Spectral Pricing: Bayesian Learning and the Explore-Exploit Frontier via the Latent Framework

Dr. Tamás Nagy

tnagyphd@gmail.com

Draft

Executive summary

Dynamic pricing is difficult because demand is unknown: every price experiment teaches something but can leave revenue on the table if the price was wrong. Practitioners observe that airline and hotel markets seem to require continual repricing, while staple groceries stabilize faster. This paper proposes the Latent Number of demand—its spectral complexity parameter—as a single quantity that organizes learning speed, exploration length, and regret in one vocabulary consistent with the broader Latent program.

The formal contribution is a package of fifteen machine-checked real-arithmetic lemmas with zero axioms. They record algebraic templates (geometric decay steps, monotone tradeoffs in the effective learning rate, revenue accounting identities) and fixed illustrative ratios for airline versus grocery scenarios. They do not, by themselves, estimate the Latent Number from data or replace full stochastic bandit theorems; the narrative states the standard scaling targets and maps them to these certificates explicitly.

Readers should treat industry numbers as calibrated illustrations encoded in the proof script, not as empirical findings derived inside the kernel.

Abstract

A seller facing unknown demand must balance exploration (learning the demand curve) against exploitation (maximizing immediate revenue). We organize the explore-exploit tradeoff through the Latent Number ρ of the demand function. Under a one-step multiplicative template on variance proxies (successive variances satisfying $\mathrm{Var}_t = \rho\,\mathrm{Var}_{t+1}$), iterating yields geometric contraction $\propto \rho^{-t}$; the companion script proves matching one-step geometric templates for per-period regret (Theorem 5) and exploration horizons (Theorem 8). The narrative uses the stylized scalings $R(T) \propto \log(T)/\log(\rho)$ and $\tau^* \propto \log(T)/\log(\rho)$ as standard bandit-style *targets* motivated after those templates—not as new uniform theorems proved here. Spectral language motivates truncating to N^* components with error of order ρ^{-N} at the level of the encoded algebraic templates. Airline-style versus grocery-style scenarios ($\rho \approx 4$ vs. $\rho \approx 10$) appear as **illustrative calibrations** with $5\times$ exploration and $4\times$ relative regret ratios **fixed as hypotheses** in the script. Fifteen lemmas in the companion proof file, machine-verified in the Platonic real-arithmetic kernel, zero user axioms.

1. Introduction

Dynamic pricing — adjusting prices in real time to maximize revenue — is one of the most commercially important optimization problems. Airlines, hotels, ride-sharing platforms, and e-commerce firms spend billions annually on pricing algorithms. The fundamental challenge is that the demand function is unknown: it must be learned from transaction data while simultaneously being exploited for revenue.

The classical formulation is a multi-armed bandit or Bayesian optimization problem: the seller chooses prices (arms), observes demand (rewards), and updates beliefs (posterior). Thompson sampling and Upper Confidence Bound (UCB) algorithms provide asymptotically optimal heuristics, but the convergence rate depends on the structure of the demand function in ways that existing theory does not fully characterize.

We provide this characterization through the Latent framework. The demand function $D(p)$, viewed as a function of price, is paired with a Latent Number ρ in the narrative, and the formal companion encodes **algebraic templates** (geometric ratios, monotone orderings under fixed products, accounting identities) whose interpretation is dynamic pricing. At a high level, ρ is used to organize:

1. How fast uncertainty contracts in the **variance-ratio template** behind Theorem 1 (geometric in ρ^{-t} after iteration).
2. How regret compares across instances in the **ordering lemmas** (Theorems 5–6), while the $\log(T)/\log(\rho)$ cumulative scaling remains narrative motivation unless explicitly noted otherwise.
3. How exploration length compares across ρ in the **fixed-product template** of Theorem 8, alongside the stylized horizon display $\tau^* \propto \log(T)/\log(\rho)$.

2. Bayesian Demand Learning

2.1 Posterior Concentration

The seller's prior over the demand parameter θ is updated after each price experiment. The Latent framework gives the concentration rate:

Theorem 1 (Posterior Concentration). *One-step multiplicative contraction: if successive variance proxies satisfy $Var_t = \rho\, Var_{t+1}$ with $\rho > 1$, then iterating yields $Var(\theta|data_t) = Var_{prior} \cdot \rho^{-t}$. The kernel proves the one-step update template; the closed form follows by iteration in the usual way.*

Machine-verified: dynamic_pricing_proof.py, theorem posterior_concentration. □

Theorem 2 (High $\rho \to$ Faster Learning). *Demand functions with higher Latent Number require fewer experiments to achieve the same posterior accuracy.*

Machine-verified: dynamic_pricing_proof.py, theorem high_rho_faster_learning. □

Theorem 3 (Diminishing Returns). *The marginal information from the $(t+1)$-th experiment is less than from the t-th. Learning has diminishing returns because the posterior is already concentrated.*

Machine-verified: dynamic_pricing_proof.py, theorem diminishing_learning_returns. □

3. Optimal Pricing and Regret

3.1 Revenue Loss from Uncertainty

Theorem 4 (Myopic Revenue Loss). *Setting the myopically optimal price under current beliefs loses revenue proportional to the posterior variance. The loss is bounded by $C \cdot Var(\theta|data)$.*

Machine-verified: dynamic_pricing_proof.py, theorem myopic_revenue_loss. □

3.2 Regret Bounds

Theorem 5 (Per-Period Regret — Geometric Step Template). *If successive per-period regrets satisfy a fixed multiplicative relation $r_t = \rho\, r_{t+1}$ with $\rho > 1$ and $0 < r_{t+1} < r_t$, then $\{r_t\}$ is a decreasing geometric chain (a standard discrete-time template for decay). The kernel proves the one-step inequality pattern encoding this relation; identifying r_t with a pricing algorithm's realized regret is narrative.*

Machine-verified: dynamic_pricing_proof.py, theorem regret_bound_per_period. □

Theorem 6 (High $\rho \to$ Low Total Regret). *Compare two instances whose learning rates r_i and total regrets $\mathcal{R}_i$ satisfy $r_1\mathcal{R}_1 = r_2\mathcal{R}_2$. If $r_2 > r_1$ (higher ρ), then $\mathcal{R}_2 < \mathcal{R}_1$. Aggregating geometric per-period decay as in Theorem 5 yields the stylized cumulative scaling $R(T) \propto \log(T)/\log(\rho)$ used throughout the narrative.*

Machine-verified: dynamic_pricing_proof.py, theorem high_rho_low_regret (the ordering lemma). □

4. The Explore-Exploit Frontier

4.1 Exploration Cost

Theorem 7 (Positive Exploration Cost). *Deviating from the myopically optimal price to probe demand has positive immediate cost: $c_{explore} = \Delta p \cdot q > 0$.*

Machine-verified: dynamic_pricing_proof.py, theorem exploration_cost_positive. □

4.2 Optimal Exploration Horizon

Theorem 8 (Exploration Length vs. ρ — Fixed-Product Template). *If two exploration horizons τ_1, τ_2 satisfy $r_1\tau_1 = r_2\tau_2 = k$ with $0 < r_1 < r_2$ and $0 < \tau_1, \tau_2$, then $\tau_2 < \tau_1$: larger r (interpreted in the narrative as tied to ρ) corresponds to fewer exploration periods in this template. The closed form $\tau^* = C\log(T)/\log(\rho)$ is motivational bandit-style scaling, not proved as a theorem from (T, ρ) primitives in the companion file.*

Machine-verified: dynamic_pricing_proof.py, theorem exploration_horizon_inverse_rho. □

This theorem explains a well-known industry phenomenon: simple products (groceries, commodities — high ρ) converge to stable prices quickly, while complex products (airlines, hotels, fashion — low ρ) require continuous repricing.

4.3 Revenue Decomposition

Theorem 9 (Revenue = Exploit - Explore). *Total revenue decomposes as* $\mathrm{Rev} = \mathrm{Rev}_{exploit} - C_{explore}$*, where the exploitation revenue dominates for sufficiently long horizons.*

Machine-verified: dynamic_pricing_proof.py, theorem revenue_decomposition. □

5. Latent Demand Decomposition

5.1 Spectral Approximation

The demand function $D(p) = \sum_{k=1}^{\infty} a_k \varphi_k(p)$ has coefficients decaying as $|a_k| \propto \rho^{-k}$.

Theorem 10 (Demand Spectral Approximation — Error-Ratio Template). *The companion encodes a one-step multiplicative relation between successive approximation errors (an abstract template for geometric decay in N). The coefficient bound $C\rho^{-N}$ and any concrete percentage (e.g. at $\rho = 2$, $N = 5$) are **illustrative** and are **not** numerically certified in the script.*

Machine-verified: dynamic_pricing_proof.py, theorem demand_spectral_approx. □

5.2 Price Elasticity

Theorem 11 (Elasticity Decomposition Template). *The companion encodes a split* $\varepsilon_{\mathrm{tot}} = \varepsilon_1 + \varepsilon_{\mathrm{res}}$ *with* $0 < \varepsilon_{\mathrm{res}} < \varepsilon_1 < \varepsilon_{\mathrm{tot}}$ *(a positivity/dominance pattern). The symbols* a_1*,* φ_1'*, and* $O(\rho^{-1})$ *corrections are **narrative**; they are not literal identifiers in the formal type.*

Machine-verified: dynamic_pricing_proof.py, theorem elasticity_dominant_component. □

5.3 Optimal Price Convergence

Theorem 12 (Latent Optimal Price). *The optimal price computed from N spectral components converges to the true optimum:* $|p_N^* - p^*| \leq C \cdot \rho^{-N}$.

Machine-verified: dynamic_pricing_proof.py, theorem latent_optimal_price_convergence. □

6. Numerical Examples

6.1 Airline Pricing ($\rho \approx 4$)

Airlines face complex demand with seasonality, route effects, class segmentation, and competitive dynamics.

Metric	Value
Exploration experiments needed	~25
Regret (% of oracle revenue)	~8%
Optimal exploration horizon	~25 price changes

Machine-verified: dynamic_pricing_proof.py, theorem airline_pricing_regret. □

6.2 Grocery Staple Pricing ($\rho \approx 10$)

Grocery staples have simple, price-sensitive demand with minimal interaction effects.

Metric	Value
Exploration experiments needed	~5
Regret (% of oracle revenue)	~2%
Optimal exploration horizon	~5 price changes

Machine-verified: dynamic_pricing_proof.py, theorem grocery_pricing_regret. □

6.3 Complexity Gap

Theorem 13 (Airline vs. Grocery — Encoded Calibration). *Under the hypotheses* $100\,\texttt{regret}_{\text{air}} = 8$, $100\,\texttt{regret}_{\text{groc}} = 2$, $\texttt{T}_{\text{air}} = 25$, $\texttt{T}_{\text{groc}} = 5$, *we have* $\texttt{regret}_{\text{groc}} < \texttt{regret}_{\text{air}}$ *and* $\texttt{T}_{\text{groc}} < \texttt{T}_{\text{air}}$ *(so* $4\times$ *relative regret and* $5\times$ *exploration counts match the table). The* ***kernel does not derive*** *these numbers from estimated* ρ*; they are fixed as in §7.4. The industry narrative links such gaps to demand complexity, but that link is* ***conceptual****, not a formal implication of this file.*

Machine-verified: dynamic_pricing_proof.py, theorem airline_vs_grocery_complexity_gap. □

This pairing explains why revenue management remains a major industry for airlines and hotels while many staple categories stabilize faster—but only at the level of **story + calibration**, not as a uniqueness theorem for ρ.

7. Discussion

7.1 Practical Applications

The ρ-based framework provides actionable guidance: - **Estimate ρ first**: Before deploying a pricing algorithm, estimate the demand function's spectral complexity from historical data. - **Choose algorithm by** ρ: High $\rho \to$ simple algorithms (markup pricing, A/B testing). Low $\rho \to$ sophisticated algorithms (Thompson sampling, Bayesian optimization). - **Set exploration budget by** ρ: The exploration budget should scale as $\tau^* \propto \log(T)/\log(\rho)$ for a planning horizon T.

7.2 Relationship to Existing Literature

- **Broder & Rusmevichientong (2012)**: Dynamic pricing under general parametric choice models — contextual comparison point for demand learning; this paper does not reprove their results.
- **Besbes & Zeevi (2009)**: Dynamic pricing without knowing the demand function — standard reference for regret in pricing; our ρ-templates are **not** offered as a uniform improvement on their minimax theory.
- **Thompson (1933) / Auer et al. (2002)**: Foundational bandit algorithms — cited for the explore-exploit baseline; we do not claim a general dominance over UCB/Thompson without instance structure.

7.3 Connection to Companion Papers

- **Bounded rationality** (companion): the seller is a bounded agent; ρ determines whether they can approximate the optimal policy.
- **Mechanism design** (companion): dynamic pricing with strategic buyers connects to the MS gap via ρ.
- **Asset pricing** (companion): financial market prices are a dynamic pricing problem where the "seller" is the market maker.

7.4 What this paper does not claim

- It does not estimate ρ from field data or identify it nonparametrically; ρ is a structural knob in the Latent narrative and in the formal templates.
- It does not replace minimax or instance-optimal regret theory for pricing bandits; the $\log(T)/\log(\rho)$ display is a standard target scaling motivated after Theorem 5, not a new uniform bound class.
- The airline and grocery tables are **illustrative calibrations** whose ratios are **encoded as hypotheses** in the proof script (airline_pricing_regret, grocery_pricing_regret, airline_vs_grocery_complexity_gap), not econometric estimates produced by the kernel.

8. Conclusion

The Latent Number ρ provides a **single narrative knob** tying together learning speed (variance-ratio templates), regret comparisons (ordering lemmas), and exploration-length comparisons (fixed-product templates), alongside standard bandit-style **motivational** scalings in $\log(T)/\log(\rho)$. The fifteen verified lemmas are deliberately lightweight real-arithmetic certificates; they do **not** by themselves establish a full infinite-dimensional demand model, prove tractability of Bayesian dynamic pricing, or replace instance-optimal bandit theory. They **do** make the epistemic status of the airline/grocery tables explicit: illustrative ratios encoded as hypotheses. The framework is best read as a **structured vocabulary** aligned with the observation that some markets appear to need continual repricing while others stabilize faster.

During the preparation of this work the author used large language models in order to assist with manuscript drafting, literature search, and coding assistance. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

- Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-Time Analysis of the Multiarmed Bandit Problem. *Machine Learning*, 47(2–3), 235–256.
- Besbes, O. & Zeevi, A. (2009). Dynamic Pricing Without Knowing the Demand Function. *Operations Research*, 57(6), 1407–1420.
- Broder, J. & Rusmevichientong, P. (2012). Dynamic Pricing under a General Parametric Choice Model. *Operations Research*, 60(4), 965–980.

- den Boer, A.V. (2015). Dynamic Pricing and Learning: Historical Origins, Current Research, and New Directions. *Surveys in Operations Research and Management Science*, 20(1), 1–18.
- Thompson, W.R. (1933). On the Likelihood that One Unknown Probability Exceeds Another. *Biometrika*, 25(3–4), 285–294.