t_e, i = 1, 2,
\]

which implies that \( |y_{i,t+1} - p_{i,t}| \leq \frac{\epsilon}{6} \), as desired. Here, the first inequality holds because:

\[
\sigma_i^2 (y_{i,t+1} - p_{i,t})^2 \overset{(a)}{\leq} (R_i'(y_{i,t+1}) - R_i'(p_{i,t})) (y_{i,t+1} - p_{i,t}) \leq |R_i'(y_{i,t+1}) - R_i'(p_{i,t})| \cdot |y_{i,t+1} - p_{i,t}|,
\]

where (a) follows from summing up \( R_i(y_{i,t+1}) - R_i(p_{i,t}) \geq R_i'(p_{i,t})(y_{i,t+1} - p_{i,t}) + \frac{\sigma_i^2}{2} (y_{i,t+1} - p_{i,t})^2 \) and \( R_i(p_{i,t}) - R_i(y_{i,t+1}) \geq R_i'(y_{i,t+1})(p_{i,t} - y_{i,t+1}) + \frac{\sigma_i^2}{2} (y_{i,t+1} - p_{i,t})^2 \) due to strong convexity.

By Equation (6), for \( t > t_e \),

\[
|p_{i,t+1} - p_{i,t}| \leq |p_{i,t+1} - y_{i,t+1}| \left( \mathbb{I}\{y_{i,t+1} < p\} + \mathbb{I}\{y_{i,t+1} \in \mathcal{P}\} + \mathbb{I}\{y_{i,t+1} > \bar{p}\} \right) + \frac{\epsilon}{6}
\]

\[
= |p_{i,t+1} - y_{i,t+1}| \left( \mathbb{I}\{y_{i,t+1} < p\} + \mathbb{I}\{y_{i,t+1} > \bar{p}\} \right) + \frac{\epsilon}{6}, \quad (7)
\]

where the equality holds because under the event \( y_{i,t+1} \in \mathcal{P} \), no projection occurs and hence, \( y_{i,t+1} = p_{i,t+1} \). In the first of the proof, we bound the first two terms in the right hand side, i.e.,

\[
|p_{i,t+1} - y_{i,t+1}| \mathbb{I}\{y_{i,t+1} < p\} \quad \text{and} \quad |p_{i,t+1} - y_{i,t+1}| \mathbb{I}\{y_{i,t+1} > \bar{p}\}.
\]

To bound \( |p_{i,t+1} - y_{i,t+1}| \mathbb{I}\{y_{i,t+1} < p\} \), similar to Equation (6) we use \( |p_{i,t} - y_{i,t+1}| \leq \frac{\epsilon}{6} \) for \( t > t_e \) which implies \( p_{i,t} - y_{i,t+1} \leq \frac{\epsilon}{6} \). Thus,

\[
y_{i,t+1} \geq p_{i,t} - \frac{\epsilon}{6} \geq \bar{p} - \frac{\epsilon}{6}, \quad (8)
\]

where (a) holds because \( p_{i,t} \geq \bar{p} \) for any \( i, t \). On the other hand, under the event \( y_{i,t+1} < p \), projection occurs and therefore we have \( p_{i,t+1} = p \). This yields

\[
|p_{i,t+1} - y_{i,t+1}| \mathbb{I}\{y_{i,t+1} < p\} = (p - y_{i,t+1}) \mathbb{I}\{y_{i,t+1} < p\} \overset{(a)}{\leq} \left( p - p + \frac{\epsilon}{6} \right) = \frac{\epsilon}{6}, \quad (9)
\]

where (a) follows from Equation (8).

Using a similar argument as above to bound \( |p_{i,t+1} - y_{i,t+1}| \mathbb{I}\{y_{i,t+1} > \bar{p}\} \), we have \( |p_{i,t+1} - y_{i,t+1}| \mathbb{I}\{y_{i,t+1} > \bar{p}\} \leq \frac{\epsilon}{6} \) under the event \( y_{i,t+1} > \bar{p} \). Plugging these upper bounds back into Equation (6), we can show that for any \( \epsilon > 0 \) and \( t > t_e \)

\[
|p_{i,t+1} - p_{i,t}| \leq \frac{\epsilon}{6} + \frac{\epsilon}{6} + \frac{\epsilon}{6} = \frac{\epsilon}{2}, \quad i = 1, 2. \quad (10)
\]
This shows prices converge, i.e., \( p_t \) converges.

Finally, we use the convergence of \( p_t \) to show that reference prices also converge. When \( t > t_\varepsilon \),

\[
|r_{t+2} - r_{t+1}| \leq a |r_{t+1} - r_t| + (1 - a) \sum_{i=1,2} \theta_i |p_{i,t+1} - p_{i,t}| \leq a |r_{t+1} - r_t| + \frac{(1 - a)\varepsilon}{2},
\]

where (a) follows from the convergence of \( p_{i,t} \) in Equation (10). Rearranging terms, this yields

\[
|r_{t+2} - r_{t+1}| - \frac{\varepsilon}{2} \leq a \left(|r_{t+1} - r_t| - \frac{\varepsilon}{2}\right).
\]

Using a telescoping argument, we have

\[
|r_{t+2} - r_{t+1}| - \frac{\varepsilon}{2} \leq a^t \left(|r_2 - r_1| - \frac{\varepsilon}{2}\right) \leq a^t \left(\bar{p} - p - \frac{\varepsilon}{2}\right).
\]

Since \( a \in (0, 1) \), there exists \( \exists \epsilon_a \in \mathbb{N} \) such that \( |r_{t+2} - r_{t+1}| - \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} \) for all \( t > t_\epsilon \). Therefore, for any \( t > \max\{t_\epsilon, \epsilon_a\} \), \( |p_{i,t+1} - p_{i,t}| \leq \frac{\varepsilon}{2} \) and \( |r_{t+2} - r_{t+1}| \leq \varepsilon \). This implies \((p_t, r_t)\) converges as \( t \to \infty \).

**Second part: Convergence to the SNE.** Now we show, via a contradiction argument, that if the conditions \( \lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t} = \infty \) and \( \lim_{T \to \infty} \sum_{t=1}^{T} \epsilon_{i,t}^2 < \infty \) hold, then \((p_t, r_t)\) converges to the interior SNE \((p^*, r^*)\). Here, we first provide a roadmap of the proof:

We start by evoking Corollary D.1.1 for any \( i = 1, 2, n \) and \( t \in \mathbb{N^+} \), we get

\[
D_i(p^*_t, p_{i,t+1}) \leq D_i(p^*_i, p_{i,t}) - \epsilon_{i,t} (g^*_i - g_{i,t}) (p^*_t - p_{i,t}) + \frac{\epsilon_{i,t}^2}{2} g_{i,t}^2 / \sigma_i,
\]

where we recall \( g_{i,t}^* \) is the partial derivative of the cost function \( \bar{r}_i \) w.r.t. \( p_t \) evaluated at the interior SNE \((p^*, r^*)\). If contrary to our claim, prices for firm \( i \) converge to some point other than the SNE prices, we can lower bound \((g^*_i - g_{i,t}) (p^*_t - p_{i,t})\) with some positive constant. Then, by utilizing the aforementioned step size conditions, we show that, roughly speaking, the distance between \( p_{i,t} \) and \( p^*_i \), measured by \( D_i(p^*_t, p_{i,t}) \), decreases by a constant amount for each time period after large enough \( t \). Eventually, this tells us that the distance between \( p_{i,t} \) and \( p^*_i \) will become negative, leading to a contradiction regarding the Bregman Divergence being positive. We conduct our proof in 3 steps. In Step 1, we provide bounds for \((g^*_i - g_{i,t}) (p^*_t - p_{i,t})\) are relatively easy to show; in Step 2 we lower bound \((g^*_i - g_{i,t}) (p^*_t - p_{i,t})\) for large enough \( t \) if firm \( i \)'s prices converge to some point other than the SNE; and in Step 3, we show that the positive lower bound for \((g^*_i - g_{i,t}) (p^*_t - p_{i,t})\) will eventually lead to a contradiction regarding Bregman Divergence being positive.
To begin with, contrary to our claim, assume that \((p_\ell, r_\ell) \xrightarrow{t \to \infty} (\bar{p}, \bar{r}) \neq (p^*, r^*)\), and define \(\Delta_i = p_i^* - \bar{p}_i\) for \(i = 1, 2, n\). Fix \(\ell \in \{1, 2\}\) such that \(|\Delta_\ell| \geq |\Delta_{-\ell}| \geq 0\) that is, let \(\ell\) be the firm whose corresponding \(|\Delta_\ell|\) is most far away from 0. Since \((\bar{p}, \bar{r}) \neq (p^*, r^*)\), we know that \(|\Delta_\ell| > 0\) and hence \(\Delta_\ell \neq 0\). Note that it is possible that the other firm \(-\ell\) has \(\Delta_{-\ell} = 0\). Define \(\eta\) such that

\[
0 < \eta < \min \left\{ \left| \frac{(2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\gamma_\ell \theta_{-\ell} + \delta_\ell) \Delta_{-\ell}}{2\beta_\ell + \delta_\ell + \gamma_\ell} \right|, \left| \frac{2\beta_\ell - \gamma_\ell \theta_\ell}{2\beta_\ell + \delta_\ell + \gamma_\ell} |\Delta_\ell| \right| \right\} .
\]

(12)

First, note that such an \(\eta\) exists because \(|\Delta_\ell| > 0\) and \(|(2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\gamma_\ell \theta_{-\ell} + \delta_\ell) \Delta_{-\ell}| \geq 0\). Here, (a) is due to the fact that \(\beta_\ell \geq m(\delta_\ell + \gamma_\ell)\) with \(m \geq 2\), and \(\theta_\ell \in (0, 1)\), which implies \(2\beta_\ell - \gamma_\ell \theta_\ell > \gamma_\ell \theta_{-\ell} + \delta_\ell > 0\), and

\[
| (2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\gamma_\ell \theta_{-\ell} + \delta_\ell) \Delta_{-\ell} | > (2\beta_\ell - \gamma_\ell \theta_\ell) \cdot |\Delta_\ell| - (\gamma_\ell \theta_{-\ell} + \delta_\ell) \cdot |\Delta_{-\ell}| > 0,
\]

where in (b) we also used the fact that \(|\Delta_\ell| \geq |\Delta_{-\ell}| \geq 0\) and \(|\Delta_\ell| > 0\). Also, note that definition of \(\eta\) implies that we have \(0 < \eta < |\Delta_\ell|\) as for any \(i = 1, 2\), \(0 < \frac{2\beta_\ell - \gamma_\ell \theta_\ell}{2\beta_\ell + \delta_\ell + \gamma_\ell} < 1\).

Furthermore, because \((p_\ell, r_\ell)\) converges to \((\bar{p}, \bar{r})\), let \(t_\eta \in \mathbb{N}^+\) be the period such that \(|\bar{p}_i - p_{i,t}| < \eta\) for all \(t > t_\eta\) and \(i = 1, 2, n\). This allows us to characterize the distance between prices and the SNE for \(i = 1, 2\) and \(t > t_\eta\):

\[
p_i^* - p_{i,t} = p_i^* - \bar{p}_i + \bar{p}_i - p_{i,t} < \Delta_i + \eta \quad (13)
\]

\[
p_i^* - p_{i,t} = p_i^* - \bar{p}_i + \bar{p}_i - p_{i,t} > \Delta_i - \eta . \quad (14)
\]

**Step 1.** Here, we will show the following lower and upper bounds for \(g_\ell^* - g_\ell, t\) for any \(t > t_\eta\):

\[
g_\ell^* - g_\ell, t \leq (2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} + \eta (2\beta_\ell + \delta_\ell + \gamma_\ell) \quad (15)
\]

\[
g_\ell^* - g_\ell, t \geq (2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} - \eta (2\beta_\ell + \delta_\ell + \gamma_\ell) . \quad (16)
\]

- **Lower bounding** \(g_\ell^* - g_\ell, t\). Recall \(g_\ell(p_\ell, p_{-\ell}, r) = 2\beta_\ell p_\ell - (\alpha_\ell + \delta_\ell p_{-\ell} + \gamma_\ell r), g_\ell, t = g_\ell(p_\ell, r_\ell)\) and \(g_\ell^* = g_\ell(p^*, r^*)\). Then,

\[
g_\ell^* - g_\ell, t = 2\beta_\ell (p^*_\ell - p_\ell, t) - \delta_\ell (p^*_{-\ell} - p_{-\ell, t}) - \gamma_\ell (r^* - r_\ell) . \quad (17)
\]
Therefore, for $t > t_\eta$ we have
\[
g^*_t - g_{\ell,t} = 2\beta_\ell (p^*_\ell - \bar{p}_\ell + \bar{p}_\ell - p_{\ell,t}) - \delta_\ell (p^*_\ell - \bar{p}_\ell + \bar{p}_\ell - p_{\ell,-}) - \gamma_\ell (r^* - \bar{r} + \bar{r} - r_\ell) \\
\geq 2\beta_\ell (\Delta_\ell - \eta) - \delta_\ell (\Delta_{-\ell} + \eta) - \gamma_\ell (\Delta_n + \eta) \\
equiv (b) (2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} - \eta (2\beta_\ell + \delta_\ell + \gamma_\ell) ,
\]
where (a) follows from the fact that $|\bar{p}_i - p_{i,t}| \leq \eta$ for any $t > t_\eta$ and $i = 1, 2$, and (b) follows from the relationship $\Delta_n = \theta_1 \Delta_1 + \theta_2 \Delta_2$, which holds due to the following:
\[
\bar{r} = \lim_{t \to \infty} r_{t+1} = \lim_{t \to \infty} ar_t + (1 - a) (\theta_1 p_{1,t} + \theta_2 p_{2,t}) = a\bar{r} + (1 - a) (\theta_1 \bar{p}_1 + \theta_2 \bar{p}_2) .
\]
Combining this with $r^* = (\theta_1 p^*_1 + \theta_2 p^*_2)$ (due to the definition of the SNE), and the definition of $\Delta_i = p^*_i - \bar{p}_i$ for $i = 1, 2, n$ yields the desired lower bound.

- **Upper bounding $g^*_t - g_{\ell,t}$**. Similar to the lower above, for $t > t_\eta$ we have
\[
g^*_t - g_{\ell,t} \leq 2\beta_\ell (\Delta_\ell + \eta) - \delta_\ell (\Delta_{-\ell} + \eta) - \gamma_\ell (\Delta_n + \eta) \\

= (2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} + \eta (2\beta_\ell + \delta_\ell + \gamma_\ell) .
\]

**Step 2.** If $\Delta_\ell > 0$, then since $2\beta_\ell - \gamma_\ell \theta_\ell > \gamma_\ell \theta_{-\ell} + \delta_\ell > 0$ and $|\Delta_i| \geq |\Delta_{-i}| \geq 0$, we have $(2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} > 0$. Using the lower bound in Equation (16), we get for any $t > t_\eta$:
\[
g^*_t - g_{\ell,t} \geq (2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} + \eta (2\beta_\ell + \delta_\ell + \gamma_\ell) ,
\]
where (a) follows from the definition of $\eta$ such that $\eta < \frac{|(2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell})\Delta_{-\ell}|}{2\beta_\ell + \delta_\ell + \gamma_\ell}$. With $\Delta_\ell > 0$, the lower bound in Equation (14) implies $p^*_\ell - p_{\ell,t} > \Delta_\ell - \eta > 0$ because $0 < \eta < |\Delta_\ell|$. Hence, by combining these lower bounds, we get for any $t > t_\eta$
\[
(p^*_\ell - p_{\ell,t}) (g^*_t - g_{\ell,t}) > |\Delta_\ell + \eta| \cdot |(2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} + \eta (2\beta_\ell + \delta_\ell + \gamma_\ell)| := c_\eta > 0 .
\]

Similarly, if $\Delta_\ell < 0$, then using the upper bounds in Equations (15) and (13), we have
\[
g^*_t - g_{\ell,t} \leq (2\beta_\ell - \gamma_\ell \theta_\ell) \Delta_\ell - (\delta_\ell + \gamma_\ell \theta_{-\ell}) \Delta_{-\ell} - \eta (2\beta_\ell + \delta_\ell + \gamma_\ell) < 0
\]
and \( p_i^* - p_{i,t} < \Delta_t + \eta < 0 \), which again allows us to obtain the lower bound for \((p_i^* - p_{i,t}) (g_i^* - g_{i,t})\) in Equation (18).

**Step 3.** Telescoping Equation (11) from \( t_\eta + 1 \) to some large time period \( T + 1 > t_\eta + 1 \), and applying the lower bound in Equation (18), i.e., \((p_i^* - p_{i,t}) (g_i^* - g_{i,t}) \geq c_\eta > 0 \) for all \( t > t_\eta \), we get

\[
D_t(p_i^*, p_{i,T+1}) \leq D_t(p_i^*, p_{i,T}) - c_\eta \epsilon_{i,t} + c_2 \sum_{t=1}^T \epsilon_{i,t} \leq D_t(p_i^*, p_{i,t_\eta+1}) - c_\eta \sum_{t=t_\eta+1}^T \epsilon_{i,t} + c_2 \sum_{t=t_\eta+1}^T \epsilon_{i,t}^2, \tag{19}
\]

where we take the absolute constant \( c_2 > \max_{t \in \mathbb{N}} \frac{g_i^2}{2c_i} \), which is possible as the gradient \( g_{i,t} \) is a linear function in all arguments and prices/reference prices are bounded, and in (a) we used a telescoping argument. Since \( D_t(p_i^*, p_{i,T+1}) \geq 0 \), rearranging terms in Equation (19) and dividing both sides by \( \sum_{t=t_\eta+1}^T \epsilon_{i,t} \) we get the following:

\[
-\frac{D_t(p_i^*, p_{i,t_\eta+1})}{\sum_{t=t_\eta+1}^T \epsilon_{i,t}} \leq -c_\eta + c_2 \frac{\sum_{t=t_\eta+1}^T \epsilon_{i,t}^2}{\sum_{t=t_\eta+1}^T \epsilon_{i,t}}.
\]

Since \( D_t(p_i^*, p_{i,t_\eta+1}) \) and \( \epsilon_{i,t} \) are finite, the condition \( \lim_{T \to \infty} \sum_{t=1}^T \epsilon_{i,t} = \infty \) implies that \( \lim_{T \to \infty} \sum_{t=t_\eta+1}^T \epsilon_{i,t} = \infty \), and the condition \( \lim_{T \to \infty} \sum_{t=1}^T \epsilon_{i,t}^2 < \infty \) implies that \( \lim_{T \to \infty} \sum_{t=t_\eta+1}^T \epsilon_{i,t}^2 < \infty \). Hence, \( \frac{\sum_{t=t_\eta+1}^T \epsilon_{i,t}^2}{\sum_{t=t_\eta+1}^T \epsilon_{i,t}} = 0 \) as \( T \to \infty \). Finally, because \( c_\eta > 0 \), as \( T \to \infty \), the above left hand side goes to zero and the right hand side goes to \(-c_\eta < 0 \). This is a contradiction, implying that our original assumption \(|\Delta_t| > 0 \) cannot hold, and hence \( \Delta_t = \Delta_{-t} = 0 \). □

### 6.3. Proof of Theorem 5.2

Recall Equation (17): \( g_i^* - g_{i,t} = 2 \beta_i (p_i^* - p_{i,t}) - \delta_i (p_{i,-i}^* - p_{i,-i,t}) - \gamma_i (r^* - r_t) \). Combining this with Corollary D.1.1 we obtain

\[
D_i(p_i^*, p_{i,i+1})
\]

\[
\leq D_i(p_i^*, p_{i,t}) - \epsilon_{i,t} \left( 2 \beta_i (p_i^* - p_{i,t})^2 - \delta_i (p_{i,-i}^* - p_{i,-i,t}) (p_i^* - p_{i,t}) - \gamma_i (r^* - r_t) (p_i^* - p_{i,t}) \right) + \frac{(\epsilon_{i,t} g_{i,t})^2}{2 \sigma_i},
\]

\[
\leq D_i(p_i^*, p_{i,t}) - \epsilon_{i,t} \left( \frac{4 \beta_i - \delta_i}{2} (p_i^* - p_{i,t})^2 - \frac{\delta_i}{2} (p_{i,-i}^* - p_{i,-i,t})^2 \right) + \epsilon_{i,t} \gamma_i (r^* - r_t) (p_i^* - p_{i,t}) + \frac{(\epsilon_{i,t} g_{i,t})^2}{2 \sigma_i}.
\]

where in (a) we used the basic inequality \( AB \leq (A^2 + B^2)/2 \) for \( A = p_{i,-i}^* - p_{i,-i,t} \) and \( B = p_i^* - p_{i,t} \).

Now, consider the step-size sequences \( \{\epsilon_{i,t}\}_i \) that satisfy

\[
\frac{1}{t+1} \cdot \frac{10}{4 \beta_i - \delta_i} \leq \epsilon_{i,t} \leq \frac{1}{t+1} \cdot \frac{2}{\max(\delta_i, \gamma_i)}, \quad i = 1, 2.
\]

(20)
Equation (20) holds due to the fact that $\beta_i > m(\delta_i + \gamma_i) > 0$ and $m \geq 2$, which further implies $2(4\beta_i - \delta_i) > 8m(\delta_i + \gamma_i) - 2\delta_i > 10(\delta_i + \gamma_i) > 10 \max \{\delta_i, \gamma_i\}$. This leads to

$$D_i(p^*_i, p_{i,t+1})$$

$$\leq D_i(p^*_i, p_{i,t}) - \frac{5}{t+1} (p^*_i - p_{i,t})^2 + \frac{1}{t+1} (p^*_i - p_{i,t})^2 + \frac{2}{t+1} \rho_i (p^*_i - p_{i,t}) + \frac{(\epsilon_i,t_g_{i,t})^2}{2\sigma_i}$$

Where in (a) we take some $c_2 > \max_{i \in \{1,2\}, t \in \mathbb{N}^+} \frac{4g_{i,t}^2}{\sigma_i \max \{\delta_i, \gamma_i\}}$ for all $i, t$ by using the fact that $p_{i,t}, r_t \in \mathcal{P}$.

Summing across $i = 1, 2$, we have

$$\sum_{i=1,2} D_i(p^*_i, p_{i,t+1}) \leq \sum_{i=1,2} D_i(p^*_i, p_{i,t}) - \frac{4}{t+1} \|p^*_i - p_i\|^2 + \frac{2}{t+1} \rho_i (p^*_i - p_{i,t} + p^*_i - p_{2,t}) + \frac{c_2}{(t+1)^2}$$

Where in inequality (a) we applied $C(A + B) \leq \frac{c_2^2}{2} + \frac{(A+B)^2}{2} \leq \frac{c_2^2}{2} + A^2 + B^2$ for $A = p^*_i - p_{i,t}, B = p^*_i - p_{2,t}$ and $C = r^* - r_t$.

When $R_t(z) = z^2$, we have $D_i(p, p') = (p - p')^2$. Therefore, denoting $x_t = \sum_{i=1,2} D_i(p^*_i, p_{i,t}) = \|p^*_i - p_i\|^2$ and $x_{n,t} = (r^* - r_t)^2$, the equation above yields

$$x_{t+1} \leq \left(1 - \frac{2}{t+1}\right)x_t + \frac{1}{t+1} x_{n,t} + \frac{c_2}{(t+1)^2}. \quad (21)$$

We will show via induction that $x_t \leq \frac{c}{t}$ for some $c > 0$ and any $t \in \mathbb{N}^+$. The proof is constructive and will rely on the following definitions, whose motivations will later be clear.

Fix $\rho_o = \left\lceil \frac{a}{1-a} \right\rceil + 1$, $t_o = \left\lceil \frac{\rho_o + 1}{\rho_o} \right\rceil$, and take any $\bar{\theta}$ so that $\max \{\theta_1, \theta_2\} < \bar{\theta} < 1$. Here, $[x] = \min \{y \in \mathbb{N}^+ : y \geq x\}$ for any $x \in \mathbb{R}$. Note that $\rho_o$ is bounded as $a$ is bounded away from 1. Define

$$t_{\theta} := \min \left\{ t \in \mathbb{N}^+ : t \geq \rho_o \quad \text{and} \quad \frac{(\rho_o + 1) \log(\tau - \rho_o - 1)}{\rho_o + \rho_o + 1 - \frac{\rho_o}{\tau}} \leq \frac{\bar{\theta}}{\max \{\theta_1, \theta_2\}} \right\}$$

$$u := (1-a) \max \{\theta_1, \theta_2\} \sum_{t=1}^{\rho_o + t_{\theta} - 1} \frac{a-\tau}{\tau}. \quad (22)$$

Note that $t_{\theta}$ is bounded because $\max \{\theta_1, \theta_2\}$ is bounded away from one. Further, since $\rho_o$ and $\bar{\theta}$ are constant, and $\log(\tau) = o(\tau)$, it is easy to see that $t_{\theta}$ exists. Furthermore, define $c := \frac{2k(\bar{p} - p)^2 + c_2 + 1}{1 - \bar{\theta}}$ and

$$\bar{t} := \min \left\{ \tau > \max \{\rho_o + t_o, t_{\theta}\} : (t - \rho_o) \cdot \left(2(\bar{p} - p)^2 + u \cdot \frac{2t \cdot (\bar{p} - p)^2 + c_2 + 1}{1 - \bar{\theta}} \right) < a^{-\tau}, \right\} \quad (23)$$

Note that $t_{\theta}$ is bounded because $\max \{\theta_1, \theta_2\}$ is bounded away from one. Further, since $\rho_o$ and $\bar{\theta}$ are constant, and $\log(\tau) = o(\tau)$, it is easy to see that $t_{\theta}$ exists. Furthermore, define $c := \frac{2k(\bar{p} - p)^2 + c_2 + 1}{1 - \bar{\theta}}$ and
Such $\bar{t}$ must exist because the left hand side is quadratic in $t$, while the right hand side is exponential in $t$ for $a \in (0, 1)$. We provide an illustration for the size of $c$ w.r.t. memory parameter $a$ and $\max\{\theta_1, \theta_2\}$ in Figure 2b of Appendix C.4. These definitions of $\bar{t}$ and $c$ imply the following:

\[ a^t \left(2(\bar{p} - p)^2 + uc\right) < \frac{1}{t - \rho_a} \quad \forall t \geq \bar{t} \tag{24} \]

\[ 2\bar{t}(\bar{p} - p)^2 + c_2 + 1 + \bar{\theta}c = c \tag{25} \]

\[ c > 2\bar{t}(\bar{p} - p)^2. \tag{26} \]

Here, Equation (24) is due to the following: plugging the definition of $c$ into that of $\bar{t}$ we get $(\bar{t} - \rho_a) \cdot (2(\bar{p} - p)^2 + uc) < a^{-\bar{t}}$, and since $c = \frac{2\bar{t}(\bar{p} - p)^2 + c_2 + 1}{1 - \theta} \leq \frac{2t(\bar{p} - p)^2 + c_2 + 1}{1 - \theta}$ for any $t \geq \bar{t} > \rho_a$, we have

\[ (t - \rho_a) \left(2(\bar{p} - p)^2 + uc\right) \leq (t - \rho_a) \left(2(\bar{p} - p)^2 + u \cdot \frac{2t(\bar{p} - p)^2 + c_2 + 1}{1 - \theta}\right) < a^{-t}, \]

where (a) follows from the definition of $\bar{t}$. Equation (25) directly follows from the definition of $c$. Equation (26) follows because $\bar{\theta} \in (0, 1)$ and hence $c = \frac{2\bar{t}(\bar{p} - p)^2 + c_2 + 1}{1 - \theta} > 2\bar{t}(\bar{p} - p)^2 + c_2 + 1 > 2\bar{t}(\bar{p} - p)^2$.

Hence, this implies that $x_t \leq 2(\bar{p} - p)^2 < \frac{c}{\bar{t}}$ for any $t = 1 \ldots \bar{t}$, where we recall $x_t = \|p^* - p_t\|^2$.

Consider $t \geq \bar{t}$. We will now show via induction that $x_{t+1} \leq c/(t+1)$ using our induction hypothesis that $x_{\tau} \leq c/\tau$ holds for all $\tau = 1, \ldots, t$. Note that the base case $x_t \leq 2(\bar{p} - p)^2 < \frac{c}{\bar{t}}$ for any $t = 1 \ldots \bar{t}$ is trivially true as we just discussed. Then, multiplying $t(t+1)$ on both sides of the recurrence relation in Equation (21) and telescoping from $\bar{t}$ to $t$, we have

\[ t(t+1)x_{t+1} \leq (t-1)tx_t + tx_{n,t} + c_2 \leq (t-2)(t-1)x_{t-1} + \sum_{\tau=t-1}^{t} tx_{n,\tau} + 2c_2 \]

\[ \leq (\bar{t} - 1)\bar{t} . x_{\bar{t}} + \sum_{\tau=\bar{t}}^{t} \tau x_{n,\tau} + (t - \bar{t} + 1)c_2 \leq (\bar{t} - 1)\bar{t} . x_{\bar{t}} + \sum_{\tau=\bar{t}}^{t} \tau x_{n,\tau} + tc_2. \tag{27} \]

Here, (a) follows from telescoping. We will now bound $x_{n,\tau}$ for all $\tau = \bar{t} \ldots t$. Using the definition $r^* = \theta_1 p_{1}^* + \theta_2 p_{2}^*$, we get

\[ r^* - r_{\tau+1} = r^* - ar_{\tau} - (1-a) (\theta_1 p_{1,\tau} + \theta_2 p_{2,\tau}) = a (r^* - r_{\tau}) - (1-a) (\theta_1 (p_{1}^* - p_{1,\tau}) + \theta_2 (p_{2}^* - p_{2,\tau})). \]
By convexity, we further have for any $\tau = 1 \ldots t$,

$$x_{n,\tau} = (r^* - r_{\tau+1})^2 \leq ax_{n,\tau} + (1-a) (\theta_1 (p_1^* - p_{1,\tau}) + \theta_2 (p_2^* - p_{2,\tau}))^2$$

$$\leq ax_{n,\tau} + (1-a) \left( \theta_1 (p_1^* - p_{1,\tau})^2 + \theta_2 (p_2^* - p_{2,\tau})^2 \right)$$

$$\leq ax_{n,\tau} + (1-a) \max\{\theta_1, \theta_2\} x_{\tau} \overset{(a)}{\leq} ax_{n,\tau} + (1-a) \max\{\theta_1, \theta_2\} \frac{c}{\tau},$$

where (a) follows from the induction hypothesis, i.e., $x_{\tau} \leq c/\tau$ holds for all $\tau = 1 \ldots t$. Using a telescoping argument, we then have for any $t \geq \tilde{t}$,

$$x_{n,t+1} \leq ax_{n,t} + (1-a) \max\{\theta_1, \theta_2\} \frac{c}{t} \leq a'x_{n,1} + (1-a)c \max\{\theta_1, \theta_2\} \sum_{\tau=1}^{t} \frac{a^{\tau-t}}{\tau}$$

$$= a'x_{n,1} + (1-a)c \max\{\theta_1, \theta_2\} \left( \sum_{\tau=1}^{\rho_a+t_a-1} \frac{a^{\tau-t}}{\tau} + \sum_{\tau=\rho_a+t_a}^{t} \frac{a^{\tau-t}}{\tau} \right)$$

$$\overset{(a)}{=} a' (x_{n,1} + uc) + (1-a)c \max\{\theta_1, \theta_2\} \sum_{\tau=\rho_a+t_a}^{t} \frac{a^{\tau-t}}{\tau}$$

$$\leq a' \left( 2(\bar{p} - p)^2 + uc \right) + (1-a)c \max\{\theta_1, \theta_2\} \sum_{\tau=\rho_a+t_a}^{t} \frac{a^{\tau-t}}{\tau}$$

$$\overset{(b)}{\leq} \frac{1}{t - \rho_a} + (1-a)c \max\{\theta_1, \theta_2\} \sum_{\tau=\rho_a+t_a}^{t} \frac{a^{\tau-t}}{\tau} \overset{(c)}{\leq} \frac{1 + \max\{\theta_1, \theta_2\} c}{t - \rho_a}.$$

Here, (a) follows from the definition of $u$ in Equation (23); (b) follows from Equation (24); and (c) evokes Lemma D.3 since $t \geq \tilde{t} \geq \rho_a + t_a$. Applying this upper bound on $x_{n,t}$ in Equation (27) we have

$$t(t+1)x_{t+1} \leq (t-1)\tilde{t} \cdot x_{\tilde{t}} + (1 + \max\{\theta_1, \theta_2\} c) \sum_{\tau=t}^{t} \frac{\tau}{t - \rho_a - 1} + tc_2$$

$$= (t-1)\tilde{t} \cdot x_{\tilde{t}} + (1 + \max\{\theta_1, \theta_2\} c) \sum_{\tau=t}^{t} \left( 1 + \frac{\rho_a + 1}{\tau - \rho_a - 1} \right) + tc_2$$

$$\leq (t-1)\tilde{t} \cdot x_{\tilde{t}} + (1 + \max\{\theta_1, \theta_2\} c) (t + (\rho_a + 1) \log(t - \rho_a - 1)) + tc_2.$$
Here, (a) follows from the fact that \( t \geq \tilde{t} \), so \[ \frac{2(\tilde{t}-1)(\bar{p}-p)^2}{t} + c_2 \leq 2\bar{t}(\bar{p}-\tilde{p})^2 + c_2; \] (b) follows from \( t \geq \tilde{t} \geq t_\theta \) so that \[ \frac{(\rho a+1)\log(t-\rho a-1)}{t} \leq \frac{\delta}{\max\{\theta_1,\theta_2\}} - 1 \] according to Equation (22); finally, (c) follows from Equation (25). □

References


Appendices for No-regret Learning in Price Competitions under Consumer Reference Effects

Appendix A: Appendix for Section 3

A.1. Proof of Lemma 3.2

Let \((p^*, r^*) \in (\bar{p}, \bar{p})^3\) be an interior SNE, whose existence is guaranteed by Assumption 1. Since revenue functions are quadratic, first order conditions at the interior best-response profiles should hold, which means the derivative of revenue functions at \(p_1^* = \psi_1(p_2^*, r^*)\) and \(p_2^* = \psi_2(p_1^*, r^*)\) should be 0: \(\frac{\partial \pi_1(p^*, r^*)}{\partial p_1} = \frac{\partial \pi_2(p^*, r^*)}{\partial p_2} = 0\).

This leads to the relationship \(\alpha_1 - 2\beta_1 \psi_1(p_2^*, r^*) + \delta_1 p_2^* + \gamma_1 r^* = \alpha_2 - 2\beta_2 \psi_2(p_1^*, r^*) + \delta_2 p_1^* + \gamma_2 r^* = 0\). Solving for the best-response equations, we get \(p_1^* = \psi_1(p_2^*, r^*) = \frac{\alpha_1 + \delta_1 p^* + \gamma_1 r^*}{2\beta_1}\) for \(i = 1, 2\). Finally, the definition of an SNE guarantees \(r^* = \theta_1 p_1^* + \theta_2 p_2^*\). Thus, solving for \((p^*, r^*)\), we obtain the unique solution:

\[
p_1^* = \frac{2\alpha_i \beta_i - \alpha_i \theta_i \gamma_i + \alpha_i \beta_i \gamma_i + \alpha_i (\delta_i + \delta_i \gamma_i)}{(2\beta_1 - \theta_1 \gamma_1) (2\beta_2 - \theta_2 \gamma_2) - (\theta_2 \gamma_1 + \delta_2) (\theta_1 \gamma_2 + \delta_2)} \quad i = 1, 2
\]

\[
r^* = \frac{\theta_1 (2\alpha_1 \beta_2 + \alpha_2 \delta_1) + \theta_2 (2\alpha_2 \beta_1 + \alpha_1 \delta_2)}{(2\beta_1 - \theta_1 \gamma_1) (2\beta_2 - \theta_2 \gamma_2) - (\theta_2 \gamma_1 + \delta_2) (\theta_1 \gamma_2 + \delta_2)}
\]

This implies that under Assumption 1 the interior SNE is unique. We remark that for any \(i = 1, 2\), because \(\beta_i \geq m(\delta_i + \gamma_i) > 0\) and \(m \geq 2 > 1\) we have \(2\beta_i - \theta_1 \gamma_i > \beta_i - \theta_1 \gamma_i > \delta_i + \gamma_i - \theta_1 \gamma_i = \theta_{-1} \gamma_i + \delta_i\). Hence, \(p_i^*, r^* > 0\).

Appendix B: Appendix for Section 4

B.1. Proof of Proposition 4.1

First of all, it is easy to see prices at the first period are identical between Algorithm 1 and 2 \(p_{i,1} = \arg \max_{p \in \mathcal{P}} R_i\) for \(i = 1, 2\) and \(p_{n,1} = r_1\). We now use induction to show price trajectories of the two algorithms are identical via considering the induction hypothesis that prices and reference prices are the same up to period \(t \in \mathbb{N}^+\). Note that \(R_n(z) = \frac{1}{2} z^2\) implies \(R_n(z) = z\). Then, the proxy variable update step for nature is:

\[
y_{n,t+1} = p_{n,t} - (1 - a) \frac{\partial \pi_n(p)}{\partial p_n} \bigg|_{p = p_{1,t}p_{2,t}p_{n,t}} = p_{n,t} - (1 - a) (p_{n,t} - \theta_1 p_{1,t} - \theta_2 p_{2,t})
\]

\[
y_{n,t+1} = ar_t + (1 - a) (\theta_1 p_{1,t} + \theta_2 p_{2,t}) = r_{t+1}.
\]

Since \(y_{n,t+1} = r_{t+1} \in \mathcal{P}\) the projection step for nature is trivial, which means \(p_{n,t+1} = y_{n,t+1} = r_{t+1}\). Furthermore, it is not difficult to see that prices \(p_{1,t+1} = \Pi_{p \in \mathcal{P}} (y_{1,t+1})\) and \(p_{2,t+1} = \Pi_{p \in \mathcal{P}} (y_{2,t+1})\) are identical between the two algorithms under the induction hypothesis. This implies that Algorithm 2 indeed recovers the prices and reference prices produced by Algorithm 1. □
B.2. Proof of Proposition 4.2

Directly considering first order conditions for the cost functions \(\{\tilde{p}_i\}_{i=1,2,n}\), we have the system of equations

\[
0 = \frac{\partial \tilde{p}_i(p_i, p_{-i}, p_n)}{\partial p_i} = 2\beta_i p_i - (\alpha_i + \delta_i p_{-i} + \gamma_i r_i), \quad i = 1, 2 \quad \text{and} \quad 0 = \frac{\partial \tilde{p}_n(p_1, p_2, p_n)}{\partial p_n} = p_n - (\theta_1 p_1 + \theta_2 p_2).
\]

Solving these equations results in a unique solution that is identical to that in Equation (28), which is the unique interior SNE according to Lemma 3.2. Since the SNE is an interior point of \((\bar{p}, \bar{p})^3\), it is the unique PSNE of the induced static 3-firm game. □

Appendix C: Appendix for Section 5

C.1. Proof of Theorem 5.3

Here, we first provide a roadmap for the proof. Evoking Corollary D.1.1 we get

\[
D_i(p_i^*, p_{i,t+1}) \leq D_i(p_i^*, p_{i,t}) - \epsilon_i t (g_i^* - g_{i,t})(p_i^* - p_{i,t}) + \frac{(\epsilon_i t)^2 (g_i^* - g_{i,t})^2}{2\sigma_i}.
\]  

(29)

By bounding the first order term \((g_i^* - g_{i,t})(p_i^* - p_{i,t})\) and the second order term \(\frac{(g_i^* - g_{i,t})^2}{2\sigma_i}\), we achieve a recursive relation in the form of \(\sum_{i=1,2,n} D_i(p_i^*, p_{i,t+1}) \leq \sum_{i=1,2,n} D_i(p_i^*, p_{i,t}) + \sum_{i=1,2,n} \kappa_i \epsilon_i x_{i,t}\), where we recall the definition \(x_{i,t} = (p_i^* - p_{i,t})^2\) for \(i = 1, 2, n\) as in the proof of Theorem 5.2 and \(\kappa_i\) is some constant that takes negative values if the conditions in the theorem’s statement are satisfied. We then argue if \((p_t, r_t)\) does not converge to the SNE, \(\sum_{i=1,2,n} D_i(p_i^*, p_{i,t})\) will be greater than some positive constant \(\epsilon > 0\) for all large enough \(t\). Combining this with the above recursive relationship, this further implies that the distance between the price profile \((p_t, r_t)\) and the SNE decreases by a positive constant for each period. This will eventually contradict the fact that Bregman divergence is positive.

We start our proof by recalling Equation (17): \(g_i^* - g_{i,t} = 2\beta_i (p_i^* - p_{i,t}) - \delta_i (p_{-i}^* - p_{-i,t}) - \gamma_i (r^* - r_t)\).

Furthermore, for \(i = 1, 2\), we have

\[
(g_i^* - g_{i,t})^2 \leq 8\beta_i^2 x_{i,t} + 4\delta_i^2 x_{-i,t} + 4\gamma_i^2 x_{n,t},
\]

(30)

where we used \((A + B + C)^2 \leq 2A^2 + 2(B + C)^2 \leq 2A^2 + 4B^2 + 4C^2\) for \(A = 2\beta_i (p_i^* - p_{i,t}),\ B = \delta_i (p_{-i}^* - p_{-i,t}),\) and \(C = \gamma_i (r^* - r_t)\). Hence, we have

\[
D_i(p_i^*, p_{i,t+1}) \leq D_i(p_i^*, p_{i,t}) - \epsilon_i t \left(2\beta_i x_{i,t} - \delta_i \left(p_{-i}^* - p_{-i,t}\right)(p_i^* - p_{i,t}) - \gamma_i (r^* - r_t)(p_i^* - p_{i,t})\right) + \frac{(\epsilon_i t g_i^*)^2}{2\sigma_i}
\]

\[
\leq D_i(p_i^*, p_{i,t}) - \epsilon_i t \left(2\beta_i x_{i,t} - \delta_i \left(x_{i,t} + x_{-i,t}\right) - \frac{\gamma_i}{2} \left(x_{n,t} + x_{i,t}\right)\right) + \frac{(\epsilon_i t g_i^*)^2}{2\sigma_i}
\]

\[
= D_i(p_i^*, p_{i,t}) - \epsilon_i t \left(\frac{4\beta_i - \delta_i - \gamma_i}{2} x_{i,t} - \frac{\delta_i}{2} x_{-i,t} - \frac{\gamma_i}{2} x_{n,t}\right) + \frac{(\epsilon_i t g_i^*)^2}{2\sigma_i}
\]
where for\

\[
D_i(p_{i}^*, p_{i+1}) \leq D_i(p_i^*, p_{i+1}) - (1 - a) \left( \frac{1}{2} x_{n,t} - \frac{\theta_1}{2} x_{1,t} - \frac{\theta_2}{2} x_{2,t} \right) + \frac{(1 - a) x_{n,t}}{2} + \frac{1}{2} (\theta_1 x_{1,t} + \theta_2 x_{2,t}),
\]

where (a) follows from \( \theta_1 + \theta_2 = 1 \) and convexity.

Summing up Equations (31) (over \( i = 1, 2 \)) and (32), and collecting terms yields

\[
\sum_{i=1,2,n} D_i(p_i^*, p_{i+1}) \leq \sum_{i=1,2,n} D_i(p_i^*, p_{i+1}) + \sum_{i=1,2,n} \kappa_i x_{i,t},
\]

where the coefficient for \( x_{i,t} \) is

\[
\kappa_{i,t} = \begin{cases} 
\frac{(4\beta_i - \delta_i - \gamma_i)z}{2\beta_i} (1-a)\sigma_i + 4z^2(1-a)^2\sigma_i + \frac{\delta i - \gamma_i}{2\beta_i} (1-a)\sigma_i - \frac{2\delta_i^2z^2(1-a)^2\sigma_i}{3\beta_i} + \frac{(1-a)\theta_1}{2} + \frac{(1-a)\theta_2}{4}, & i = 1, 2 \\
\frac{1-a}{2} + \frac{(1-a)^2}{4} + \frac{\gamma_i z^2}{2\sigma_i} + \frac{\gamma_i^2 z^2}{2\sigma_i^2} + \frac{2\delta_i^2 z^2}{2\sigma_i^2}, & i = n
\end{cases}
\]

Now, for \( i = 1, 2 \), consider taking step size \( \epsilon_i = \frac{z\epsilon}{\beta_i} (1-a) \), for some constant \( z > 0 \) that will be determined later, and denote the corresponding \( \kappa_{i,t} \) as \( \kappa_i(z) \) (we drop the dependence on \( t \) as step sizes are constant), where for \( i = 1, 2 \),

\[
\kappa_i(z) \overset{(a)}{=} \frac{(4\beta_i - \delta_i - \gamma_i)z}{2\beta_i} (1-a)\sigma_i + 4z^2(1-a)^2\sigma_i + \frac{\delta i - \gamma_i}{2\beta_i} (1-a)\sigma_i - \frac{2\delta_i^2z^2(1-a)^2\sigma_i}{3\beta_i} + \frac{(1-a)\theta_1}{2} + \frac{(1-a)\theta_2}{4} \\
\leq \frac{(4 - \frac{1}{2})z}{2 (1-a)\sigma_i + 4z^2(1-a)^2\sigma_i + \frac{z}{2m} (1-a)\sigma_i - \frac{2z^2}{m^2} (1-a)\sigma_i - \frac{3(1-a)}{4} \\
= (1-a) \left( \frac{4\sigma_i + \frac{2\sigma_i}{m^2}}{2m} \right) z^2 - \left( \frac{2 - \frac{1}{2m}}{2m} \right) \sigma_i - \frac{\sigma_i}{2m} z + \frac{3}{4} \right) = (1-a) f_{i,m}(z).
\]

Here, in (a) we substitute \( \epsilon_i = \frac{z\epsilon}{\beta_i} (1-a) \) for \( i = 1, 2 \); in (b) we use the fact that \( \theta_i, a \in (0, 1) \) (which implies \( (1-a)^2 \leq 1-a \)) and \( \beta_i > m(\delta_i + \gamma_i) > m \max \{\delta, \gamma\} \). We follow a similar argument as above and obtain

\[
\kappa_n(z) \overset{(a)}{=} \frac{1-a}{2} + \frac{(1-a)^2}{4} + \frac{z\gamma_i \sigma_i}{2\beta_i} (1-a) + \frac{2\gamma_i^2 \sigma_i}{\beta_i^2} (1-a)^2 + \frac{z\gamma_2 \sigma_2}{2\beta_2} (1-a) + \frac{2\gamma_2^2 \sigma_2^2}{\beta_2^2} (1-a)^2
\]
where in (a) we substitute \( \epsilon_{i,t} = \frac{\alpha_i}{n_t} (1 - a) \) for \( i = 1, 2 \); in (b) we used the fact that \( \theta_i, a \in (0, 1) \) and 
\[ \beta_i > m(\delta_i + \gamma_i) > m \max \{\delta_i, \gamma_i\} \] 
for any \( i = 1, 2 \).

Now, recall the definition \( S_{i,m} = \{ z > 0 : f_{i,m}(z) < 0 \} \). Then, if we have \( \cap_{i=1,2,n} S_{i,m} \neq \emptyset \), taking any \( s \in \cap_{i=1,2,n} S_{i,m} \) yields \( \kappa_i(s) < 0 \) for \( i = 1, 2, n \). Hence, Equation (33) now becomes

\[
\sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t+1}) \leq \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t}) + \sum_{i=1,2,n} \kappa_i(s) x_{i,t}, \quad \kappa_i(s) < 0.
\]

Therefore, we know that

\[
\sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t+1}) < \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t}).
\]

Furthermore, by strong convexity, \( \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t}) \geq \sum_{i=1,2,n} \frac{\sigma_i^2}{2} (p_{i}^* - p_{i,t})^2 \geq \frac{\min_{i=1,2,n} \sigma_i^2}{2} ||p^* - p_t|| \). Hence, for any small \( \epsilon > 0 \), if there exists some \( t_\epsilon \in \mathbb{N}^+ \) such that \( \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t_\epsilon}) \leq \frac{\min_{i=1,2,n} \sigma_i^2}{2} \), then by Equation (38), \( \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t}) \leq \frac{\min_{i=1,2,n} \sigma_i^2}{2} \) for all \( t \geq t_\epsilon \), which further implies \( \|p^* - p_t\| \leq \epsilon \) for all \( t \geq t_\epsilon \). Hence \( (p_t, r_t) \overset{t \to \infty}{\to} (p^*, r^*) \).

Thus, it remains to show that for any small \( \epsilon > 0 \), there exists \( t_\epsilon > 0 \) such that \( \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t_\epsilon}) < \epsilon \). We will prove this by contradiction. If this is not the case, there exists \( \epsilon > 0 \), and \( \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t}) \geq \epsilon \) for all \( t \geq 0 \). Define \( R(z_1, z_2, z_3) = \sum_{i=1,2,3} R_i(z_i) \) for any \( z_1, z_2, z_3 \in \mathbb{R} \), and slightly abuse the notation to define \( D : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R} \) as the Bregman divergence with respect to \( R \). In the rest of this proof for simplicity we also write \( p^* = (p_{1}^*, p_{2}^*, r^*) \) and \( p_t = (p_{1,t}, p_{2,t}, r_t) \). A simple analysis shows \( D(p^*, p_t) = \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t}) \).

Since \( R_i \) is continuously differentiable (by definition of Bregman divergence), \( R \) is also continuously differentiable, and hence it is easy to see for any \( x, y \in \mathbb{R}^3 \) there exists \( \delta > 0 \) such that \( D(x, y) < \epsilon \) for \( \forall \|x - y\| < \delta \). Since we assumed \( D(p^*, p_t) = \sum_{i=1,2,n} D_i(p_{i}^*, p_{i,t}) \geq \epsilon \) for all \( t \geq 0 \), the above implies \( \|p^* - p_t\| \geq \delta \) for all \( t \geq 0 \). Hence following Equation (37),

\[
D(p^*, p_{t+1}) \leq D(p^*, p_t) + \sum_{i=1,2,n} \kappa_i(s) (p_{i}^* - p_{i,t})^2 \leq D(p^*, p_t) + \max_{i=1,2,n} \kappa_i(s) \sum_{i=1,2,n} (p_{i}^* - p_{i,t})^2
\]

\[
= D(p^*, p_t) + \max_{i=1,2,n} \kappa_i(s) \cdot \|p^* - p_t\|^2 \overset{(a)}{\leq} D(p^*, p_t) + \delta^2 \max_{i=1,2,n} \kappa_i(s)
\]

\[
\overset{(b)}{\leq} D(p^*, p_1) + t\delta^2 \max_{i=1,2,n} \kappa_i(s),
\]
where (a) follows because \( \kappa_i(s) < 0 \) for \( i = 1, 2, n \) and \( \| p_t - p_i \| > \delta \) for all \( t \geq 0 \); (b) follows from a telescoping argument. Finally, \( \max_{i=1,2,n} \kappa_i(s) < 0 \) implies the right hand side in the above inequality goes to negative infinity as \( t \) goes to infinity. This implies that \( D(p^*, p_{t+1}) = \sum_{i=1,2,n} D_i(p_i^*, p_{i,t}) \leq -\infty \), which contradicts nonnegativity of Bregman divergence. Hence, for any small \( \epsilon > 0 \), there exists \( t_\epsilon > 0 \) such that \( \sum_{i=1,2,n} D_i(p_i^*, p_{i,t}) < \epsilon \), concluding the proof. \( \square \)

### C.2. Proof of Corollary 5.3.1

When \( \sigma_1 = \sigma_2 = \sigma \), Equation (5) becomes \( f_{i,m}(z) = \begin{cases} 2\sigma \left( 2 + \frac{1}{m\sigma} \right) z^2 - \sigma \left( 2 - \frac{1}{m\sigma} \right) z + \frac{3}{4} & \text{if } i = 1, 2 \\ \frac{4\sigma}{m\sigma} z^2 + \frac{m}{2\sigma} z - \frac{1}{4} & \text{if } i = n \end{cases} \).

Since in this case the function \( f_{i,m}(z) \) and \( f_{j,m}(z) \) are identical, we only consider \( f_{i,m}(z) \). Note that the function \( f_{i,m}(z) \) has two distinct zero roots if and only if its discriminant is strictly greater than 0, i.e., \( (1 - \frac{1}{2m\sigma})^2 - \frac{3}{2\sigma} (2 + \frac{1}{m\sigma}) > 0 \) which is equivalent to \( \sigma > \frac{6(2m^2 + 1)}{(2m - 1)^2} \). Therefore, when \( f_{i,m} \) has two distinct zero roots, the smaller one is

\[
z_1 = \frac{1 - \frac{1}{2m\sigma} - \sqrt{(1 - \frac{1}{2m\sigma})^2 - \frac{3}{2\sigma} (2 + \frac{1}{m\sigma})}}{2 (2 + \frac{1}{m\sigma})} = \frac{1}{\sigma} \cdot \frac{3/4}{\frac{1 - \frac{1}{2m\sigma} + \sqrt{(1 - \frac{1}{2m\sigma})^2 - \frac{3}{2\sigma} (2 + \frac{1}{m\sigma})}}{\frac{1 - \frac{1}{2m\sigma} + \sqrt{(1 - \frac{1}{2m\sigma})^2 - \frac{3}{2\sigma} (2 + \frac{1}{m\sigma})}}}} > 0. \tag{39}
\]

Similarly, the discriminant of \( f_{n,m} \) (i.e., \( \frac{1}{m\sigma} + \frac{4}{\sigma m^2} \)) is always positive, so \( f_{n,m} \) always has two zero roots. The larger one is given by

\[
z_2 = \frac{-\frac{1}{m\sigma} + \sqrt{\frac{1}{m\sigma} + \frac{4}{\sigma m^2}}}{\frac{8}{m^2}} = \frac{1}{\sigma} \cdot \frac{1/2}{\frac{-\frac{1}{m\sigma} + \sqrt{\frac{1}{m\sigma} + \frac{4}{\sigma m^2}}}{\frac{-\frac{1}{m\sigma} + \sqrt{\frac{1}{m\sigma} + \frac{4}{\sigma m^2}}}}}} > 0. \tag{40}
\]

For any \( \sigma > \frac{6(2m^2 + 1)}{(2m - 1)^2} \), the two roots of \( f_{i,m}(z) \) are both positive, while \( f_{n,m}(z) \) always has one positive root and one negative root. Hence, using a simple geometric argument regarding two quadratic functions, it is easy to see that if \( z_2 > z_1 \), any \( s \in (z_1, z_2) \) satisfies \( f_{i,m}(s), f_{n,m}(s) < 0 \).

Now, consider \( z_2 - z_1 = \frac{1}{4}(B - A) \). Since we observe \( A \) is decreasing in \( \sigma \) and \( B \) is increasing in \( \sigma \), we have \( B - A \) is increasing in \( \sigma \). By direct calculations, we see that when \( \sigma = \frac{(2m^2 + 7)^2}{8m^3 - 36m + 8} > 0 \), \( z_1 = z_2 \) (i.e., \( B - A = 0 \)). Therefore because \( B - A \) is increasing in \( \sigma \), we conclude \( B - A > 0 \) for any \( \sigma > \sigma_0 := \max \left\{ \frac{6(2m^2 + 1)}{(2m - 1)^2}, \frac{(2m^2 + 7)^2}{8m^3 - 36m + 8} \right\} \), which implies \( z_2 > z_1 \) for any \( \sigma > \sigma_0 \).

In sum, we conclude for any \( m > 2 \), if \( \sigma > \sigma_0 \), there exists \( s > 0 \) that depends on \( \sigma \) and \( m \) such that \( f_{i,m}(s) < 0 \) for \( i = 1, 2, n \), and by Theorem 5.3 this implies that there exist constant step sizes under which prices and reference prices converge to the unique interior SNE. \( \square \)
C.3. Proof of Theorem 5.4

\[ R_i(x) = \frac{\sigma}{2} x^2 \text{ for } i = 1, 2 \] implies \( \sigma_1 = \sigma_2 = \sigma \), and Equation (5) becomes \( f_{i,m}(z) = \)

\[
\begin{cases}
2 \sigma \left( 2 + \frac{1}{m} \right) z^2 - \sigma \left( 2 - \frac{1}{m} \right) z + \frac{3}{4}, & i = 1, 2 \\
\frac{4\sigma}{m} z^2 + \frac{\sigma}{m^2} z - \frac{1}{4}, & i = n
\end{cases}
\]

Define \( h_{i,m}(z) := f_{i,m}(z)/2\sigma \) for \( i = 1, 2 \) and \( h_{n,m}(z) := f_{n,m}(z) \):

\[
h_{i,m}(z) = \begin{cases}
\left( 2 + \frac{1}{m} \right) z^2 - \left( 1 - \frac{1}{2m} \right) z + \frac{3}{8\sigma}, & i = 1, 2 \\
\frac{4\sigma}{m} z^2 + \frac{\sigma}{m^2} z - \frac{1}{4}, & i = n
\end{cases}
\] (41)

Note that for any \( i = 1, 2, n \), \( f_{i,m}(z) < 0 \) if and only if \( h_{i,m}(z) < 0 \). Hence, according to Corollary 5.3.1 we know that when \( m > 2 \) and \( \sigma > \sigma_0 = \max \left\{ \frac{2(2n^2 + 1) - \frac{4}{2(n+1)} + \frac{1}{4}}{8m^2 - \frac{36m + 8}{m^2}}, \frac{2m^2 + \gamma}{8m^3 - 30m + 8} \right\} \), for any \( M \in (z_1, z_2) \) (defined in Equations (39) and (40)) we have \( f_{i,m}(s) < 0 \) for \( i = 1, 2, n \), which implies \( h_{i,m}(s) < 0 \) for \( i = 1, 2, n \). Furthermore, via a simple geometric argument, the quadratic functions \( h_{1,m} \) (with two positive zero roots) and \( h_{n,m} \) (with two zero roots, one positive and one negative) have a unique intersection point \( \tilde{s} \in (z_1, z_2) \). Define \( H := h_{1,m}(\tilde{s}) = h_{2,m}(\tilde{s}) = h_{n,m}(\tilde{s}) < 0 \). Furthermore, since \( \min_{m \geq 0} h_{n,m}(z) = -\frac{1}{4} \), we have

\[ -\frac{1}{4} \leq H = h_{1,m}(\tilde{s}) = h_{2,m}(\tilde{s}) = h_{n,m}(\tilde{s}) < 0. \]

Now, note that when \( R_1(x) = R_2(x) = \frac{\sigma}{2} x^2 \), \( D_t(p, p') = D_2(p, p') = \frac{\sigma}{2} (p - p')^2 \). Also recall \( R_n(x) = \frac{1}{2} x^2 \), so \( D_n(p, p') = \frac{1}{2} (p - p')^2 \). Hence, \( \sum_{i=1,2,n} D_i(p_i, p_{i,t}) = \frac{1}{2} (\sigma x_t + x_{n,t}) \), where we define \( x_{t} = \|p^* - p_t\|^2 \) and \( x_{n,t} = (r^* - r_t)^2 \) as in the proof of Theorem 5.3. Hence, by taking \( \epsilon_{i,t} = \frac{\sigma (1 - a)}{\eta_i} \), and continuing from Equation (37), we get

\[
\frac{1}{2} \left( \sigma x_{t+1} + x_{n,t+1} \right) \leq \frac{1}{2} \left( \sigma x_t + x_{n,t} \right) + \sum_{i=1,2,n} \kappa_i(z)x_{i,t} \overset{(a)}{=} \frac{1}{2} \left( \sigma x_t + x_{n,t} \right) + (1 - a) \sum_{i=1,2,n} f_{i,m}(z)x_{i,t} \overset{(b)}{=} \frac{1}{2} \left( \sigma x_t + x_{n,t} \right) + (1 - a) \left( \sum_{i=1,2} \sigma h_{i,m}(\tilde{s})x_{i,t} + h_{n,m}(\tilde{s})x_{n,t} \right) \overset{(c)}{=} \frac{1}{2} \left( \sigma x_t + x_{n,t} \right) + (1 - a) H \cdot (\sigma x_t + x_{n,t}) \]

Here, (a) follows from upper bounding \( \kappa_i(z) \) with \( f_{i,m}(z) \) for any \( z > 0 \) and \( i = 1, 2, n \) in Equations (35) and (36) within the proof of Theorem 5.3; (b) follows from the definition of \( h_{i,m} \) in Equation (41); (c) follows from the definition of \( H := h_{i,m}(\tilde{s}) \in [-\frac{1}{4}, 0] \) for \( i = 1, 2, n \) and \( x_t = x_{1,t} + x_{2,t} \).

Using a telescoping argument, we have \( \sigma x_t < \sigma x_t + x_{n,t} \leq (1 + 2(1 - a) H) (\sigma x_t + x_{n,t}) \leq (\sigma x_t + x_{n,t}) \left( 1 + \frac{a}{2} \right) \), where the final inequality follows from \( 0 < 1 + 2(1 - a) H \leq \frac{1 + a}{2} \) since \( H \in [-\frac{1}{4}, 0] \). Finally, because \( x_1 \leq 2 (\bar{p} - \hat{p})^2 \) and \( x_{n,t} \leq (\bar{p} - \hat{p})^2 \), we have \( x_t < (x_1 + x_2, n, 1) \left( 1 + \frac{a}{2} \right) \leq \frac{1 + a}{2} (\bar{p} - \hat{p})^2 \left( 1 + \frac{a}{2} \right) \). \( \square \)
C.4. Supplementary Figures for Section 5

![Figure 2](image)

**Figure 2** (a) $\sigma_0$ as a function of sensitivity margin $m$, where $\sigma_0$ is defined in Corollary 5.3.1 (b) Illustration of absolute constant $c$ in Theorem 5.2 w.r.t. memory parameter $a$ and $\max\{\theta_1, \theta_2\}$. All other model parameters take respective values as in Example 1 and firm $i = 1, 2$ again adopts regularizer $R_i(z) \equiv \frac{1}{2}z^2$.

Appendix D: Supplementary Lemmas of Section 5

**Lemma D.1** For $i = 1, 2, n$ and any $\tilde{z} \in \mathcal{P}$, we have for any $t \in \mathbb{N}^+$,

$$D_i(\tilde{z}, p_{i,t+1}) \leq D_i(\tilde{z}, p_{i,t}) + \epsilon_{i,t} \cdot g_{i,t}(\tilde{z} - p_{i,t}) + \frac{(\epsilon_{i,t} g_{i,t})^2}{2\sigma_i}.$$  

**Proof:** In the projection step of Algorithm 1, we have $p_{i,t+1} = \Pi_P(y_{i,t+1})$. Since we are working with one-dimensional decision sets, it is easy to see that $\Pi_P(y_{i,t+1}) = \arg\min_{\tilde{p} \in \mathcal{P}} D_i(\tilde{p}, y_{i,t+1})$ due to convexity of $R_i$. Recalling the definition $R'_i(p) = \frac{dR_i(z)}{dz} \bigg|_{z=p}$, we have $p_{i,t+1} = \arg\min_{p \in \mathcal{P}} D_i(p, y_{i,t+1}) = \arg\min_{p \in \mathcal{P}} R_i(p) - R_i(y_{i,t+1}) - R'_i(y_{i,t+1})(p - y_{i,t+1})$, so

$$p_{i,t+1} = \arg\min_{p \in \mathcal{P}} R_i(p) - p \cdot R'_i(y_{i,t+1}) \overset{(a)}{=} \arg\min_{p \in \mathcal{P}} R_i(p) - p \cdot (R'_i(p_{i,t}) - \epsilon_{i,t} g_{i,t})$$

$$= \arg\min_{p \in \mathcal{P}} R_i(p) - R_i(p_{i,t}) - R'_i(p_{i,t})(p - p_{i,t}) + p \cdot \epsilon_{i,t} g_{i,t} = \arg\min_{p \in \mathcal{P}} D_i(p, p_{i,t}) + p \cdot \epsilon_{i,t} g_{i,t}.$$  

Here (a) follows from the proxy update step in Algorithm 1. Now, evoking Lemma D.2 (ii) by taking $x = p$, $f(p) = p \cdot \epsilon_{i,t} g_{i,t}$, $z = p_{i,t}$, $y = \tilde{z} \in \mathcal{P}$, we have $D_i(\tilde{z}, p_{i,t+1}) \leq D_i(\tilde{z}, p_{i,t}) + \epsilon_{i,t} g_{i,t}(\tilde{z} - p_{i,t}) - D_i(p_{i,t}, p_{i,t+1})$. So

$$D_i(\tilde{z}, p_{i,t+1}) \leq D_i(\tilde{z}, p_{i,t}) + \epsilon_{i,t} g_{i,t}(\tilde{z} - p_{i,t}) + \epsilon_{i,t} g_{i,t}(p_{i,t} - p_{i,t+1}) - D_i(p_{i,t}, p_{i,t+1}) \overset{(a)}{=} D_i(\tilde{z}, p_{i,t}) + \epsilon_{i,t} g_{i,t}(\tilde{z} - p_{i,t}) + \epsilon_{i,t} g_{i,t}(p_{i,t} - p_{i,t+1}) - \frac{\sigma_i}{2} (p_{i,t} - p_{i,t+1})^2$$

$$\leq D_i(\tilde{z}, p_{i,t}) + \epsilon_{i,t} g_{i,t}(\tilde{z} - p_{i,t}) + \frac{(\epsilon_{i,t} g_{i,t})^2}{2\sigma_i},$$

where (a) follows from strong convexity of $R_i$. \qed
Corollary D.1.1 Under Assumption 1, let \((p^*, r^*)\) be the unique interior SNE as illustrated in Lemma 3.2, then for \(i = 1, 2, n\), \(g_i^* = \frac{\partial \tilde{g}_i}{\partial x_i} \big|_{p=p^*, r=r^*} = 0\), and for any \(t \in \mathbb{N}^+\), \(D_i(p_i^*, p_{i,t+1}) \leq D_i(p_i^*, p_{i,t}) - \epsilon, (g_i^* - g_i^*) (p_i^* - p_{i,t}) + \frac{(e, x_i, t) \sum a_i}{2\sigma_i}\).

Proof: Similar to the proof of Lemma 3.2 and Proposition 4.2, the SNE \((p^*, r^*)\) must satisfy first order conditions w.r.t. quadratic cost function \(\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_n\), respectively, due to the fact that it lies in the interior of the decision set. So \(g_i^* = 0\) for \(i = 1, 2, n\). Furthermore, Evoking Lemma D.1 by replacing \(z\) with \(p_i^*\) and combining \(g_i^* = 0\) yields the second part of the proof. \(\square\)

Lemma D.2 (Lemma 3.1 and 3.2 of Chen and Teboulle (1993)) Let \(D: \mathbb{C} \times \mathbb{C} \to \mathbb{R}^+\) be the Bregman divergence w.r.t convex function \(R\) on the convex set \(\mathbb{C}\). Then, (i) For any \(x, y, z \in \mathbb{C}\), \(D(x, y) + D(y, z) = D(x, z) + (R'(z) - R'(y)) (x - y)\); (ii) Let \(f: \mathbb{C} \to \mathbb{R}\) be any convex function and \(z \in \mathbb{C}\). If \(x^* = \arg\min_{x \in \mathbb{C}} \{f(x) + D(x, z)\}\), then for any \(y \in \mathbb{C}\), we have \(f(y) + D(y, z) \geq f(x^*) + D(x^*, z) + D(y, x^*)\).

The proofs for Lemma D.2 are very standard and we will omit them in this paper.

Lemma D.3 Let \(a \in (0, 1)\), \(\rho_a = \left[\frac{a}{1-a}\right] + 1\), and \(t_a = \left[\frac{a(\rho_a+1)}{\rho_a + t_a}\right]\). Then, for any \(t \geq \rho_a + t_a\), we have
\[
\sum_{r=\rho_a + t_a}^{t} \frac{a^{-r}}{t} \leq \frac{1}{1-a} \cdot \frac{a^{-t}}{t - \rho_a}.
\]

Proof: We adopt an induction argument with hypothesis \(\sum_{r=\rho_a + t_a}^{t} \frac{a^{-r}}{t} \leq \frac{1}{1-a} \cdot \frac{a^{-t}}{t - \rho_a}\). For the base case, consider \(t = \rho_a + t_a\). We can easily see \(\frac{a^{-\rho_a-t_a}}{\rho_a + t_a} < \frac{1}{1-a} \cdot \frac{a^{-\rho_a-t_a}}{t_a}\). Now assume that the induction hypothesis holds for some \(t \geq \rho_a + t_a\). We will show \(\sum_{r=\rho_a + t_a}^{t+1} \frac{a^{-r}}{t} \leq \frac{1}{1-a} \cdot \frac{a^{-t+1}}{t - \rho_a} \cdot \frac{a^{-t+1}}{t+1} = \frac{a^{-t+1}}{1-a} \cdot \frac{a}{t - \rho_a} + \frac{1-a}{t+1}\).

Furthermore,
\[
\frac{a}{t - \rho_a} + \frac{1-a}{t+1} - \frac{1}{t - \rho_a + 1} = a \left(\frac{1}{t - t - \rho_a + 1}\right) + (1-a) \left(\frac{1}{t+1 - t - \rho_a + 1}\right) = \frac{1}{t - \rho_a + 1} \cdot \frac{t - \rho_a (1-a)}{t - \rho_a + 1} = \frac{1}{t - \rho_a + 1} \cdot \frac{(a - \rho_a)(1-a)}{t - \rho_a + 1} \leq 0,
\]
where (a) follows from \(\rho_a = \left[\frac{a}{1-a}\right] + 1 > \frac{a}{1-a}\) and the fact that \(\frac{(1-a)p_a^2 + a}{\rho_a(1-a) - a} = \rho_a + \frac{a}{1-a} \cdot \frac{(1-a)p_a^2 + a}{\rho_a(1-a) - a} < \rho_a + t_a \leq t\).

Therefore, we can conclude that \(\sum_{r=\rho_a + t_a}^{t+1} \frac{a^{-r}}{t} \leq \frac{1}{1-a} \cdot \frac{a^{-t+1}}{t - \rho_a + 1}\), which is the desired result. \(\square\)