We first thank all reviewers for their thoughtful comments, and we wish everyone health during these hard times.

[R1]: 1. We acknowledge the simplicity in our linear demand and reference price update models. We made these modelling decisions due to two reasons: (i) We intended to focus on the convergence of prices set by agents (who run mirror descent) when they may be using step sizes of different orders; e.g. firms use step size $\mathcal{O}(1/\sqrt{t})$ while nature adopts constant step sizes. We believe that this will serve as a first step to future research on more complex models. (ii) Linear demand and reference price update models are well studied, and have many practical motivations. For example in marketing, reference prices are used to model consumer price expectations for particular products, and such expectations are adjusted based on past prices consumers experience; please see [28] for a comprehensive survey for applications of linear demand models, and [24,45,50] for motivations/empirical validations for the linear reference price model. These references are also discussed in Section 2 of the paper.

2. Regarding practical aspects of gradient feedback, real-world firms typically have a good understanding of price elasticity of demand due to operational experiences. The gradient of revenue can be calculated using estimated elasticity, observed sales (i.e. demand) and prices. That being said, we agree that a more practical setting would be to assume firms obtain noisy gradients or zero-order feedback. Nevertheless, we point out that the main message we hope to convey through this paper is the fact that firms can converge to the SNE under limited information by running simple yet practical OMD algorithms, which we argue is a non-trivial result even when firms have exact gradient feedback due to heterogeneity in step sizes (see Section 5.3). We are definitely interested in generalizing our results to other feedback structures in future work.

3. Assumption 1 is invoked in all theorems and lemmas of Section 5, and we will clearly state this in the revised paper. Regarding footnote 7, consider two different price profiles: $(\boldsymbol{p}, r)$ which is obtained through first order conditions w.r.t. the demand and reference price update models (see proof of Lemma 3.2 in Appendix B); and the SNE $(\boldsymbol{p}^*, r^*)$, which by definition satisfies the best-response property. These profiles are identical iff $(\boldsymbol{p}, r)$ lies in the decision set. In the proof of Lemma 3.2, we show that $(\boldsymbol{p}, r)$ are strictly positive constants given model parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}$ and reference memory parameter a. Footnote 7 intends to say that as long as the decision set lower (upper) bound is small (large) enough while remaining positive, $(\boldsymbol{p}, r)$ will lie in the decision set. This means if firms are willing to consider both prices near zero and those sufficiently large, Assumption 1 holds. Nevertheless, we agree that footnote 7 may cause confusion, and we will modify this in the revised paper.

[R2 & R3]: We indeed believe the simple linear demand and reference update models capture many realistic aspects of consumer behavior and firm competition in various markets; please see response 1. for Reviewer #1. Echoing our expanded literature review in Appendix A, we realized there are not many works in the CS-econ and behavioral economics literature that analyze market dynamics under reference effects from an online learning perspective. We think it is important to consider demand models which not only depend on pricing decisions, but also some time-dependent market characteristic (e.g. consumer expectations and choice behavior). Thus, we view this paper as a first step for analyzing more complex competitions under reference effects, and believe the learning and game theoretic aspects of this work make it a good fit for NeurIPS. We also thank Reviewer #2 for suggestions regarding experiments. We are definitely interested in comparing our theoretical convergence rates to those obtained in practice, despite many practical challenges, e.g. the proprietary nature of data from firms that run similar algorithms. We are certain that this will be valuable in future extensions of this work. Regarding Reviewer #3's comments on the practical aspects of firms running OMD, we point out that OMD is an off-the-shelf algorithm that many firms may consider using in practice due to its simplicity. In this work, we wanted to analyze the impact of running this widely used algorithm in dynamic pricing competitions.

[R4]: 1. We agree that uniqueness of SNE can be proved using variational inequalities (VI). However, we would like to point out that uniqueness, as well as proving uniqueness, is not the focal point of our paper. Our main focus is to understand whether multiple firms running OMD would imply convergence to the unique SNE under varying reference prices. In fact, our proof for showing uniqueness under Assumption 1 (Lemma 3.2) is also very straightforward. Moreover, we point out the differences between an Nash Equilibrium (NE) w.r.t. a fixed reference price, and the SNE. Fixing a reference price, the two firms admit a unique NE, and VI indeed implies OMD convergence to this particular NE if the reference price remains fixed. However, reference prices vary according to firms' decisions, so the respective NE changes as firms progress. Hence, the 3-agent system may diverge and oscillate (Figure 1 (c) gives such an example). Mathematically, in order to use VI to show the update $x \leftarrow x - \textit{step} * (Jx + b)$ implies convergence, *step* should be identical across all 3 agents (two firms and nature). In our setting, nature takes constant step sizes, while firms may take decreasing step sizes, so VI is not directly applicable. Furthermore, heterogeneous step sizes better model reality, as firms are unaware of how reference prices update and run OMD independently. On a separate note, we agree that it is beneficial to add discussions that our model admits a diagonally dominant Jacobian, and will revise the paper accordingly.

2. As discussed in the introduction and Section 5.3, we are surprised that previous results for multi-agent OMD convergence (e.g. [9, 33, 41]), whose methodologies rely on variational inequalities, cannot be applied in our setting, primarily because agents (firms and nature) may be using step sizes of different orders. The fact that convergence still occurs under such heterogeneity is not intuitive, and we find this phenomena to be interesting. Finally, to the best of our knowledge, our proof techniques for convergence, as well as the perspective of modeling reference prices as decisions by nature running OMD, are novel.

3. Thank you for allowing us to clarify the sentence *"...... games are stateless and all agents are required to take decreasing step sizes"*. When saying the 2-firm game has a "state", we mean agents' utility functions are not only functions of their current decisions, but also of the time-dependent reference price. We can view our utility functions of interest being parameterized by varying reference prices. The two works Reviewer #4 pointed out consider each agent optimizing a fixed utility function over time, and these functions only depend on agents' current decisions. This means the gradient or whatever feedback a firm obtains after making a decision is with respect to these fixed functions. However, in our case, the gradient feedback can be considered as associated with "different functions" which are parameterized by each period's reference price. This technicality makes previous algorithms and convergence results break down in our setting. When we state "all agents are required to take decreasing step sizes", we were particularly referring to results presented in [9, 33, 41]. Nevertheless, we agree that the sentence pointed out as well as similar sentences are rather inaccurate, and we will modify the way we explain this in the revised version of the paper.