Conjectural Variations in Competitive Dynamic Pricing: A Learning Foundation via Experimentation Design and Feedback Structure

Introduction. The digitization of commerce, the growth of online marketplaces, and rapid improvements in technology have made it feasible to adjust prices at high frequency and to measure demand responses at scale, accelerating the use of experimentation and algorithmic pricing in practice. Recent papers document large-scale randomized pricing interventions, including discounts and promotions, e.g., a randomized field experiment on Alibaba involving more than 100 million customers Zhang et al. (2020), as well as a pricing meta-experiment on Airbnb Holtz et al. (2025). In addition to large-scale experimentation by platforms themselves, a substantial number of third-party sellers deploy algorithmic pricing (e.g., Calzolari and Hanspach (2025)).

These developments have motivated a growing operations literature on learning demand through price experimentation under practical constraints, but most theoretical work studies a single seller in isolation rather than in competitive environments (e.g., Besbes and Zeevi (2015)). Meanwhile, a growing computational literature documents that competitive learning can generate supra-competitive outcomes, yet theoretical results characterizing the long-run outcomes of practical pricing algorithms in competitive markets remain limited.

This paper takes a step toward bridging this gap by studying a simple dynamic pricing algorithm in a general framework that accommodates a broad class of demand systems, experimental designs, and feedback structures. We prove that the learning dynamics converge to a static equilibrium and show that the limiting equilibrium is pinned down by the learning bias induced by correlated experimentation and the market feedback structure.

Problem Setting. Specifically, we study a repeated dynamic pricing game with multiple sellers, where demand depends on the full vector of sellers' prices and is subject to random shocks. Sellers observe their own prices and realized demands, but they may differ in the competitor-price information available to them. In particular, each seller may observe the prices of only a subset of rivals, and sellers do not observe rivals' demands or know the true demand system. Thus, our framework accommodates heterogeneous and potentially asymmetric information structures, ranging from bandit feedback to full competitor-price observability.

Given this information environment, we model sellers as learning through local price experimentation around a current focal price. Experimentation is organized in batches: during a batch, each seller repeatedly perturbs her focal price, observes demand and the rival prices available under her feedback structure. At the end of the batch, she uses simple Ordinary Least Squares (OLS) to estimate demand as a function of her own price and the rival prices she observes, and updates her focal price toward the estimated revenue-maximizing price.

Crucially, experimentation may be statistically dependent across sellers even without explicit coordination. Such statistical dependence can arise endogenously from the structural features of modern marketplaces. Platform-imposed API throttling, reliance on common third-party pricing services, shared responses to market-wide demand shocks, and platform-wide promotional events such as ‘Black Friday or ‘11-11’ can all lead to cross-player experimentation correlation.

Main Results. Our analysis connects these learning dynamics to the classic Conjectural Variations (CV) equilibrium. In a CV equilibrium, each seller optimizes given current rival prices and conjectures about how those prices locally co-move with her own, summarized by a matrix A. The entry A_{ij} represents seller i's conjecture that a one-dollar increase in her own price is accompanied by a change of A_{ij} dollars in seller j's price. Particularly, $A_{ij} = 0$ implies no expected co-movement. When all off-diagonal entries are zero, sellers treat rival prices as locally fixed, and the CV equilibrium reduces to the Nash equilibrium.

Our main theoretical result shows that, under certain stationary conditions on demand, when each seller follows the partial-feedback pricing procedure described above, the resulting focal price sequence converges to a CV equilibrium, even though sellers do not know the true demand system and do not explicitly reason about competitors' reactions. This convergence is not restricted to linear demand and accommodates nonlinear systems such as multinomial logit demand. Notably, the limiting conjecture matrix A is not imposed as an exogenous behavioral primitive; instead, it is endogenously determined by the interaction between the feedback structure and the correlation structure of experimentation.

When sellers' experiments are correlated, a seller's price changes may systematically co-occur with changes in unobserved rival prices. Because these unobserved rival prices are omitted from the seller's demand regression, their effects are partially absorbed into the coefficients on the included regressors, including her own price, thereby biasing the learned demand response. We show that, as batch sizes grow and perturbations shrink, this bias induces a specific conjecture matrix. Thus, conjectural variations emerge endogenously from the learning process, even though sellers do not explicitly form conjectures about unobserved rivals.

This result bridges the gap between an abstract “mental” solution concept and a realizable outcome of learning algorithms. It shows that correlated experimentation, partial feedback, and even the same statistical dependence that generates bias in demand learning can all determine the conjectures that characterize the static equilibrium approached by the learning dynamics, providing a learning foundation for conjectural variations. Interestingly, when the bias vanishes, as under full price feedback or asymptotically uncorrelated experimentation by unobserved rivals, the induced conjectures are zero and the dynamics instead converge to the Nash equilibrium. We further show that, with appropriately chosen perturbation scales and batch sizes, the mean squared price error decays on the order of $T^{-1/2}$ up to logarithmic factors. This is the same order as the optimal mean-squared rate established for related Nash-convergence bandit-feedback games Ba et al. (2025).

Turning to the implications of the CV mechanism, we analyze how information and experimental design affect equilibrium outcomes through the conjectures they induce. Under strategic complementarity, CV equilibrium prices increase with conjectures, so positively correlated experimentation under bandit feedback can induce prices above the Nash benchmark. Under the broader partial feedback, however, conjectures are partial projection coefficients whose signs depend on which rivals prices are observed; for example, when smaller sellers observe a major seller, positive pairwise experimentation correlations can induce negative conjectures and prices below Nash. More generally, observing additional rivals or changing experimentation correlations has market-specific and potentially non-monotonic effects on individual sellers' revenues.

Notably, the CV mechanism is not an artifact of the OLS used by our algorithm; it arises from the demand relationship a seller can learn from her observed data. We extend our analysis to the case when the sufficient stability conditions for convergence fail. We find that the same conjecture matrix continues to govern the direction of the limiting price adjustments. When some rival prices are unobserved, this learned demand relationship averages over those omitted prices conditional on the observed price variation. As a result, other learning rules that estimate marginal demand from the observed experimental variation can inherit the same conjectural-variation component. Thus, more generally, the feedback and experimentation structures determine the price dynamics followed by the learning process, not only the CV equilibrium reached under the convergence conditions.

Taken together, our results identify feedback and experimentation as joint determinants of equilibrium selection. With unobserved rival prices, correlated experimentation induces an endogenous conjecture matrix that can shift the market from Nash to a CV equilibrium. Under strategic complementarities, positive conjectures raise prices above Nash without communication or coordination, whereas zero or negative conjectures can produce Nash or below-Nash outcomes. Thus, feedback and experimentation shape both demand learning and the equilibrium reached by decentralized pricing algorithms.

References

Ba, W., Lin, T., Zhang, J., and Zhou, Z. (2025). Doubly optimal no-regret online learning in strongly monotone games with bandit feedback. *Operations Research*.

Besbes, O. and Zeevi, A. (2015). On the (Surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Science*, 61(4):723–739.

Calzolari, G. and Hanspach, P. (2025). Pricing algorithms out of the box: a study of the repricing industry. *Journal of Competition Law & Economics*, 21(2):163–187.

Holtz, D., Lobel, F., Lobel, R., Liskovich, I., and Aral, S. (2025). Reducing interference bias in online marketplace experiments using cluster randomization: Evidence from a pricing meta-experiment on airbnb. *Management Science*, 71(1):390–406.

Zhang, D. J., Dai, H., Dong, L., Qi, F., Zhang, N., Liu, X., Liu, Z., and Yang, J. (2020). The long-term and spillover effects of price promotions on retailing platforms: Evidence from a large randomized experiment on alibaba. *Management Science*, 66(6):2589–2609.