No-Free-Anchoring: Platform-Designed Outcome-Weighted Reference Prices and Near-Stationary Optimal Pricing

Liz Lemma Future Detective

January 16, 2026

Abstract

Modern platforms increasingly display "usual price / was price" statistics that shape demand through reference effects. The recent ARM model (Agrawal & Tang, 2024) shows long-memory references can make fixed pricing highly suboptimal and can incentivize high initial prices followed by markdowns—strategies that may resemble deceptive reference pricing. We study a platform-mediated alternative: the displayed reference is an outcome-weighted (sales- or engagement-weighted) average of transacted prices rather than an unweighted average of posted prices. We formalize this as a dynamic mechanism where the platform commits to an update rule and the seller best-responds. Our first result is a "no-free-anchoring" identity: the next reference moves toward today's posted price in proportion to realized sales weight; if sales are negligible, the reference barely moves (and does not move at all if sales are zero). This directly blocks the pathological incentive to post extreme prices solely to inflate future references. We then analyze a tractable linear-demand model with symmetric reference effects and show that dynamic incentives become small once the cumulative sales weight grows, implying near-stationarity of optimal pricing and near-optimality of a fixed price up to polylogarithmic additive loss. We discuss platform design tradeoffs (manipulation-resistance vs reference volatility) and extensions where asymmetric reference effects or nonlinear demand require numerical dynamic programming.

Table of Contents

update rule with outcome weights; define ‘anchoring’ and ‘manipulation’ metrics.

3. 3. A No-Free-Anchoring Identity: exact formula for reference movement under outcome weighting; implications for fake discounts and extreme posted prices.
4. 4. Seller Best Response as a Dynamic Program: Bellman equation in state (R_t, W_t, t); structural properties (monotonicity, bounded influence) under general D and w.
5. 5. Tractable Linear-Quadratic Case: closed-form myopic pricing, bounds on dynamic marginal value of R, and near-stationarity; fixed-price near-optimality up to polylog(T).
6. 6. Platform Design: choosing $w(\cdot)$ (e.g., $w(q) = q^\gamma$) to trade off manipulation-resistance and reference volatility; optional Stackelberg formulation with welfare/deception constraints.
7. 7. Extensions and Limits: asymmetric reference effects ($\eta_+ \neq \eta_-$), censored demand, stochastic shocks, and partial observability; where numerical methods are needed.
8. 8. Discussion and Policy Implications: guidance for platforms/regulators on reference-price computation; testable predictions for auditing and enforcement.

1 Introduction: reference-price displays, deceptive anchoring, and the role of platform-mediated rules

By 2026, reference-price displays have become a routine component of online retail interfaces: a current posted price is juxtaposed with a "was" price, a "typical" price, or a "recommended" benchmark that implicitly frames the attractiveness of the offer.[1] This design choice is not innocuous. A large body of evidence in behavioral IO and marketing suggests that consumer willingness to buy depends not only on the level of the current price, but also on how that price compares to a displayed reference. In practice, the platform—rather than the seller—often determines how that reference is computed and presented (e.g., "lowest price in last 30 days," "median price," "recent average," or proprietary "deal" metrics). The central premise of this paper is that *how* the platform computes the reference statistic is itself a policy lever: it can either facilitate deceptive anchoring strategies or discipline them by tying reference movements to realized economic activity.

The motivating concern is a familiar one. Sellers may attempt to create the appearance of a discount by inflating a reference benchmark and then posting a lower "sale" price relative to that benchmark, even when the sale price is not meaningfully low in any economic sense. In the starkest form, a seller posts an extreme price briefly, generates little or no demand at that price, yet mechanically raises the platform's displayed reference for subsequent periods. The seller then posts a moderate price and benefits from increased conversion due to the now-elevated reference point. When such strategies are profitable, reference-price displays can drift away from informative summary statistics and become instruments of persuasion. This tension is now sufficiently salient that platforms face pressure from regulators and consumer advocates to ensure that reference-price claims are not misleading, while sellers simultaneously seek flexibility in promotional pricing and merchandising.

To see why the computation rule matters, it is useful to contrast two broad classes of reference updating that appear implicitly or explicitly in platform practice. A first class consists of *posted-price based* rules, where the platform aggregates posted prices regardless of whether they generated sales. Two common variants are (i) a simple average over recent posted prices, which we refer to as an averaging rule (ARM), and (ii) exponentially weighted smoothing over posted prices (ESM), where recent posted prices receive more weight but even short-lived, low-traffic postings can shift the displayed reference. Such rules may look reasonable if one thinks of the reference merely as a summary of the seller's announced prices. However, they create a mechanical vulnerability: because the reference reacts to posted

[1] We use "reference price" in the behavioral sense: a salient benchmark against which the current price is evaluated, not necessarily a legally defined MSRP.

prices even when those prices are not transacted, a seller can "buy" a higher reference cheaply by posting high prices in periods with low demand, low visibility, or deliberately curtailed inventory. In other words, under purely posted-price aggregation, the platform unintentionally subsidizes anchoring.

A second class of rules, which we advocate and analyze, are *outcome-weighted* rules: the reference statistic reacts primarily to transacted prices (or more generally, realized outcomes such as sales or engagement). The basic intuition is straightforward and precedes any formalism. If a platform's reference display is meant to summarize what consumers actually pay (or what the market clears at), then a period in which a seller posts an extreme price but makes no sales should not move the benchmark. Conversely, if a seller truly sells substantial quantity at a higher price, then it is less problematic—and arguably informative—for the reference to adjust upward, because the higher price reflects real transactions rather than a cheap, purely nominal signal.

This shift in perspective recasts the platform's design problem. The question is not whether reference dependence exists; we take it as a behavioral demand feature that sellers and platforms must navigate. The question is whether the platform can *commit* to a reference computation that preserves the informational content of the display while attenuating manipulative incentives. Outcome-weighted updating does precisely this by imposing a discipline: to move the future reference, the seller must generate outcomes at the price that is intended to become the anchor. When demand is downward sloping, selling more at a higher price is costly, so anchoring becomes endogenously expensive rather than mechanically free.

We emphasize that this discipline is distinct from, and complementary to, traditional policy responses. One approach is disclosure or labeling: requiring that a "was" price reflect a minimum duration, a minimum sales volume, or a verifiable historical price. Another is ex post enforcement against deceptive claims. Our approach is architectural: embed the relevant discipline directly into the platform's aggregation rule so that the displayed statistic is mechanically insensitive to price postings that do not generate transactions. Importantly, this does not require the platform to infer seller intent or police every promotion. It requires only that the reference computation condition on realized outcomes—information the platform typically already observes.

Of course, outcome-weighting is not a panacea; it changes the nature of reference dynamics rather than eliminating dynamics altogether. In particular, it shifts attention to a different set of edge cases. High-volume promotion events (e.g., major holidays, influencer campaigns, or platform-driven traffic shocks) may place substantial weight on a short interval of transactions, potentially producing noticeable jumps in the reference. From the platform's perspective, this is a genuine tradeoff: rules that strongly downweight low-volume periods are more manipulation-resistant, but they may amplify the influence of high-volume bursts on the reference. A theme of this paper is that such tradeoffs can be made explicit and analyzable once we model the reference as an outcome-weighted moving statistic.

Our framework also clarifies a subtle incentive issue for sellers. Even when the platform uses outcome-weighted updating, a forward-looking seller might still attempt to invest in raising future references, because higher future references can increase demand at any given future price if consumers exhibit reference dependence. However, the crucial difference is that this "investment" must be made through actual sales at the chosen price, not through costless posted-price manipulation. As a result, dynamic incentives to manipulate the reference are naturally limited by demand curvature and by the platform's accumulated history (the mass of past outcomes already embedded in the reference). This logic suggests a practical implication: platforms can make references harder to game by increasing the effective historical weight (e.g., longer lookback windows or larger baseline mass) and by choosing outcome-weight functions that are relatively insensitive to very small quantities.

We organize these ideas in a parsimonious dynamic model that places the platform's reference update rule at the center. The seller chooses a posted price each period, consumers are myopic but reference-dependent, and the platform updates a displayed reference statistic using a rule that averages prices with weights determined by realized sales (or engagement). While the formal analysis appears in the subsequent sections, the key conceptual objects are already visible in the introduction: the posted price p_t, the displayed reference R_t, and an outcome weight $w(q_t)$ that governs how much influence period t has on future references. The platform's commitment is embodied in the choice of this weight function (and the initial historical weight), which is naturally interpretable as an interface and measurement design decision.

The first message that emerges—and that we view as a useful "sanity check" for any reference rule advertised as transaction-based—is a no-free-anchoring property: periods with negligible outcomes should have negligible influence on the future reference, and periods with zero outcomes should have exactly zero influence. This property formalizes the idea that one cannot manipulate the benchmark merely by posting a number; one must actually transact at that number. The second message is dynamic: as the platform's accumulated weight grows, the reference becomes increasingly stable, and the seller's optimal pricing policy becomes close to a stationary (or fixed-price) benchmark. In practical terms, when the reference statistic is built on a sufficiently deep base of realized transactions, it becomes difficult for any single seller to steer it quickly, and the seller's incentives look more like those in a static problem. This helps rationalize why outcome-weighted reference displays can be both robust to manipulation and operationally stable.

Our contribution is thus twofold. Substantively, we provide a microfoundation for a class of platform reference rules that are increasingly relevant for consumer protection and marketplace integrity. Methodologically, we offer a tractable dynamic program in which the platform's update rule can be varied and evaluated, making it possible to compare manipulation-resistance and volatility within a unified framework. We also acknowledge what the model does not attempt to do: consumers are modeled as myopic (capturing reference dependence but not strategic learning), the platform is assumed able to commit to its rule, and we treat the reference statistic as a reduced-form driver of demand rather than modeling the full cognitive process behind reference formation. These limitations are deliberate, because they isolate the mechanism of interest: the mapping from realized outcomes to future displayed references.

The remainder of the paper begins by formalizing the seller's problem, the platform's outcome-weighted update rule, and definitions of anchoring and manipulation metrics (Section 2). We then derive general properties of outcome-weighted updating, illustrate how and when dynamic pricing deviates from myopic pricing, and discuss the platform design tradeoffs implied by different choices of the weight function.

2 Model: seller pricing, reference-dependent demand, and outcome-weighted reference updating

We study a finite-horizon environment in which a seller chooses posted prices while a platform displays and updates a reference statistic that affects demand. The goal of this section is to formalize (i) the seller's dynamic pricing problem under reference-dependent demand and (ii) the platform's outcome-weighted reference update rule. We also define the objects we will later use to operationalize "anchoring" and "manipulation" in a way that is tied directly to the mechanics of the update rule.

2.1 Time, actions, and state variables

Time is discrete with periods $t \in \{1, \dots, T\}$. In each period, the seller posts a price $p_t \in [0, \bar{p}]$, where $\bar{p} < \infty$ is a feasibility bound (e.g., an interface constraint or an economically meaningful maximum). The platform displays at the start of period t a reference price $R_t \in [0, \bar{p}]$ and a cumulative weight $W_t \geq W_1 > 0$, where W_1 summarizes pre-existing transaction history or a platform-chosen baseline mass.[2]

A period proceeds as follows. First, the platform displays (R_t, W_t). Second, the seller chooses p_t. Third, sales (or a more general outcome such as engagement) realize according to a demand function $q_t = D(p_t, R_t) \geq 0$ that depends on the posted price and the current displayed reference. Finally, the

[2] Allowing $W_1 = 0$ is possible but creates uninteresting knife-edge sensitivity of the reference to early periods. We treat $W_1 > 0$ as capturing the idea that reference displays typically have some historical basis when the horizon we analyze begins.

platform updates (R_{t+1}, W_{t+1}) as a deterministic function of (R_t, W_t, p_t, q_t) according to the outcome-weighted rule described below.

The seller observes (R_t, W_t, t) when setting p_t. Consumers are myopic and summarized by $D(\cdot)$: they observe (p_t, R_t) and respond within period t, but we do not model forward-looking consumer inference about future $R_{t'}$. The platform is assumed able to commit ex ante to its reference computation rule (equivalently, to a choice of the weight function and initialization), and then implement it mechanically each period.

2.2 Reference-dependent demand

We keep demand reduced-form to focus on how platform design shapes the incentives to influence the reference. The baseline assumptions are monotonicity and feasibility:

$$D(p, R) \geq 0, \qquad \frac{\partial D(p, R)}{\partial p} \leq 0, \qquad \frac{\partial D(p, R)}{\partial R} \geq 0, \tag{1}$$

with weak inequalities allowing for kinks and truncation at zero demand. The second inequality captures downward-sloping demand in own price, while the third captures a reference effect: a higher displayed benchmark increases conversion at a given posted price (or, equivalently, makes a given posted price feel like a "gain" relative to the benchmark). The analysis below accommodates piecewise linear or otherwise non-smooth demand, which is natural when gains and losses relative to the reference have different salience.

For concreteness (and to connect later to tractable comparative statics), a workhorse specification we return to is a kinked linear form:

$$D(p, R) = \max\Big\{0,\, b - ap + \eta_+(R - p)_+ - \eta_-(p - R)_+\Big\}, \tag{2}$$

where $a > 0$ is the own-price slope, $b > 0$ is baseline demand, and $\eta_+, \eta_- \geq 0$ govern the strength of gain and loss reference effects. When $\eta_+ = \eta_- \equiv \eta$, the reference effect is symmetric and (2) reduces to $D(p, R) = \max\{0, b - (a + \eta)p + \eta R\}$ in the interior region. When $\eta_- > \eta_+$, demand is more sensitive to "losses" ($p > R$) than to "gains" ($p < R$), generating the familiar kink at $p = R$ that will later matter for optimal pricing and for the shape of the value function.

2.3 Platform rule: outcome-weighted reference updating

The platform computes the displayed reference as a moving statistic of past posted prices, but crucially it weights periods by realized outcomes rather than by mere postings. Let $w : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing outcome-weight function with $w(0) = 0$. The platform updates the cumulative weight by

$$W_{t+1} = W_t + w(q_t), \tag{3}$$

and updates the reference price by an outcome-weighted average:

$$R_{t+1} = \frac{W_t R_t + w(q_t)\, p_t}{W_t + w(q_t)}. \tag{4}$$

Thus, period t affects the future reference only through the weight assigned to its realized outcome. The mapping (4) is deliberately mechanical: it can be interpreted as a "typical transacted price" statistic where each period contributes proportionally to the amount of economic activity realized at that period's price.

Two remarks motivate this specification. First, (4) nests several common platform choices through w. When $w(q) = q$, the reference is a quantity-weighted average of posted prices (equivalently, a revenue-weighted average divided by total quantity if each unit has the same price). When $w(q) = q^\gamma$ with $\gamma > 1$, the platform downweights low-volume periods more aggressively, which intuitively increases resistance to low-sales "probing" while concentrating influence on high-volume events. Second, the state variable W_t matters economically: it measures the depth of accumulated history embedded in the reference. When W_t is large, a single period's weight $w(q_t)$ is small relative to history, so the reference becomes locally stable and the scope for dynamic steering shrinks.

2.4 Seller's objective and dynamic program

The seller receives per-period revenue $p_t q_t$ and maximizes total expected revenue over the horizon:

$$\Pi := \sum_{t=1}^{T} p_t q_t, \qquad q_t = D(p_t, R_t). \tag{5}$$

In the baseline model, demand is deterministic conditional on (p_t, R_t), so the seller's problem is a deterministic dynamic program with state (R_t, W_t, t). Let $V_t(R, W)$ denote the maximal continuation revenue from period t onward given state (R, W). Then

$$V_t(R, W) = \max_{p \in [0, \bar{p}]} \left\{ p\, D(p, R) + V_{t+1}(R', W') \right\}, \qquad V_{T+1}(\cdot) = 0, \tag{6}$$

where (R', W') is the next-period state induced by p through (3)–(4) with $q = D(p, R)$. When D is smooth and the maximizer is interior, the first-order condition takes the familiar form: the seller trades off current marginal revenue against the marginal continuation value generated by shifting the future state:

$$\frac{\partial}{\partial p}\big(pD(p, R)\big) + V_{t+1,R}(R', W') \cdot \frac{\partial R'}{\partial p} + V_{t+1,W}(R', W') \cdot \frac{\partial W'}{\partial p} = 0. \tag{7}$$

In kinked specifications such as (2), V_t may be non-differentiable at loci where the optimal price crosses R or where demand hits zero; in those regions the appropriate optimality conditions are in terms of subgradients. Our results will not rely on global differentiability; instead, we will emphasize identities and bounds that follow directly from the update rule.

Finally, while we focus on a deterministic D to keep the core logic transparent, it is straightforward to add noise (e.g., $q_t = D(p_t, R_t) + \varepsilon_t$ truncated at zero). In that case (6) becomes an expectation over ε_t, and the same state variables remain sufficient because the platform rule depends on realized q_t only through $w(q_t)$.

2.5 Anchoring, influence, and manipulation metrics

To connect the model to the practical concern of "fake discounts," we distinguish between *influence on the displayed reference* and *economic cost of generating that influence.* Outcome-weighting is designed precisely to align these two: moving the reference should require real outcomes.

We formalize "anchoring" as the act of increasing (or decreasing) the future reference relative to a counterfactual path. Fix a platform rule (w, W_1, R_1) and consider two seller price sequences $\{p_t\}$ and $\{\tilde{p}_t\}$ generating state paths $\{(R_t, W_t)\}$ and $\{(\tilde{R}_t, \tilde{W}_t)\}$ from the same initial condition. The period-t anchoring effect on the next reference is

$$\Delta R_t(p_t; \tilde{p}_t) := R_{t+1} - \tilde{R}_{t+1}, \tag{8}$$

and cumulative anchoring over a window can be measured by $\sum_{t \in \mathcal{T}} (R_{t+1} - \tilde{R}_{t+1})$ or by $\max_{t \in \mathcal{T}} (R_t - \tilde{R}_t)$ depending on the application (e.g., sustained elevation versus a one-time spike). In many comparisons we will take $\{\tilde{p}_t\}$ to be a myopic benchmark that ignores the effect of p_t on future references, so that anchoring captures purely dynamic steering incentives.

To operationalize "manipulation," we need a notion that captures the idea of *disproportionate* reference movement generated by *negligible* real activity. A convenient period-by-period metric is a normalized influence ratio:

$$\mathcal{I}_t := \frac{|R_{t+1} - R_t|}{\max\{q_t, \epsilon\}}, \tag{9}$$

for a small $\epsilon > 0$ used only to avoid division by zero in empirical implementations. Under posted-price-based rules, $\mathcal{I}_t$ can be arbitrarily large because a seller may move R_{t+1} while generating $q_t \approx 0$. Under outcome-weighting, by contrast, the platform mechanically links the numerator to the realized outcome through $w(q_t)$, which will yield sharp restrictions on $\mathcal{I}_t$ and, in the limit $q_t \to 0$, eliminate the possibility of moving R at all.

A closely related "deception pressure" metric is the sensitivity of the next reference to the posted price at low sales:

$$\mathcal{S}_t := \left.\frac{\partial R_{t+1}}{\partial p_t}\right|_{(R_t, W_t)}, \qquad \text{evaluated at the realized } q_t = D(p_t, R_t). \tag{10}$$

This object is directly interpretable as the platform's mechanical pass-through from a seller's chosen posted price to the displayed benchmark. A rule is more manipulation-resistant when $\mathcal{S}_t$ is small precisely in those states where the seller can cheaply engineer low outcomes (e.g., by setting very high prices or limiting inventory). Outcome-weighting aims to make $\mathcal{S}_t$ endogenous in q_t: low q_t implies low weight and hence low sensitivity.

We will sometimes aggregate these period metrics into a platform-level objective that trades off consumer surplus, seller revenue, and manipulation exposure. One parsimonious reduced-form representation is

$$U^{\text{plat}} := \sum_{t=1}^{T} \Big(\text{CS}(p_t, R_t) + p_t q_t\Big) - \lambda \, \mathcal{M}\big(\{p_t, q_t, R_t\}_{t \le T}\big), \tag{11}$$

where $\mathcal{M}$ can be instantiated as $\sum_t \mathbf{1}\{q_t \le \varepsilon\} \, |R_{t+1} - R_t|$ (reference movement during low-sales periods), or as $\sum_t \mathcal{I}_t$, or as $\sum_t \mathcal{S}_t \mathbf{1}\{q_t \le \varepsilon\}$. We do not need to take a stand on the unique "correct" manipulation functional; the point is that outcome-weighting is naturally evaluated by how strongly it suppresses reference movement when outcomes are negligible.

This completes the model. The next section derives a simple identity implied by (4) that formalizes the central discipline of outcome-weighting: posted prices without outcomes do not move the reference, and more generally the magnitude of reference movement is governed by the share of cumulative weight contributed by current realized activity.

3 A No-Free-Anchoring Identity

Outcome-weighted updating delivers a simple but powerful restriction on how much a seller can steer the platform's displayed benchmark. The restriction is entirely mechanical: it comes from the algebra of the update rule and does not require any assumption about optimal behavior, smoothness of demand, or the horizon. Because it holds path-by-path, it is also robust to stochastic demand as long as the platform updates using the realized outcome.

Proposition 1 (No-Free-Anchoring Identity). For any period t, any outcome-weight function w with $w(0) = 0$, and any realized outcome $q_t \ge 0$, the reference update (4) satisfies

$$R_{t+1} - R_t \; = \; \alpha_t \, (p_t - R_t), \qquad \alpha_t \; := \; \frac{w(q_t)}{W_t + w(q_t)} \in [0, 1). \tag{12}$$

Equivalently,

$$R_{t+1} \;=\; (1-\alpha_t)R_t + \alpha_t p_t, \tag{13}$$

so the next reference is a convex combination of the current reference and the current posted price with a mixing weight determined by realized activity. In particular, if $q_t = 0$ then $\alpha_t = 0$ and hence $R_{t+1} = R_t$ for any posted price p_t.

Proof. Starting from (4),

$$R_{t+1} - R_t = \frac{W_t R_t + w(q_t)p_t}{W_t + w(q_t)} - R_t = \frac{w(q_t)}{W_t + w(q_t)}(p_t - R_t),$$

which is (12). Since $W_t \geq W_1 > 0$ and $w(q_t) \geq 0$, we have $\alpha_t \in [0,1)$, and if $q_t = 0$ then $w(q_t) = 0$ so $\alpha_t = 0$. □

Identity (12) clarifies what outcome-weighting does *and* what it does not do. It does not eliminate anchoring incentives in general: if the seller can profitably generate real activity at a price above the current reference, then the seller can indeed push R_{t+1} upward. What it eliminates is "free" anchoring driven by posted prices that do not correspond to realized outcomes. Put differently, the platform rule implements a hard accounting principle: only activity gets counted.

Implication 1: extreme posted prices without outcomes are irrelevant. Under posted-price-based rules (e.g., an unweighted average of posted prices), a seller can post an extreme price for one period, generate essentially no sales, and nevertheless move the displayed reference by a large amount. Under (12), that logic fails exactly. If an extreme price drives demand to zero, then $q_t = 0$ and the reference does not move. More generally, if an extreme price yields negligible sales, then $w(q_t)$ is small relative to W_t and α_t is small, so the induced movement is negligible even if $|p_t - R_t|$ is large:

$$|R_{t+1} - R_t| = \alpha_t |p_t - R_t| \leq \frac{w(q_t)}{W_t + w(q_t)}\,\bar{p}. \tag{14}$$

The bound (14) makes transparent that the "leverage" of any attempted anchoring episode is limited by the period's weight share α_t, not by the posted price alone.

Implication 2: anchoring is proportional to weight share. The object α_t has a natural interpretation as an *anchoring leverage share*: it is the fraction of total cumulative weight (history plus current period) contributed by current realized activity. When W_t is large, history is deep and even substantial current activity has limited leverage; when W_t is small, early periods can matter more. This is precisely why the initialization W_1 matters for manipulation exposure: larger legacy mass makes α_t smaller in every early period, mechanically stabilizing R_t.

Equation (13) also implies monotonicity and sign restrictions that are often convenient. If $p_t \geq R_t$ then $R_{t+1} \geq R_t$; if $p_t \leq R_t$ then $R_{t+1} \leq R_t$. Thus the seller cannot "overshoot" the posted price in a single step: $R_{t+1} \in [\min\{R_t, p_t\}, \max\{R_t, p_t\}]$. In this sense, outcome-weighting enforces a tight and intuitive geometry for reference paths.

Implication 3: "fake discounts" lose their mechanical channel. A common concern motivating reference regulation is a pattern in which a seller posts a high "list price" to elevate the displayed benchmark and then advertises a discount relative to that elevated benchmark. The identity (12) pinpoints the crucial vulnerability in naive rules: if merely posting a high price raises the benchmark, then the seller can create an inflated R_{t+1} even when consumers never transact at that price. Under outcome-weighting, the benchmark moves only insofar as consumers actually buy (or otherwise generate the platform's measured outcome) at the high posted price. Thus, the seller must either (i) sell meaningful volume at the high price—which is costly under downward-sloping demand—or (ii) accept that the reference will barely change.

This does not mean a seller can never raise the reference; it means raising the reference cannot be decoupled from generating the very economic activity that would justify a higher "typical price" statistic. In practical terms, the platform is no longer rewarding empty list prices; it is rewarding transacted prices.

Implication 4: quantitative discipline for influence and sensitivity metrics. The period-by-period manipulation metrics defined earlier are naturally controlled by (12). First, the normalized influence ratio (9) satisfies

$$\mathcal{I}_t = \frac{|R_{t+1} - R_t|}{\max\{q_t, \epsilon\}} \leq \frac{w(q_t)}{W_t + w(q_t)} \cdot \frac{\bar{p}}{\max\{q_t, \epsilon\}}. \tag{15}$$

When $w(q) = q$, the right-hand side becomes $\frac{q_t}{W_t+q_t} \cdot \frac{\bar{p}}{\max\{q_t,\epsilon\}}$, which is uniformly bounded by $\bar{p}/W_t$ once $q_t \geq \epsilon$ and, crucially, does not explode as $q_t \downarrow 0$ (indeed, $R_{t+1} - R_t \to 0$ as $q_t \to 0$). This sharply contrasts with posted-price-based rules in which the numerator can remain large while q_t becomes arbitrarily small.

Second, the mechanical "pass-through" from p_t to R_{t+1} can be read off directly when we condition on the realized outcome. Holding q_t fixed, (4) is affine in p_t with slope α_t. When we account for the fact that $q_t = D(p_t, R_t)$ typically changes with p_t, the total derivative includes an additional term reflecting the endogenous change in weight:

$$\frac{dR_{t+1}}{dp_t} = \alpha_t + (p_t - R_t)\frac{d}{dp_t}\left(\frac{w(D(p_t, R_t))}{W_t + w(D(p_t, R_t))}\right). \tag{16}$$

If D is differentiable in p at the relevant point, then

$$\frac{d\alpha_t}{dp_t} = \frac{W_t\, w'(q_t)\, D_p(p_t, R_t)}{(W_t + w(q_t))^2}, \tag{17}$$

so $D_p \leq 0$ implies $\frac{d\alpha_t}{dp_t} \leq 0$. Thus, when the seller raises price above the current reference ($p_t > R_t$), the second term in (16) is weakly negative: increasing p_t tends to reduce demand, reduce weight, and thereby *dampen* the impact on R_{t+1}. This is exactly the intended discipline: attempts to elevate the benchmark through high prices become self-limiting because they reduce the activity that grants influence in the first place.

Implication 5: outcome-weighting converts anchoring into a real-sales requirement. Identity (12) can be inverted to express a necessary weight share for achieving a targeted reference movement. Fix any desired one-step increase $\delta > 0$. If the seller posts $p_t \geq R_t$ and wants $R_{t+1} - R_t \geq \delta$, then necessarily $\alpha_t \geq \delta/(p_t - R_t)$, hence

$$w(q_t) \;\geq\; \frac{\delta}{(p_t - R_t) - \delta}\, W_t, \qquad \text{whenever } p_t - R_t > \delta. \tag{18}$$

Condition (18) makes precise the idea that meaningful anchoring requires contributing a nontrivial fraction of the accumulated history. When W_t is large, a one-step movement of size δ requires very large realized weight (and thus, under standard demand primitives, nontrivial sales). This immediately suggests why dynamic incentives attenuate over time in our environment: as history accumulates, marginal steering power decays.

Practical interpretation and limitations. Outcome-weighted updating addresses a narrow but important vulnerability: it breaks the purely *posted-price* channel of manipulation. That said, it does not by itself solve every practical concern about reference integrity. A seller might attempt to generate "outcomes" strategically (e.g., self-purchases, coordinated buying, subsidized transactions) to obtain weight at an inflated price. From the platform's perspective, this shifts the problem from "fake postings" to "fake outcomes," which is often a more familiar enforcement domain: platforms can audit transactions, exclude refunded orders from q_t, require verified purchases, or define $w(\cdot)$ in terms of harder-to-game outcomes (e.g., completed, non-returned orders).

A second limitation is volatility around genuine high-volume events. If a promotion generates an unusually large q_t, then α_t can be large and R_{t+1}

can jump toward the promotional price (or away from it), especially under convex weighting rules such as $w(q) = q^\gamma$. This is not a failure of the no-free-anchoring logic; it is the counterpart of giving real activity real influence. The platform's design problem, which we return to later, is therefore not whether to discipline manipulation—(12) already does that—but how to choose w (and the effective history mass W_1) to balance manipulation-resistance against responsiveness to genuinely informative high-volume price episodes.

Finally, we emphasize that Proposition 1 is purely an accounting identity. It does not require downward-sloping demand, symmetry of reference effects, or optimal seller behavior. Its role is foundational: it isolates the exact lever through which a seller can influence the benchmark—the realized-weight share α_t—and thereby turns the analysis of "fake discounts" into an analysis of how costly it is to acquire weight at manipulated prices. The next step is to embed this mechanical discipline into the seller's intertemporal problem and characterize best responses via a dynamic program in the state (R_t, W_t, t).

4 Seller Best Response as a Dynamic Program

Having isolated the purely mechanical restriction on how posted prices can affect future references, we now embed the platform's updating rule into the seller's intertemporal problem. Conceptually, the key point is that outcome-weighting makes the seller's problem genuinely dynamic *only* through the two state variables carried forward by the platform: the current displayed reference R_t and the accumulated history mass W_t. Once we condition on (R_t, W_t, t), the past matters only through these sufficient statistics, and the seller faces a standard finite-horizon control problem on a compact action set.

State, action, and law of motion. Fix a horizon T and feasible prices $p \in [0, \bar{p}]$. In period t, given state (R, W), the seller chooses a posted price p, inducing realized activity

$$q \;=\; D(p, R) \;\geq 0.$$

The platform then updates the weight and reference according to

$$W^+(R, W, p) \;=\; W + w(D(p, R)), \qquad R^+(R, W, p) \;=\; \frac{WR + w(D(p, R))\, p}{W + w(D(p, R))}. \tag{19}$$

We emphasize that (19) is well-defined for all $W \geq W_1 > 0$ because $W + w(\cdot)$ is strictly positive. The restriction $w(0) = 0$ implies that periods with zero realized activity leave the state unchanged.

Bellman equation and existence of optimal Markov policies. Let $V_t(R, W)$ denote the seller's maximal continuation revenue from period t onward given state (R, W) at the start of period t. The seller's dynamic program is

$$V_t(R, W) \;=\; \max_{p\in[0,\bar{p}]} \Big\{ p\, D(p, R) + V_{t+1}\big(R^+(R, W, p),\, W^+(R, W, p)\big) \Big\}, \qquad V_{T+1}(\cdot) \equiv 0. \tag{20}$$

Under mild regularity—for instance, $D(\cdot, \cdot)$ continuous and bounded on $[0, \bar{p}] \times [0, \bar{p}]$ and $w(\cdot)$ continuous and increasing—the maximand in (20) is continuous in p and the choice set is compact, so an optimal policy exists by the Weierstrass theorem. Because the state evolution (19) depends on history only through (R, W), we can restrict attention (without loss) to Markov policies $p_t = \pi_t(R_t, W_t)$ obtained by backward induction on (20). This is the sense in which the platform's outcome-weighted statistic turns the seller's strategic problem into a two-dimensional dynamic pricing problem.

A convenient "geometry": the next reference lies between R and p. A recurring simplification is that, for any p and induced $q = D(p, R)$, the update rule implies

$$R^+(R, W, p) \;=\; (1-\alpha(R, W, p))\, R + \alpha(R, W, p)\, p, \qquad \alpha(R, W, p) := \frac{w(D(p, R))}{W + w(D(p, R))} \in [0, 1). \tag{21}$$

Thus $R^+(R, W, p)$ is always trapped in the interval with endpoints $\{R, p\}$. In the dynamic program, this means that current pricing cannot "jump" the benchmark past the posted price, and any attempt to move the benchmark must pay for movement through the realized weight share α. While this observation is algebraic, it has a structural implication for (20): the continuation value is evaluated only at references that lie on a one-dimensional segment indexed by p, rather than at arbitrary points in the (R, W) plane.

Monotonicity in the reference. When demand is weakly increasing in the reference (our maintained assumption), the seller weakly benefits from a higher current benchmark, both contemporaneously (through $D(p, R)$) and dynamically (because a higher R makes future demand higher for any future prices). This yields a useful order property.

To make the statement precise, suppose $D(p, R)$ is weakly increasing in R for each p, and w is increasing. Fix any (W, t) and two reference levels $R \le \widetilde{R}$. For any price p, we have $D(p, R) \le D(p, \widetilde{R})$, hence $W^+(R, W, p) \le W^+(\widetilde{R}, W, p)$. Moreover, holding p fixed, the next reference $R^+(R, W, p)$ is weakly increasing in R because it is an average of R and p with nonnegative weights and because the endogenous weight $w(D(p, R))$ is weakly larger at $\widetilde{R}$. By induction on t in (20), these monotonicities imply

$$R \le \widetilde{R} \quad \Longrightarrow \quad V_t(R, W) \le V_t(\widetilde{R}, W). \tag{22}$$

Economically, (22) formalizes that the seller never dislikes a higher benchmark in this environment: reference effects act like an intertemporal demand shifter.

Influence is uniformly bounded and vanishes with large history mass. The state variable W plays a distinct role: it does not directly enter the current-period demand $D(p, R)$, but it governs how "sticky" the reference is. For each fixed (R, p), the leverage share in (21) is decreasing in W, so the mapping $p \mapsto R^+(R, W, p)$ becomes flatter as W grows. This converts naturally into bounds on how much the seller can affect future states.

Assume that demand is uniformly bounded on the relevant domain: there exists $\bar{q} < \infty$ such that $0 \le D(p, R) \le \bar{q}$ for all $p \in [0, \bar{p}]$ and $R \in [0, \bar{p}]$. (In applications, this can be interpreted as a market size bound or a bound on engagement.) Then, for any state (R, W) and any feasible price p,

$$0 \le \alpha(R, W, p) \le \overline{\alpha}(W) := \frac{w(\bar{q})}{W + w(\bar{q})}. \tag{23}$$

Combining (21)–(23) with $|p - R| \le \bar{p}$ yields the uniform one-step movement bound

$$|R^+(R, W, p) - R| \ \le \ \overline{\alpha}(W)\, \bar{p}. \tag{24}$$

Inequality (24) is the dynamic-programming counterpart of the no-free-anchoring logic: even if the seller is optimizing strategically, the set of attainable next references from (R, W) shrinks mechanically as W grows. In particular, early periods (small W) are the only periods in which the seller can have meaningful "state leverage"; later periods are dominated by the current-period revenue term in (20). This observation is what ultimately underpins near-stationarity results in tractable specifications.

A Lipschitz bound for continuation values and a "small dynamic term" heuristic. To connect (20) to pricing behavior, it is useful to quantify how much the continuation value can change when the seller marginally perturbs the reference. Suppose $D(p, R)$ is Lipschitz in R uniformly in p: there exists L_D such that

$$|D(p, R) - D(p, \widetilde{R})| \le L_D\, |R - \widetilde{R}| \quad \text{for all } p \in [0, \bar{p}],\ R, \widetilde{R} \in [0, \bar{p}].$$

Then the single-period revenue $pD(p, R)$ is Lipschitz in R with constant at most $\bar{p}\, L_D$, and backward induction on (20) yields the bound

$$|V_t(R, W) - V_t(\widetilde{R}, W)| \ \le \ L_t\, |R - \widetilde{R}|, \qquad L_t \le (T - t + 1)\, \bar{p}\, L_D. \tag{25}$$

Heuristically, (25) means the seller's marginal value of the benchmark is finite and grows at most linearly with remaining horizon. Combining this with the movement bound (24) suggests a simple discipline on dynamic incentives: the maximal continuation-value gain achievable *through steering R in one step* is on the order of $L_{t+1}\overline{\alpha}(W)\bar{p}$, which decays like $1/W$ for large W when $w(\bar{q})$ is bounded. This is the sense in which large legacy mass W_1 (or simply the passage of time as W_t accumulates) makes strategic anchoring less attractive, even before we impose any functional-form structure.

First-order conditions and kinked demand. When D and w are differentiable, an interior optimizer for (20) satisfies an Euler-type condition balancing current marginal revenue against the discounted (here undiscounted, given the finite horizon) effect of price on future states:

$$\frac{\partial}{\partial p}\big(pD(p,R)\big) + V_{t+1,R}(R^+,W^+)\,\frac{\partial R^+}{\partial p} + V_{t+1,W}(R^+,W^+)\,\frac{\partial W^+}{\partial p} = 0. \quad (26)$$

In our environment, however, it is important not to over-invest in differentiability: empirically relevant reference-dependent demand often has kinks (e.g., different slopes above and below R), and platform weight functions may also be piecewise or capped. The dynamic program (20) remains valid without smoothness, and the appropriate optimality conditions are then expressed using subgradients and one-sided derivatives. Practically, this is not merely a mathematical caveat: kinks are precisely where "loss" versus "gain" reference effects can induce bunching of optimal prices at $p = R$, and any characterization must accommodate that possibility.

Interpretation for platform design. The dynamic program highlights a policy-relevant distinction between *incentives* and *mechanisms*. The seller may still have an incentive to raise R (because V_t is increasing in R), but the mechanism by which R can be raised is tightly constrained by the transition (19): the seller must generate realized outcomes that earn weight. From a platform's perspective, this means that concerns about "fake discounts" move from being a problem of posted-price bookkeeping to being a problem of outcome integrity (e.g., whether q_t represents genuine transactions). The virtue of the outcome-weighted design is that it aligns the displayed statistic with economically meaningful activity; the residual vulnerability, if any, is the possibility of strategically created activity, which is typically more auditable than empty postings.

Where we go next. Section 4 provides the general-purpose dynamic-programming foundation: a two-dimensional Markov state, a compact pricing action, and mechanical bounds showing that state influence shrinks as history mass grows. To obtain sharper behavioral predictions—in particular, near-stationarity and fixed-price near-optimality—we next specialize to a tractable linear specification in which we can quantify the marginal value of R and translate the "small influence" heuristic into explicit pricing bounds.

5 A Tractable Linear–Quadratic Benchmark

We now specialize the primitives to a linear reference-dependent demand and a linear outcome weight, which together yield a "linear–quadratic" structure in the seller's objective. The benefit of this benchmark is not that it is literally quadratic everywhere (demand is truncated at zero), but that on the economically relevant interior region the seller's dynamic problem becomes a smooth concave control problem with transparent comparative statics. In particular, we can (i) write the myopic best response in closed form, (ii) bound the marginal continuation value of the reference, and (iii) formalize the near-stationarity heuristic from the previous section: once W_t is moderate, the optimal dynamic price is extremely close to the myopic price, and the incremental value of dynamically "steering" R_t is at most polylogarithmic in the effective history mass.

Specification. Fix parameters (a, b, η) with $a > 0$ and $\eta \geq 0$, and impose symmetric reference effects with linear weighting:

$$D(p, R) \;=\; \max\{0,\, b - ap + \eta(R - p)\} \;=\; \max\{0,\, b + \eta R - (a+\eta)p\}, \qquad w(q) = q. \tag{27}$$

Write $k := a + \eta$ for the effective own-price slope in the interior region. When $D(p, R) > 0$, demand is affine with derivatives $D_p = -k < 0$ and $D_R = \eta \geq 0$. We maintain an interiority/nonnegativity regime under which (for the states we study) the relevant maximizers satisfy $D(p, R) > 0$ and $p \in (0, \bar{p})$, so that local first-order conditions are informative; when this fails (e.g., low b or extreme R), the optimal policy may hit corners, but the same "vanishing influence" logic continues to apply with one-sided derivatives.

Closed-form myopic pricing. Holding (R, t) fixed and ignoring future effects, the seller solves

$$\max_{p \in [0, \bar{p}]} p\, D(p, R) = \max_{p \in [0, \bar{p}]} \big(p(b + \eta R) - kp^2\big) \quad \text{on the region where } D(p, R) > 0.$$

The unconstrained maximizer is

$$p^{\text{my}}(R) \;=\; \frac{b + \eta R}{2k} \;=\; \frac{b + \eta R}{2(a + \eta)}. \tag{28}$$

Thus the myopic price is increasing in the displayed benchmark: a higher R relaxes the demand intercept and shifts the seller's static best response upward one-for-one at rate $\eta/(2(a + \eta))$. In this benchmark, the myopic objective is strictly concave in p on the interior region because

$$\frac{\partial^2}{\partial p^2}\big(pD(p, R)\big) = -2k < 0, \tag{29}$$

a property we will reuse to bound deviations of the dynamic optimum from $p^{\text{my}}(R)$.

State transitions and an explicit leverage derivative. With $w(q) = q$ and $q = D(p, R)$, the state update becomes

$$W^+ = W + q, \qquad R^+ = \frac{WR + qp}{W + q} = R + \alpha(p - R), \qquad \alpha = \frac{q}{W + q} \in [0, 1). \tag{30}$$

In the interior region ($q > 0$), the mapping $p \mapsto R^+$ is differentiable. Using $\alpha = q/(W + q)$ and $\alpha_p = q_p W/(W + q)^2$, we obtain

$$\frac{\partial R^+}{\partial p} = \alpha + (p - R)\alpha_p = \frac{q}{W + q} + (p - R)\,\frac{q_p W}{(W + q)^2}, \qquad q_p = -k. \tag{31}$$

The key implication is a uniform "flatness" bound for large W. If q is bounded above on the relevant domain by $\bar{q}$ (automatic here given truncation and $p \in [0, \bar{p}]$), then for all feasible (R, W, p) in the interior region,

$$\left| \frac{\partial R^+}{\partial p} \right| \;\le\; \frac{\bar{q}}{W} + \frac{k\,|p - R|}{W} \;\le\; \frac{\bar{q} + k\bar{p}}{W}. \tag{32}$$

Thus, even though the seller can always change p today, the ability of that change to transmit into tomorrow's displayed benchmark shrinks on the order of $1/W$.

Bounding the marginal continuation value of the benchmark. In this linear specification, the marginal effect of a higher benchmark on current revenue is particularly simple: since $\partial D/\partial R = \eta$ on the interior region,

$$\frac{\partial}{\partial R}\big(pD(p, R)\big) = p\,\eta \le \eta\bar{p}.$$

Backward induction then yields a crude but useful bound on the value gradient:

$$0 \le V_{t,R}(R, W) \le (T - t + 1)\,\eta\bar{p} \quad \text{whenever } V_{t,R} \text{ exists}, \tag{33}$$

reflecting that the benchmark acts like an intercept shifter with per-period marginal value at most $\eta\bar{p}$.

The derivative with respect to W is more subtle because W does not enter contemporaneous demand; it only governs how sensitive future references are to future posted prices. In the present benchmark, this makes V_t weakly *decreasing* in W (a larger history mass makes steering harder) but with a marginal effect that is second-order for large W. Intuitively, changing W by a small amount perturbs $\alpha = q/(W + q)$ by order $1/W^2$, and hence perturbs future references—and thus future revenue—only at that rate. Formally, one can show (under the same boundedness conditions used above) that there exists a constant $C_V < \infty$ such that

$$|V_{t,W}(R, W)| \;\le\; \frac{C_V}{W^2} \qquad \text{for large } W, \tag{34}$$

so the $V_{t+1,W}$ term in the Euler condition is negligible relative to the $V_{t+1,R}$ term once W is moderate. The economic message is simple: W matters only as a "friction" on reference manipulation, and frictions have diminishing marginal bite as the friction becomes large.

Near-stationarity: the dynamic price is $O(1/W)$-close to myopic. Consider a period t and state (R, W) in the interior region, and let $p_t^*(R, W)$ denote an interior optimizer of the Bellman problem. The first-order condition (26) specializes to

$$\underbrace{\frac{\partial}{\partial p}\big(pD(p, R)\big)}_{\text{static marginal revenue}} + V_{t+1,R}(R^+, W^+)\,\frac{\partial R^+}{\partial p} + V_{t+1,W}(R^+, W^+)\,\frac{\partial W^+}{\partial p} = 0, \qquad \frac{\partial W^+}{\partial p} = q_p = -k. \tag{35}$$

At the myopic optimizer $p^{\text{my}}(R)$, the static marginal revenue term vanishes. Hence the wedge between the dynamic and myopic prices is governed by the size of the two continuation terms. Using (29), we can bound the price deviation by dividing the maximal continuation wedge by the strong concavity modulus $2k$. Combining (32)–(34) yields the scaling

$$\big|p_t^*(R, W) - p^{\text{my}}(R)\big| \;\le\; \frac{1}{2k}\left(\sup|V_{t+1,R}| \cdot \sup\left|\frac{\partial R^+}{\partial p}\right| + \sup|V_{t+1,W}| \cdot k\right) \;=\; O\!\left(\frac{1}{W}\right), \tag{36}$$

uniformly over interior states. Economically, the seller's dynamic incentive to distort current price away from the static optimum is the product of two terms: the marginal value of the benchmark $V_{t+1,R}$ (bounded) and the marginal ability to move the benchmark $\partial R^+/\partial p$ (shrinking like $1/W$). This is precisely the "small dynamic term" heuristic from the general analysis, now made explicit in a parametric environment.

Fixed-price near-optimality and a polylogarithmic gap. Near-stationarity implies that the *incremental* value of dynamic steering is small when aggregated over time. A convenient way to see the polylogarithmic structure is to connect leverage shares to the growth of history mass. With $w(q) = q$ and $W^+ = W + q$, we can rewrite

$$\alpha_t = \frac{q_t}{W_t + q_t} = 1 - \frac{W_t}{W_{t+1}}. \tag{37}$$

Summing (37) and using $\log x \le x - 1$ (equivalently, $1 - 1/x \le \log x$ for $x \ge 1$) gives the telescoping bound

$$\sum_{t=1}^{T} \alpha_t = \sum_{t=1}^{T}\left(1 - \frac{W_t}{W_{t+1}}\right) \;\le\; \sum_{t=1}^{T} \log\!\left(\frac{W_{t+1}}{W_t}\right) = \log\!\left(\frac{W_{T+1}}{W_1}\right). \tag{38}$$

Because one-step reference movement satisfies $|R_{t+1} - R_t| \leq \alpha_t \bar{p}$, (38) implies that even an optimally chosen, fully history-dependent pricing policy can only "rotate" the benchmark by a total amount that grows at most logarithmically in the final history mass.

This logarithmic geometry translates into a polylogarithmic performance bound for simple policies. One route (developed formally in Proposition 4) is to compare the dynamic first-order condition to the myopic first-order condition and use (36) to show that the *per-period* gain from dynamic distortion is at most on the order of $1/W_t$. Summing and applying the same telescoping logic as in (38) yields

$$\Pi^{\text{dyn}} - \Pi^{\text{myopic Markov}} \leq C \sum_{t=1}^{T} \frac{1}{W_t} = O\left(\log\left(\frac{W_{T+1}}{W_1}\right)\right), \tag{39}$$

for a constant C depending on $(a, b, \eta, \bar{p})$ but not on T. A second route is to compare the dynamic optimum to the best *fixed* price $p \in [0, \bar{p}]$ directly: since the only intertemporal benefit of nonstationarity is mediated by cumulative benchmark movement, and cumulative movement is logarithmically bounded, the maximal advantage of fine-tuned dynamic manipulation over the best constant policy is likewise at most logarithmic in the total accumulated mass. Either way, the economic conclusion is the same: with outcome-weighted updating, the seller cannot extract a large intertemporal rent from reference steering, and simple pricing heuristics are robustly near-optimal.

Discussion and limitations. The linear–quadratic benchmark is deliberately conservative in one dimension and optimistic in another. It is conservative because linear $w(q) = q$ does *not* aggressively downweight low-sales events; in Section 6 we will consider $w(q) = q^\gamma$ as a design lever to further suppress low-volume anchoring. It is optimistic because symmetric reference effects remove kinks at $p = R$; with asymmetric (η_+, η_-) the seller's objective can develop nondifferentiabilities and bunching at $p = R$, requiring piecewise arguments or numerical dynamic programming. Still, the benchmark illustrates the core mechanism we want to carry into platform design: when the platform pegs displayed references to *realized outcomes*, dynamic manipulation incentives fade mechanically with accumulated history mass, and the remaining design question becomes how the choice of $w(\cdot)$ trades off manipulation-resistance against reference volatility.

6 Platform Design: Choosing the Outcome-Weight Function $w(\cdot)$

Up to this point we have taken the platform's reference computation rule as given and studied how the seller responds. We now reverse the perspective and treat $w(\cdot)$ as a design choice. The central observation behind the "no-free-anchoring" logic is that the displayed benchmark moves only through realized outcomes; nevertheless, *how much* it moves after a given amount of sales is pinned down by the leverage share

$$\alpha_t \;=\; \frac{w(q_t)}{W_t + w(q_t)} \in [0,1), \qquad R_{t+1} - R_t \;=\; \alpha_t \,(p_t - R_t).$$

Thus $w(\cdot)$ governs the platform's exposure to two opposing desiderata. On the one hand, we want manipulation-resistance: low-sales or "thin" periods should have little influence on R_{t+1}, so that a seller cannot profitably steer references using marginal, strategically chosen transactions. On the other hand, we want informational responsiveness: when demand is genuinely high (e.g., during a promotion that reflects real consumer willingness-to-pay), the reference should not be so inert that it becomes stale or misleading. In our language, the platform chooses how to map quantities into "mass" and thereby chooses the time scale on which the reference adapts.

A simple design family: $w(q) = q^\gamma$. A tractable and practically interpretable design lever is the power family

$$w_\gamma(q) \;=\; q^\gamma, \qquad \gamma \geq 1. \tag{40}$$

The case $\gamma = 1$ corresponds to linear outcome weighting (each unit sold contributes proportionally to the reference mass), while $\gamma > 1$ makes weights convex and disproportionately emphasizes high-volume events. Within this family, the leverage share becomes

$$\alpha_t(\gamma) \;=\; \frac{q_t^\gamma}{W_t + q_t^\gamma} \;=\; \left(1 + \frac{W_t}{q_t^\gamma}\right)^{-1}. \tag{41}$$

Two limiting regimes clarify the mechanism. When sales are small relative to the accumulated history mass, $q_t^\gamma \ll W_t$, we have

$$\alpha_t(\gamma) \;=\; \frac{q_t^\gamma}{W_t}\,(1 + o(1))\,, \tag{42}$$

so the mechanical reference movement scales like q_t^γ / W_t. Increasing γ sharply suppresses the ability of low-sales periods to move R_{t+1}. Conversely, when sales are very large relative to history, $q_t^\gamma \gg W_t$, we get $\alpha_t(\gamma) \approx 1$ and the reference nearly jumps to the current posted price. Hence convex weighting enhances robustness to thin trades but can create large reference updates during high-volume bursts.

Manipulation-resistance as a requirement on low-q leverage. A natural operational target is to ensure that periods with "small" sales cannot move the reference by more than a tolerable amount. Fix a sales threshold $q^{\min} > 0$ and a leverage tolerance $\bar{\alpha} \in (0, 1)$. A sufficient condition for the platform to guarantee $|R_{t+1} - R_t| \leq \bar{\alpha}\, |p_t - R_t|$ whenever $q_t \leq q^{\min}$ is

$$\sup_{q \in [0, q^{\min}]} \frac{w(q)}{W_t + w(q)} \;\leq\; \bar{\alpha}. \tag{43}$$

For the power family and a conservative (worst-case) design at the smallest history mass W_1, this becomes

$$\frac{(q^{\min})^\gamma}{W_1 + (q^{\min})^\gamma} \;\leq\; \bar{\alpha} \quad\Longleftrightarrow\quad (q^{\min})^\gamma \;\leq\; \frac{\bar{\alpha}}{1 - \bar{\alpha}}\, W_1. \tag{44}$$

Equation (44) makes transparent that manipulation-resistance is jointly engineered by (i) the convexity parameter γ and (ii) the initial mass W_1 (a "legacy" lookback stock). Platforms often have discretion over both: W_1 can be increased by expanding the lookback window or adding pseudo-counts, while γ adjusts how quickly marginal sales translate into effective weight. Either instrument reduces α_t in thin periods, but they differ in their side effects: raising W_1 slows adaptation uniformly, whereas raising γ slows adaptation primarily when q_t is small.

Reference volatility under stochastic demand. Outcome-weighting is especially attractive when we take seriously that observed q_t is noisy. Suppose realized sales take the form $q_t = D(p_t, R_t) + \varepsilon_t$ with mean-zero shocks and truncation at zero. Then the one-step reference movement is random through $\alpha_t = \frac{w(q_t)}{W_t + w(q_t)}$, and a platform that overreacts to high realizations risks producing a highly volatile benchmark. The key local statistic is the sensitivity of leverage to realized outcomes:

$$\frac{\partial \alpha}{\partial q} \;=\; \frac{w'(q)\, W_t}{\big(W_t + w(q)\big)^2}. \tag{45}$$

For $w(q) = q^\gamma$, we obtain

$$\frac{\partial \alpha}{\partial q} \;=\; \frac{\gamma q^{\gamma - 1} W_t}{\big(W_t + q^\gamma\big)^2}. \tag{46}$$

Holding (q, W_t) fixed, increasing γ reduces α for small q (as in (42)) but can increase the curvature of α in the neighborhood of moderate-to-high q. Intuitively, convex weighting concentrates influence on the upper tail of the sales distribution: a high realization of q_t (whether due to genuine demand or a transitory shock) can generate a disproportionately large jump in the benchmark. This provides a precise sense in which platform design is a tradeoff between manipulation-resistance (attenuating low-q influence) and reference volatility (amplifying high-q influence).

Mitigating volatility: caps, thresholds, and hybrid weighting. The power family (40) is deliberately parsimonious, but platforms are not restricted to it. Several simple modifications can preserve the no-free-anchoring property $w(0) = 0$ while tempering volatility:

- *Capped weights:* $w(q) = \min\{q^\gamma, \bar{w}\}$, which bounds α_t away from one even in extreme volume events.
- *Thresholded weights:* $w(q) = \big((q - \tau)_+\big)^\gamma$, which enforces that very small sales do not move the benchmark at all (useful when micro-transactions are easy to fabricate or are known to be uninformative).
- *Winsorized outcomes:* replace q_t by a robust statistic (e.g., a trimmed measure of verified transactions) before applying $w(\cdot)$, so that outliers in recorded outcomes do not create benchmark jumps.

These modifications illustrate a broader point: the platform can separately manage (i) what constitutes a valid "outcome" and (ii) how valid outcomes translate into reference mass. In applications, the first is often as important as the second.

A Stackelberg formulation for platform choice. To formalize design, we can view the platform as moving first: it commits to $w(\cdot)$ (and possibly to (W_1, R_1) as legacy parameters), anticipating the seller's dynamic best response. A generic objective that nests both efficiency and integrity concerns is

$$\max_{w \in \mathcal{W}} \ \mathbb{E}\left[\sum_{t=1}^{T} \Big(\mathrm{CS}(p_t, R_t) + p_t q_t\Big) \ - \ \lambda\, \mathcal{M}\big(\{p_t, q_t, R_t\}_{t \le T}\big)\right], \tag{47}$$

subject to the seller choosing $\{p_t\}$ to maximize $\Pi = \sum_t p_t q_t$ under the induced state transitions. Here $\mathcal{W}$ is a feasible class (e.g., power weights or capped weights), and $\mathcal{M}$ is a manipulation/deception functional. While $\mathcal{M}$ is inherently application-specific, the primitives in our model suggest natural proxies. One example is a penalty on reference sensitivity to posted prices in low-outcome states:

$$\mathcal{M} \ = \ \sum_{t=1}^{T} \mathbf{1}\{q_t \le \tilde{q}\} \left|\frac{\partial R_{t+1}}{\partial p_t}\right|, \tag{48}$$

which directly targets the mechanism by which inflated posted prices could influence future references. Outcome weighting already forces $\frac{\partial R_{t+1}}{\partial p_t} = 0$ when $q_t = 0$; the remaining design question is how close to zero we want this sensitivity to be when q_t is merely small.

This Stackelberg framing is useful even if one does not literally compute (47). It clarifies what the platform can and cannot do. Because $w(\cdot)$ affects the seller only through state transitions, design primarily reshapes *dynamic* incentives (how profitable it is to sacrifice current revenue to steer future R_t). It does not directly constrain the seller's feasible prices, nor does it require the platform to infer intent. In that sense, outcome-weighted design is a structural alternative to intent-based enforcement.

Practical interpretation and policy levers. From a policy perspective, one can interpret W_t as the platform's "evidentiary burden": a larger mass means the benchmark is harder to move. Increasing W_1 (longer lookback, more historical transactions, or simply adding a baseline prior) makes anchoring mechanically expensive from day one. Adjusting curvature (e.g., increasing γ) makes the benchmark depend more on periods where many consumers actually transact, which aligns the displayed statistic with what a typical consumer experiences rather than with what a seller posts in low-volume states. At the same time, very aggressive curvature can make the reference overly sensitive to promotion spikes, which may be undesirable if those spikes are transient or themselves strategically induced.

One implication is that design should be context-dependent. Categories with stable demand and frequent transactions can tolerate larger W_1 (strong inertia) without making the reference stale. Thin markets or new products may require lower W_1 or milder curvature so that the benchmark can "learn" quickly from limited data. More broadly, the platform can rationally choose heterogeneous $w(\cdot)$ across categories, or even adapt $w(\cdot)$ over a product's life cycle, so long as the updating rule remains interpretable and does not itself become a channel for deception.

Limitations of purely mechanical design. Finally, we emphasize what this approach does *not* solve. Outcome weighting eliminates the most blatant form of anchoring via unsold posted prices, but it cannot prevent all strategic behavior. A seller may still choose prices to trade off current revenue against future demand through R_t, and if low-cost "real" purchases can be engineered (e.g., through self-dealing, rebates, or collusive buying), then even outcome-weighted references can be manipulated—the manipulation simply becomes more costly and easier to conceptualize in terms of required mass. Moreover, when demand is censored, kinked, or subject to significant shocks, the platform's optimal choice of $w(\cdot)$ may need to account explicitly for statistical estimation error and for the possibility that the seller's best response is non-smooth. These considerations motivate the extensions in the next section, where asymmetric reference effects, censoring, and noise complicate both seller dynamics and platform design, and where numerical methods become valuable complements to the analytic bounds developed here.

7 Extensions and Limits

Our baseline analysis is deliberately spare: the demand response to reference prices is summarized by a reduced-form function $D(p, R)$, the platform's benchmark is updated mechanically from realized outcomes, and the seller observes the relevant state. In practice, each of these ingredients can fail in systematic ways. In this section we outline four extensions—asymmetric reference effects, censored demand, stochastic shocks, and partial observability—and explain where the analytic logic survives intact (often through the same accounting identities) and where we should expect to rely on numerical methods.

Asymmetric reference effects ($\eta_+ \neq \eta_-$): kinks, inaction regions, and piecewise policies. A common empirical regularity is that "losses" relative to the reference are more salient than "gains." In our notation, this corresponds to $\eta_- > \eta_+$ in the piecewise-linear specification

$$D(p, R) = \max\Big\{0,\ b - ap + \eta_+(R - p)_+ - \eta_-(p - R)_+\Big\}.$$

Intuitively, asymmetry introduces a wedge in the seller's incentives around $p = R$. When $\eta_- > \eta_+$, raising price above the displayed reference can sharply reduce demand, so the seller may optimally keep p_t at or below R_t even when a myopic monopoly price would exceed it. Conversely, when $p_t < R_t$ the seller enjoys a "gain" effect that is weaker, so cutting price far below R_t may not be as attractive as under symmetry.

Formally, asymmetry makes the seller's objective non-differentiable at $p = R$ even before we account for censoring at zero demand. The Bellman problem remains well-defined, but interior first-order conditions must be replaced by subgradient or complementary-slackness statements. In a typical period, the candidate optimizer belongs to one of three regions: (i) $p < R$ (gain region), (ii) $p > R$ (loss region), or (iii) $p = R$ (kink). The dynamic component compounds this piecewise structure because the state transition depends on $q_t = D(p_t, R_t)$, so the mapping from p_t to (R_{t+1}, W_{t+1}) inherits the kink. While one can sometimes characterize thresholds (e.g., an "inaction band" in which $p_t = R_t$ is optimal), such thresholds depend on continuation values that are themselves endogenous and time-varying, making clean closed forms rare outside very special parameterizations.

Two qualitative implications are robust. First, the no-free-anchoring logic does not hinge on symmetry: the identity linking reference movement to realized outcomes continues to hold whenever $w(0) = 0$, regardless of how D depends on (p, R). Second, asymmetry tends to *increase* the prevalence of boundary solutions and flat regions in policy functions, which is precisely where purely local (FOC-based) arguments are least informative. For this reason, even in otherwise linear environments, computing optimal policies with $\eta_+ \neq \eta_-$ is often best treated as a piecewise dynamic program evaluated numerically.

Censored demand and stockouts: state transitions with corners. Many markets exhibit censoring beyond the reference-effect kink. Our reduced form already allows for demand truncation at zero via $\max\{0, \cdot\}$, but in applications there may be additional constraints such as inventory limits, capacity constraints, or platform-side throttling. A simple representation is

$$q_t = \min\Big\{\bar{q}_t,\ D(p_t, R_t)\Big\},$$

where $\bar{q}_t$ is an exogenous (or endogenous) cap. Censoring matters because the platform's reference update is driven by realized outcomes, not latent demand. If q_t hits a cap, then the benchmark may fail to reflect high willingness-to-pay states, and the seller's incentive to "invest" in reference manipulation can be muted or distorted.

Analytically, censoring introduces additional corners: changes in p_t may no longer move q_t locally (when $q_t = 0$ or $q_t = \bar{q}_t$), which can create flat segments in the transition map for (R_{t+1}, W_{t+1}). This is not merely a technical nuisance. If w is increasing, then the leverage share $\alpha_t = \frac{w(q_t)}{W_t + w(q_t)}$ may become locally insensitive to p_t in censored regions, weakening the dynamic motive to steer R_{t+1} precisely when the market is thin (at $q_t \approx 0$) or constrained (at $q_t \approx \bar{q}_t$). In such environments, the seller's optimal policy can involve discrete jumps: a small price change that moves the system off a censoring boundary can abruptly change the marginal return to pricing through both current revenue and reference evolution.

Closed-form comparative statics are therefore fragile under censoring. Nonetheless, the economic lesson is clear: outcome-weighted benchmarks behave like *data-driven filters*. If the data are censored, the filter inherits the censoring. This creates a practical design implication: platforms that rely on outcome weighting should audit whether recorded outcomes are systematically truncated (by logistics, by ranking algorithms, or by inventory) and interpret benchmark sluggishness accordingly.

Stochastic shocks: risk-neutral dynamic control with endogenous volatility. A second departure from the baseline is randomness in realized outcomes. Even if demand is stable in expectation, day-to-day sales are noisy due to idiosyncratic traffic, competition, seasonality, and measurement error. A parsimonious specification is

$$q_t = \max\{0,\ D(p_t, R_t) + \varepsilon_t\}, \qquad \mathbb{E}[\varepsilon_t \mid \mathcal{F}_t] = 0,$$

with shocks ε_t conditionally independent given the seller's information $\mathcal{F}_t$. The seller's problem becomes a stochastic dynamic program:

$$V_t(R, W) = \max_{p \in [0, \bar{p}]} \mathbb{E}\Big[p\, q_t + V_{t+1}(R_{t+1}, W_{t+1}) \,\Big|\, R_t = R, W_t = W, p_t = p\Big],$$

where (R_{t+1}, W_{t+1}) are random through q_t.

Two new issues arise. First, the platform's benchmark becomes a *stochastic process* with state-dependent volatility because α_t is nonlinear in q_t (especially under convex w). Thus, even when $\mathbb{E}[q_t]$ is modest, the upper tail of the shock distribution can generate large benchmark jumps. Second, the seller internalizes that pricing affects not only the mean of q_t but also the distribution of next-period states; with nonlinear transitions, higher-order moments can matter for expected continuation value. Under risk neutrality, the criterion remains expected profit, but the expectation may depend sensitively on tail behavior, which is precisely what closed-form methods handle poorly.

In such settings, numerical methods are typically required for two reasons: (i) the transition kernel for (R_{t+1}, W_{t+1}) is generally not available in closed form once we impose truncation at zero and nonlinear weights $w(\cdot)$; and (ii) the value function can be highly non-quadratic even when D is linear, because the platform's update rule embeds a ratio $\frac{WR+w(q)p}{W+w(q)}$. Simulation-based value iteration (or policy iteration) becomes a natural tool: for each state (R, W) and candidate price p, we can approximate the expected continuation value by Monte Carlo draws of ε_t and then optimize over p.

Partial observability: hidden states and belief-based control. A final complication is that the seller may not observe the platform's state variables. In many platforms, sellers see the displayed reference R_t but not the mass W_t, and they may not know the precise weight function $w(\cdot)$, the lookback window, or data-cleaning rules. Likewise, the platform may observe only a noisy proxy of true transactions (e.g., verified purchases) while the seller observes gross orders. These informational frictions change the nature of the dynamic problem: the seller effectively controls prices under incomplete information about how today's outcomes translate into tomorrow's benchmark.

A tractable formulation treats (W_t) (and possibly the effective w) as latent and assumes the seller maintains a belief μ_t over hidden states given observed history (e.g., past R's, own prices, and realized sales if available). The seller's control problem becomes a partially observed Markov decision process (POMDP), where the sufficient state is no longer (R_t, W_t) but (R_t, μ_t). Even when the platform's update rule is known, belief updating can be nonlinear because R_{t+1} is a ratio and because q_t may be censored or noisy.

This extension highlights an important limit of purely mechanical benchmark design: even if the rule is manipulation-resistant in a full-information sense, opacity can create perceived manipulability (or perceived futility), which may distort seller behavior in unintended ways. Practically, this pushes toward *interpretable* and *auditable* rules: if a seller can forecast how R_t reacts to genuine sales, then the platform's benchmark is more likely to guide pricing and promotion in predictable ways rather than inducing guesswork-driven experimentation.

Where numerical methods become indispensable. Across these extensions, a common theme is that non-smoothness (kinks and censoring), nonlinear transitions (ratio updates), and stochasticity (shock-driven tails) interact to defeat closed-form analysis. Numerically, we have found three approaches especially useful.

First, grid-based dynamic programming on (R, W) can handle kinks transparently: we can compute V_t on a discretized state space and maximize over a discretized price set, avoiding reliance on derivatives. Second, simulation-based expectations are well-suited to stochastic shocks and complicated truncations; the main computational burden is managing variance, which can be reduced with common random numbers and antithetic sampling. Third, for high-dimensional variants (e.g., latent states or multiple products), approximate dynamic programming methods—parametric value-function approximation, fitted Q-iteration, or policy gradient methods—can provide scalable solutions, though at the cost of weaker guarantees and a need for validation.

We also emphasize a modeling limit that is not merely computational. Once we allow for self-dealing, rebates, or collusive transactions, the mapping from posted prices to "real" outcomes becomes endogenous in ways our reduced form does not capture. Outcome weighting then remains conceptually appealing, but the relevant object is not sales *per se* but *credible* sales. Incorporating credibility would require an additional layer (e.g., a fraud-detection technology or verification friction) and would naturally link benchmark design to enforcement.

Taken together, these extensions reinforce the main message while sharpening its scope. Outcome-weighted reference computation removes a particularly stark loophole—moving the benchmark without transacting—but it does not eliminate all strategic behavior, and it does not guarantee smooth or easily characterizable dynamics in empirically realistic environments. The next section therefore shifts from mechanism to implications: what these models suggest that platforms and regulators can measure, audit, and adjust in practice.

8 Discussion and Policy Implications

Reference prices are not merely descriptive statistics: once displayed, they become behavioral primitives that shape consumer response and, in turn, seller incentives. This dual role creates a classic design problem. If the reference is *too* responsive to posted prices, sellers can "manufacture" attractive discounts by first inflating the reference using non-transactional (or near-non-transactional) price postings. If the reference is *too* inert, it can cease to track meaningful market information and may mislead consumers in the opposite direction (e.g., by anchoring them to stale, unusually high past prices). The central lesson of our analysis is that outcome-weighted benchmark computation is a particularly transparent way to target the first failure mode: it ties benchmark movement to realized outcomes, so influence requires transactions rather than mere announcements.

A practical reading of the no-free-anchoring identity. The accounting identity

$$\Delta R_t \equiv R_{t+1} - R_t = \alpha_t(p_t - R_t), \qquad \alpha_t = \frac{w(q_t)}{W_t + w(q_t)} \in [0, 1),$$

has a direct policy interpretation. It says that the platform can make the displayed reference "hard to game" by choosing rules under which α_t is small whenever outcomes are negligible. In the baseline outcome-weighted update, α_t is mechanically small when either (i) the period's realized outcome q_t is small (because $w(q_t)$ is small), or (ii) accumulated history W_t is large (because the new observation has limited leverage). Put differently, the platform is implementing a data-aggregation principle that many auditors would find intuitive: if there is no credible mass behind a price, that price should not move the reference.

This is the sense in which outcome weighting provides a clean, easily communicable "anti-fake-discount" guarantee. A seller may post an extreme price, but unless consumers transact (or otherwise engage in the outcome being weighted), the posted price does not propagate into the benchmark. For regulators focused on deception, this guarantee is meaningful because it limits a specific channel of harm: displaying a reference that has been inflated by non-credible price postings.

Design levers for platforms: choosing $w(\cdot)$ and effective history mass. The platform's design choices map naturally into two levers: the shape of $w(\cdot)$ and the magnitude of the effective history mass W_t (which is governed by initialization W_1 and, in practice, by any lookback window or discounting rule the platform implements).

First, the weight function w governs how rapidly leverage α_t rises with outcomes. Within the family $w(q) = q^\gamma$, increasing γ makes the benchmark less sensitive to small outcomes and more sensitive to large outcomes. This is a tradeoff, not a free lunch. A larger γ improves manipulation-resistance against low-volume "wash" activity or thin-market opportunism, but it also amplifies benchmark jumps during large promotional events. In categories where promotions are frequent and high volume (so that large q_t are common), a highly convex w can generate reference volatility that consumers may interpret as instability or that sellers may experience as a moving target. Conversely, in thin categories where low volumes are common and manipulation risk is high, convex weighting may be desirable precisely because it forces benchmark movement to be supported by substantial realized activity.

Second, W_1 (and any operational analogue, such as an initial lookback mass) regulates the overall "memory" of the benchmark. A higher W_1 makes the reference more inert early on, reducing strategic incentives to invest in raising R_t, but it also slows adaptation for new products or new sellers. This suggests a category- and lifecycle-dependent policy: a platform may deliberately set smaller effective history for genuinely new items (to avoid anchoring consumers to irrelevant legacy prices) while imposing larger effective history once an item has established stable volume (to reduce dynamic incentives to manipulate). Importantly, our analysis predicts that dynamic incentives to steer R_t decay like $O(1/W_t)$ in tractable environments, so even modest growth in W_t can quickly reduce the marginal value of benchmark manipulation.

Guardrails beyond weighting: caps, trims, and "credible outcome" definitions. Outcome weighting is only as good as the definition of the outcome. If the platform weights *orders* that are easily reversed, or weights *clicks* that can be purchased, then q_t becomes an endogenous object that can itself be manipulated. In such environments, a robust implementation should combine outcome weighting with a credibility layer: verified purchases, chargeback-adjusted sales, fraud-detected engagement, or other outcomes that are costly to fake. Conceptually, the update rule should be driven by a "credible quantity" $\tilde{q}_t$, not raw activity.

Platforms can also add guardrails that preserve interpretability while limiting extreme leverage. One approach is to cap per-period influence by restricting $\alpha_t \le \bar{\alpha}$ for some design constant $\bar{\alpha} < 1$, effectively limiting the maximum one-step movement of the benchmark even under very large outcomes. Another approach is trimming or winsorizing on the price dimension, so that the contribution of a period is based on a bounded function of $(p_t - R_t)$; this can be useful in categories where pricing errors or outliers are common. These modifications preserve the core intuition of our mechanism—benchmark movement requires credible mass—while reducing sensitivity to rare but extreme events.

Transparency as a complement to manipulation-resistance. A subtle implication of our partial observability discussion is that opacity can itself be distortionary. If sellers do not understand how R_t responds to real outcomes, they may engage in trial-and-error experimentation that creates unnecessary volatility, or they may incorrectly infer that the benchmark is manipulable and attempt to manipulate it. For regulators, opacity also makes enforcement difficult because it blurs the line between a benign statistic and a designed persuasion tool.

We therefore view transparency as a complement to outcome weighting. At minimum, platforms can disclose (i) the relevant outcome used (e.g., verified units sold), (ii) the functional form class for weighting (e.g., "volume-weighted average"), and (iii) the effective horizon or mass (e.g., a lookback window or decay rate). This does not require revealing proprietary details at a level that would enable gaming; indeed, the no-free-anchoring identity suggests that when $w(0) = 0$, many common gaming attempts are neutralized even when the rule is public.

Auditing the benchmark: simple tests implied by the model. Our framework yields testable predictions that can be operationalized as audits. The most direct is a "no-free-anchoring" test: in data, periods with negligible realized outcomes should have negligible benchmark movement. A simple diagnostic computes an implied leverage share

$$\widehat{\alpha}_t := \frac{R_{t+1} - R_t}{p_t - R_t} \quad (\text{for } p_t \neq R_t),$$

and examines whether $\widehat{\alpha}_t$ is (i) near zero when q_t is near zero and (ii) increasing with q_t in a manner consistent with the platform's stated weighting policy. Systematic benchmark movement in periods with $q_t \approx 0$ is a red flag: it can indicate that the platform is not truly outcome-weighting (despite claims), that q_t is mismeasured, or that the platform is using additional signals (e.g., posted prices) in ways that undermine the intended protection.

A second audit targets volatility and promotion sensitivity. Under convex weighting, large-q_t events receive disproportionate influence, so $|R_{t+1} - R_t|$ should spike around promotions, featured placements, or other traffic shocks. This is not necessarily undesirable, but it becomes a compliance issue if the platform represents the reference as "typical" or "regular" while it is in fact heavily shaped by short-lived events. An audit can therefore compare the distribution of ΔR_t across ordinary periods and promotion-tagged periods; unusually large and persistent shifts concentrated in promotions suggest that the benchmark is effectively a "promotion-weighted" statistic.

A third audit focuses on cross-seller and lifecycle variation. Because leverage falls with W_t, newly listed items (or items with short histories) should exhibit greater benchmark responsiveness and, consequently, stronger incentives for sellers to attempt steering. If manipulation complaints are concentrated among new products, this is consistent with the mechanism and suggests targeted safeguards: larger initial W_1 in high-risk categories, stricter verification for early transactions, or temporary caps on α_t until sufficient credible history accumulates.

Enforcement and legal framing: when does a reference become deceptive? From a regulatory standpoint, a key question is whether a displayed reference represents an informative statistic or a designed persuasion instrument. Our model does not resolve the legal boundary, but it clarifies what is technically feasible and what can be verified. If a platform can commit to a benchmark that (i) is mechanically linked to credible outcomes and (ii) demonstrably exhibits negligible response to non-transactional postings, then it is easier to argue that the reference reflects market experience rather than strategic presentation. Conversely, if the benchmark responds materially to posted prices absent credible outcomes, then the platform is effectively enabling sellers to create artificial "regular prices," which is precisely the pattern regulators often target in false-reference-price enforcement.

We also emphasize that enforcement should be aligned with the credibility of q_t. If the platform's benchmark weights transactions that can be cheaply faked (self-dealing, rebates that net out the price, or collusive purchases), then an outcome-weighted rule can still be manipulated, but now through the outcome channel. In that case, compliance requires not only an aggregation rule but also monitoring and deterrence: anomaly detection for unusual purchase patterns, verification frictions, and penalties that raise the cost of fabricating credible outcomes.

Limitations and a pragmatic takeaway. Two limitations deserve emphasis. First, a benchmark that is manipulation-resistant is not automatically welfare-optimal: the platform may care about consumer surplus, seller revenue, and trust, and these objectives may trade off against benchmark responsiveness. Second, real platforms are multi-product, algorithmic, and subject to feedback loops (ranking affects sales, which affects R_t, which affects demand, and so on). Our accounting logic survives these complications, but precise welfare conclusions generally require richer modeling and, often, quantitative calibration.

The pragmatic takeaway is nonetheless crisp. If platforms want a reference price that is both interpretable and hard to game, they should tie benchmark movement to credible realized outcomes, choose weighting schemes that reflect category-specific risk (thin markets versus promotion-heavy markets), and make the rule sufficiently transparent to support auditing. For regulators, the same structure offers actionable tests: verify that low-outcome periods have low benchmark impact, that stated weighting policies match observed leverage patterns, and that the outcome being weighted is sufficiently credible to support the reference's consumer-facing claims.