We thank the reviewers for their thoughtful suggestions which we'll incorporate to substantially improve our final paper.

Experiments. It is not easy to integrate bandit systems without extensive industry resources, particularly in a sensitive area like pricing. Nearly all past research on general bandits/pricing methodologies relied on simulation experiments (many papers just provide regret bounds without experiments). No matter which experiments we run, the sensitive nature of pricing necessitates provable guarantees, which is a major strength of our adversarial regret bounds. We'll emphasize our robustness analysis under misspecified d (Fig C.1) and a totally-misspecified demand-model (Fig C.2).

R1: Seems odd that the regret can be decreasing in time (which happens for figures C, D, E and F). How is this possible?

L267 states: "regret of the bandit algorithms decreases over time, indicating they begin to outperform the optimal fixed price chosen in hindsight". We'll clarify: Since our bandits can vary price over time and these environments are nonstationary, our algorithms are able to outperform *any* single fixed price-configuration (what regret is defined against). This limitation of the standard regret definition has led to alternative dynamic-regret formulations such as those of [RS2013, Z2017], although dynamic pricing literature typically measures regret against a single price as we've done.

[RS2013] Rakhlin, Sridharan. "Online Learning with Predictable Sequences" [Z2017] Zhang et al. "Improved Dynamic Regret for Non-degenerate Functions"

R1: How does FindPrice (from 0 or p_{t-1}) influence regret? GDG/OPOK/OPOL curves suspiciously similar for $d = N$.

We'll clarify: Even comparing GDG vs. itself would result in a (small statistically insignificant) visible difference in curves as bandits are internally stochastic (cf. OPOK Step 4). If $d = N$: GDG & OPOK are nearly mathematically equivalent (same regret bound, but their empirical regret is not identical for the aforementioned reason). Minor difference is action-noising step (ie. Step 5 of OPOK): ξ_t is applied in the p-space for GDG and to x for OPOK. Since $d = N \Rightarrow U$ is $N \times N$ and orthogonal, this makes no theoretical difference for rotationally-invariant uniform ξ_t. FindPrice makes no difference here because $d = N \Rightarrow U$ is invertible. When $d = N$: OPOL and OPOK are also nearly equivalent, because $\widehat{U}$ is also orthogonal $N \times N$ matrix. We'll clarify that either choice of FindPrice (from 0 or p_{t-1}) obeys our paper's regret bound (we do not find statistically-significant difference in our experiments). Choice should be based on the seller's philosophy (p_{t-1}: less dramatic price changes, 0: more stability around default price=0).

R1: I suspect chosen parameters in experiments make task too easy (influence of prices p in the demand seems marginal)

We'll clarify: If parameters were chosen to make problem too easy, our figures wouldn't depict such a statistically significant difference in different methods' regrets. We chose z, V, λ to encourage three properties underlying real-world demand curves: different products' baseline demands & demand elasticities should be highly diverse (wide range of z), prices should highly influence demands such that price-increases should severely decrease demand and affect demand for the same product more than other products ($\lambda = 10$ reflects this far better than 0 and leads to V having far bigger values than suggested by $N(0, 2)$). The optimal p^* in a stationary environment has $||p^*||_2 \approx 8$, whereas p^* would instead lie somewhere near $\mathcal{S}$-boundary ($||p^*||_2 = 20$) if price didn't substantially influence demand. We did initially set our noise variance =1 as suggested, but wanted to explore noisier settings (ie. harder problems) and found methods could handle 10 without noticeable performance degradation. Results from main text look very similar under variance=1, we'll add them to supplement. In existing supplement experiments, we already use variance =1 (see L584).

Relationship with existing work. We'll clarify: the main aspect of model (1) that is similar to cited work is the assumption of a linear demand/price relationship[1]. Existing work on dynamic pricing is unrealistic as it does not consider multiple products & nonstationary demand curves. Our work is novel because it can handle these cases and obtains even superior performance guarantees when an additional low-rank assumption holds. We do not claim the low-rank assumption is justified by existing pricing work, and instead will cite work on e-commerce recommendations, where low-rank product feature decompositions are a standard assumption that practically works [S2017, Z2016].

[S2017] Sen et al. "Contextual Bandits with Latent Confounders: An NMF Approach". [Z2016] Zhao et al. "Predictive Collaborative Filtering with Side Information".

Assumptions. Fig C.1-C.2 show our methods work well even if our assumptions are wrong. We'll include extra experiment on real demand data[2] for 1340 products sold by Grupo Bimbo over 7 weeks. We form a matrix $\mathbf{Q}$ of the total weekly demands for each product across all stores. The SVD of $\mathbf{Q}$ reveals the following percentages of variation in the observed demands are captured by top k singular vectors: $k = 1 : 97.1\%, k = 2 : 99.1\%, k = 3 : 99.9\%$, thus suggesting empirical validity of our low-rank assumption on the demand variation.

(A4) is not a strong assumption: up to scaling factors, the orthogonality condition on U does not actually really restrict the family of demand curves that can be captured via our low-rank unknown-features model (there is much flexibility by changing V_t). See also Theorem A.2 in the supplement for alternative (more general) assumptions in case of known features. As stated in L60 & Appendix D, C denotes universal constants. We'll clarify C is problem-independent and does not depend on T, d, r (our usage of C is equivalent to big O notation commonly used to present regret bounds).

[1] Historical demand data often nicely fit linear relationship, cf. Houthakker and Taylor (1966) or `towardsdatascience.com/calculating-price-elasticity-of-demand-statistical-modeling-with-python-6adb2fa7824d`

[2] `www.kaggle.com/c/grupo-bimbo-inventory-demand/`