

Appendix A: An Example Value Function for LVA

Suppose the concave envelope takes the form $\hat{v}(\ell) = \theta\sqrt{\ell}$ with $\theta > 0$. Then $\hat{v}'(\ell) = \theta/(2\sqrt{\ell})$, and the interior target in (21) is $\ell^\circ(D) = \left(\frac{\omega\theta}{2cD}\right)^2$. Substituting into (23) yields the explicit Stage-II rule

$$\tilde{\ell}(D, Q) = \left[\max \left\{ \frac{Q}{D}, \left(\frac{\omega\theta}{2cD} \right)^2 \right\} \right]_x, \quad D = \min\{\alpha r - \beta p + \epsilon, y\}.$$

In this square-root case, both the supplementation target and the marginal scarcity cost admit transparent power laws. First, $\ell^\circ(D) \propto D^{-2}$: when accepted reservations increase, the LVA target value per bag drops rapidly, reflecting diminishing marginal satisfaction from additional value. Second, for $D > 0$, the marginal scarcity cost in (22) reduces to

$$m(D, Q) = \begin{cases} \frac{\omega^2\theta^2}{4cD^2}, & Q \leq D\ell^\circ(D), \\ \frac{\omega\theta}{2} \cdot \frac{\sqrt{Q}}{D^{3/2}}, & D\ell^\circ(D) < Q < D\bar{x}, \\ 0, & Q \geq D\bar{x}. \end{cases}$$

The first regime ($Q \leq D\ell^\circ(D)$) corresponds to topping up to the target; the marginal scarcity cost is then exactly the per-bag top-up cost $c\ell^\circ(D)$, which decreases in D because the target itself falls with demand. The second regime corresponds to using surplus only; here $m(D, Q)$ captures the reputational loss from diluting a fixed surplus pool across more bags, and its $D^{-3/2}$ dependence shows that the dilution penalty becomes much smaller when demand is large. The square-root specification provides closed-form Stage-II targets and scarcity costs. The joint Stage-I pair $(\tilde{p}, \tilde{y})$ is then obtained through the bounded search in (20).

Appendix B: Proofs

B.1. Proof of Proposition 1

Let $J_k(r)$ denote the optimal k -period discounted value function with terminal value 0. Then $J_0 \equiv 0$ and, for $k \geq 0$,

$$J_{k+1}(r) = \max_{\substack{p \in [0, \bar{p}] \\ y \in [\underline{y}, \bar{y}]}} \mathbb{E}_\epsilon \left[pD(r, p, y, \epsilon) + \mathbb{E}_Q \left[\max_{\phi(\cdot)} \left\{ -C[\phi(\cdot), D(r, p, y, \epsilon), Q] + \gamma J_k(\delta(V[\phi] - p) + (1 - \delta)r) \right\} \right] \right].$$

We first show by induction that each J_k is weakly increasing in r .

Fix $r' > r$ and $k \geq 0$. Let $(p_k^*(r), y_k^*(r))$ be an optimal Stage-I pair for $J_{k+1}(r)$. For each (ϵ, Q) , let $\phi_k^*(r, \epsilon, Q)$ be an optimizer of the inner Stage-II problem at $(r, p_k^*(r), y_k^*(r), \epsilon, Q)$, and let

$$\ell^*(\epsilon, Q) := \int_{\mathcal{X}} x \phi_k^*(r, \epsilon, Q)(x) dx$$

denote its mean. At state r' , consider the feasible price

$$p' := \min \left\{ \bar{p}, p_k^*(r) + \frac{\alpha}{\beta}(r' - r) \right\}$$

and the same reservation cap $y_k^*(r)$. After observing the same realization (ϵ, Q) , use the same Stage-II allocation $\phi_k^*(r, \epsilon, Q)$. Define $d := \alpha r - \beta p_k^*(r) + \epsilon$, $d' := \alpha r' - \beta p' + \epsilon$, $D := \min\{d, y_k^*(r)\}$, and $D' := \min\{d', y_k^*(r)\}$. By construction, $d' \geq d$ and therefore $D' \geq D$ for every ϵ . Moreover, by the active-cycle support restriction in Section 3.3, $d > 0$ under the policies considered. For the continuation value, the next-period ratings are

$$r_{\text{next}} = (1 - \delta)r + \delta(V[\phi_k^*] - p_k^*(r)) \quad \text{and} \quad r'_{\text{next}} = (1 - \delta)r' + \delta(V[\phi_k^*] - p'),$$

so

$$r'_{\text{next}} - r_{\text{next}} = (1 - \delta)(r' - r) - \delta(p' - p_k^*(r)) \geq \frac{(1 - \delta)\beta - \delta\alpha}{\beta}(r' - r) \geq 0,$$

where the last inequality uses Assumption 1. By the induction hypothesis, J_k is weakly increasing, so $\gamma J_k(r'_{\text{next}}) \geq \gamma J_k(r_{\text{next}})$. For the one-period payoff, write

$$\Delta\Pi := p'D' - p_k^*(r)D - c[(\ell^*D' - Q)^+ - (\ell^*D - Q)^+].$$

Because $x \mapsto x^+$ is 1-Lipschitz and $D' \geq D$, we have $(\ell^*D' - Q)^+ - (\ell^*D - Q)^+ \leq \ell^*(D' - D)$, hence $\Delta\Pi \geq (p' - c\ell^*)D' - (p_k^*(r) - c\ell^*)D$. If $p' = p_k^*(r) + \frac{\alpha}{\beta}(r' - r) \leq \bar{p}$, then $D' = D$, so $\Delta\Pi \geq (p' - p_k^*(r))D \geq 0$. If instead $p' = \bar{p} < p_k^*(r) + \frac{\alpha}{\beta}(r' - r)$, then Assumption 1 and $\ell^* \leq \bar{x}$ imply $\bar{p} - c\ell^* \geq 0$, and therefore

$$\Delta\Pi \geq (\bar{p} - c\ell^*)D' - (p_k^*(r) - c\ell^*)D \geq (\bar{p} - c\ell^*)D - (p_k^*(r) - c\ell^*)D = (\bar{p} - p_k^*(r))D \geq 0.$$

Thus, for every realization (ϵ, Q) , the constructed feasible decision at state r' yields a weakly larger current payoff and a weakly larger continuation value than the optimal decision at state r . Taking expectations gives $J_{k+1}(r') \geq J_{k+1}(r)$. Since $J_0 \equiv 0$, every J_k is weakly increasing. It remains to pass to the infinite horizon. Fix an initial state r and any admissible policy π . Let $R_t^\pi := p_t^\pi D_t^\pi - C[\phi_t^\pi(\cdot), D_t^\pi, Q_t]$ denote the period- t payoff, and let $\ell_t^\pi := \int_{\mathcal{X}} x \phi_t^\pi(x) dx \in \mathcal{X}$ be the corresponding mean bag value. Since $V[\phi_t^\pi] \leq v(\bar{x})$ and $p_t^\pi \geq 0$,

$$r_{t+1}^\pi = (1 - \delta)r_t^\pi + \delta(V[\phi_t^\pi] - p_t^\pi) \leq (1 - \delta)r_t^\pi + \delta v(\bar{x}),$$

which implies $r_t^\pi \leq \bar{r}(r) := \max\{r, v(\bar{x})\}$ for all t . Under the active-cycle demand restriction,

$$|R_t^\pi| \leq p_t^\pi D_t^\pi + c(\ell_t^\pi D_t^\pi - Q_t)^+ \leq (\bar{p} + c\bar{x})D_t^\pi \leq (\bar{p} + c\bar{x})\bar{y}.$$

Hence $M := (\bar{p} + c\bar{x})\bar{y} < \infty$ satisfies $\mathbb{E}|R_t^\pi| \leq M$ for all t and all admissible π . Therefore,

$$\sup_{\pi} \mathbb{E} \left[\sum_{t=k}^{\infty} \gamma^t |R_t^\pi| \middle| r_0 = r \right] \leq \frac{\gamma^k}{1 - \gamma} M \xrightarrow{k \rightarrow \infty} 0.$$

So the k -period value functions converge pointwise to the infinite-horizon value function: $J_k(r) \rightarrow J^*(r)$. Since each J_k is weakly increasing, the limit $J^*(r)$ is also weakly increasing. $\square$

B.2. Proof of Corollary 1

Proof. Fix (r, p, D, Q) and let $\bar{\phi}(\cdot)$ be any optimal Stage-II distribution. Let

$$\bar{\ell} := \int_{\mathcal{X}} x \bar{\phi}(x) dx$$

be its mean bag value. Among all distributions on $\mathcal{X}$ with mean $\bar{\ell}$, choose one that maximizes $V[\phi]$; denote it by $\phi^*(\cdot)$. Such a distribution exists because the feasible set is tight on the compact interval $\mathcal{X}$ and the objective is continuous.

The Stage-II objective is

$$W(\phi(\cdot); r, p, D, Q) = -C[\phi(\cdot), D, Q] + \gamma J^*(\delta(V[\phi(\cdot)] - p) + (1 - \delta)r).$$

Because ϕ^* and $\bar{\phi}$ have the same mean, their supplementation costs are identical: $C[\phi^*(\cdot), D, Q] = C[\bar{\phi}(\cdot), D, Q]$. By construction, $V[\phi^*(\cdot)] \geq V[\bar{\phi}(\cdot)]$. The common price term $-p$ appears in both rating arguments, so Proposition 1 implies

$$J^*(\delta(V[\phi^*(\cdot)] - p) + (1 - \delta)r) \geq J^*(\delta(V[\bar{\phi}(\cdot)] - p) + (1 - \delta)r).$$

Thus $W(\phi^*(\cdot); r, p, D, Q) \geq W(\bar{\phi}(\cdot); r, p, D, Q)$. Since $\bar{\phi}$ was optimal, ϕ^* is also optimal and maximizes $V[\phi]$ among distributions with the same mean. $\square$

B.3. Proof of Proposition 2 and Corollary 2

Proof of Proposition 2. Interpret ϕ in (12) as a probability measure on $[\underline{x}, \bar{x}]$. When ϕ admits a density with respect to Lebesgue measure, we use the notation $\phi(x) dx$ as shorthand. Fix $\ell \in [\underline{x}, \bar{x}]$, and let ϕ be any feasible distribution satisfying $\int_{\mathcal{X}} x \phi(dx) = \ell$.

Because $\hat{v}$ is the upper concave envelope of v , we have $\hat{v}(x) \geq v(x) \forall x \in [\underline{x}, \bar{x}]$. Hence

$$V[\phi] = \int_{\mathcal{X}} v(x) \phi(dx) \leq \int_{\mathcal{X}} \hat{v}(x) \phi(dx).$$

Since $\hat{v}$ is concave and ϕ has mean ℓ , Jensen's inequality implies

$$\int_{\mathcal{X}} \hat{v}(x) \phi(dx) \leq \hat{v}\left(\int_{\mathcal{X}} x \phi(dx)\right) = \hat{v}(\ell).$$

Therefore every feasible ϕ satisfies $V[\phi] \leq \hat{v}(\ell)$. So the optimal value of (12) is at most $\hat{v}(\ell)$. It remains to construct a feasible distribution attaining $\hat{v}(\ell)$. Let $\mathcal{O} := \{x \in [\underline{x}, \bar{x}] : v(x) = \hat{v}(x)\}$. If $\ell \in \mathcal{O}$, then the degenerate distribution $\phi^*(\cdot) = \Delta_{\ell}$ is feasible, has mean ℓ , and satisfies $V[\phi^*] = v(\ell) = \hat{v}(\ell)$. Hence Δ_{ℓ} is optimal. Now suppose $\ell \notin \mathcal{O}$, equivalently $\hat{v}(\ell) > v(\ell)$. Since v is continuous, the set $\{x \in [\underline{x}, \bar{x}] : \hat{v}(x) > v(x)\}$ is open, so ℓ belongs to some open interval $(x^-, x^+) \subseteq [\underline{x}, \bar{x}]$. By the standard geometry of the smallest concave majorant on an interval, $\hat{v}$ is affine on $[x^-, x^+]$ and satisfies $\hat{v}(x^-) = v(x^-)$ and $\hat{v}(x^+) = v(x^+)$. Define $a := \frac{x^+ - \ell}{x^+ - x^-}$ and $1 - a := \frac{\ell - x^-}{x^+ - x^-}$, and let $\phi^*(\cdot) := a \Delta_{x^-} + (1 - a) \Delta_{x^+}$. Then ϕ^* is feasible and has mean

$$\int_{\mathcal{X}} x \phi^*(dx) = ax^- + (1 - a)x^+ = \ell.$$

Using the contact conditions and the affinity of $\hat{v}$ on $[x^-, x^+]$, we obtain

$$V[\phi^*] = av(x^-) + (1 - a)v(x^+) = a\hat{v}(x^-) + (1 - a)\hat{v}(x^+) = \hat{v}(\ell).$$

Thus ϕ^* is optimal. We conclude that the optimal value of (12) is $\hat{v}(\ell)$ and that there exists an optimal solution supported on at most two points. This selected optimizer is the $\phi^*(\cdot | \ell)$ stated in the proposition. Because v is nondecreasing, its upper concave envelope $\hat{v}$ is also nondecreasing. $\square$

Proof of Corollary 2. If $v(\cdot)$ is concave on $\mathcal{X}$, then its upper concave envelope coincides with v itself: $\hat{v}(\ell) = v(\ell)$ for all $\ell \in \mathcal{X}$. Therefore $\ell \in \mathcal{O}$ for every $\ell \in \mathcal{X}$, and the degenerate distribution $\phi^*(\cdot | \ell) = \Delta_{\ell}$ is optimal by the first case of the proof of Proposition 2. $\square$

B.4. Proof of Proposition 3

Proof. Fix (r, p, D, Q) and define the objective

$$\Psi(\ell; r, p, D, Q) := \gamma J^*(\delta(\hat{v}(\ell) - p) + (1 - \delta)r) - c(\ell D - Q)^+, \quad \ell \in \mathcal{X}.$$

Let $\mathcal{L}^*(r, p, D, Q) := \arg \max_{\ell \in \mathcal{X}} \Psi(\ell; r, p, D, Q)$ denote the maximizer set. When $\mathcal{L}^*$ is not a singleton, we use the maximal selector $\bar{\ell}^*(r, p, D, Q) := \sup \mathcal{L}^*(r, p, D, Q)$. The maximizer set is nonempty because $\mathcal{X}$ is compact and the objective is upper semicontinuous in ℓ ; continuity of J^* follows from Lemma C.1 below.

Part 1 (monotonicity in Q). Fix (r, p, D) and define $g(\ell, Q) := -c(\ell D - Q)^+$. For any $Q_2 \geq Q_1$, consider

$$\Delta(\ell) := g(\ell, Q_2) - g(\ell, Q_1) = -c(\ell D - Q_2)^+ + c(\ell D - Q_1)^+.$$

By the active-cycle assumption, $D > 0$. Using the identity

$$(\ell D - Q)^+ = \begin{cases} 0, & \ell \leq Q/D, \\ \ell D - Q, & \ell > Q/D, \end{cases}$$

we obtain

$$\Delta(\ell) = \begin{cases} 0, & \ell \leq Q_1/D, \\ c(\ell D - Q_1), & Q_1/D < \ell \leq Q_2/D, \\ c(Q_2 - Q_1), & \ell > Q_2/D, \end{cases}$$

which is nondecreasing in ℓ . Hence $g(\ell, Q)$ has increasing differences in (ℓ, Q) . Because the remaining term $\gamma J^*(\delta(\hat{v}(\ell) - p) + (1 - \delta)r)$ does not depend on Q , $\Psi(\ell; r, p, D, Q)$ also has increasing differences in (ℓ, Q) . By Topkis' Monotonicity Theorem, the maximal selector $\bar{\ell}^*(r, p, D, Q)$ is nondecreasing in Q . This proves (ii).

Part 2 (monotonicity in D). Fix (r, p, Q) and view $g(\ell, D) := -c(\ell D - Q)^+$ for $D > 0$. For $D_2 \geq D_1 > 0$, define

$$\tilde{\Delta}(\ell) := g(\ell, D_2) - g(\ell, D_1) = -c(\ell D_2 - Q)^+ + c(\ell D_1 - Q)^+.$$

For any fixed $\ell \in \mathcal{X}$, the mapping $D \mapsto c(\ell D - Q)^+$ is convex and nondecreasing, and its right-derivative equals $c\ell \mathbf{1}\{\ell D \geq Q\}$. Consequently, for $D_2 \geq D_1$,

$$c(\ell D_2 - Q)^+ - c(\ell D_1 - Q)^+ = c \int_{D_1}^{D_2} \ell \mathbf{1}\{\ell s > Q\} ds$$

is nondecreasing in ℓ , which implies $\tilde{\Delta}(\ell)$ is nonincreasing in ℓ .

Therefore $g(\ell, D)$ has decreasing differences in (ℓ, D) on $D > 0$. Since the continuation-value term is independent of D , Ψ has decreasing differences in (ℓ, D) . Applying Topkis' Monotonicity Theorem yields that the maximal selector $\bar{\ell}^*(r, p, D, Q)$ is nonincreasing in D for $D > 0$. This proves (i).

Part 3 (monotonicity in p and r). Fix (D, Q) and note that

$$\Psi(\ell; r, p, D, Q) = \gamma J^*(\delta(\hat{v}(\ell) - p) + (1 - \delta)r) - c(\ell D - Q)^+.$$

Assume now that J^* is concave, as stated in Proposition 3. Because $\hat{v}$ is nondecreasing, for any $p_2 \geq p_1$ we can write

$$\Psi(\ell; r, p_2, D, Q) - \Psi(\ell; r, p_1, D, Q) = \gamma \left(J^*(x(\ell) - \delta(p_2 - p_1)) - J^*(x(\ell)) \right),$$

where $x(\ell) := \delta(\hat{v}(\ell) - p_1) + (1 - \delta)r$ is nondecreasing in ℓ . For any fixed $\Delta \geq 0$, concavity of J^* implies that the map $z \mapsto J^*(z - \Delta) - J^*(z)$ is nondecreasing. Hence the display above is nondecreasing in ℓ , so Ψ has increasing differences in (ℓ, p) . Similarly, for any $r_2 \geq r_1$,

$$\Psi(\ell; r_2, p, D, Q) - \Psi(\ell; r_1, p, D, Q) = \gamma \left(J^*(y(\ell) + (1 - \delta)(r_2 - r_1)) - J^*(y(\ell)) \right),$$

where $y(\ell) := \delta(\hat{v}(\ell) - p) + (1 - \delta)r_1$ is nondecreasing in ℓ . For any fixed $\Delta \geq 0$, concavity of J^* implies that the map $z \mapsto J^*(z + \Delta) - J^*(z)$ is nonincreasing. Therefore the display above is nonincreasing in ℓ , so Ψ has decreasing differences in (ℓ, r) .

The cost term does not depend on p or r . Applying Topkis' Monotonicity Theorem, the maximal selector $\bar{\ell}^*$ can be chosen to be nondecreasing in p and nonincreasing in r . This proves (iii) and (iv). $\square$

Appendix C: Proofs for Section 5

All proofs use the active-cycle support condition stated in Section 3.3; equivalently, one may restrict the price set to state-dependent admissible prices that keep latent demand positive.

C.1. Proof of Proposition 4

Proof. Fix a rating r . The proof proceeds in four steps.

Step 1. Substitute the affine value approximation. Let $\tilde{J}(r) = \tilde{J}_0 + kr$ and define $\omega := \gamma\delta k$. By assumption, $\omega > 0$; recall also that $\hat{v}$ is concave and increasing. Evaluating the Bellman operator at $\tilde{J}$ and collecting terms gives

$$(\mathcal{T}\tilde{J})(r) = \gamma\tilde{J}_0 + \gamma k(1 - \delta)r + \max_{\substack{0 \leq p \leq \bar{p} \\ y \leq y \leq \bar{y}}} \mathbb{E}_{\epsilon, Q} \left[pD(r, p, y, \epsilon) - \omega p + \max_{\ell \in \mathcal{X}} \{ \omega \hat{v}(\ell) - c(\ell D(r, p, y, \epsilon) - Q)^+ \} \right].$$

Since $\gamma\tilde{J}_0 + \gamma k(1 - \delta)r$ does not depend on the controls, the control problem decomposes into a Stage-II choice of ℓ given (r, p, y, ϵ, Q) and a Stage-I choice of (p, y) . Define the maximized Stage-II value

$$\Phi(D, Q) := \max_{\ell \in \mathcal{X}} \{ \omega \hat{v}(\ell) - c(\ell D - Q)^+ \}.$$

Substituting the maximized Stage-II value back into the LVA objective gives the Stage-I problem

$$(\tilde{p}(r), \tilde{y}(r)) \in \arg \max_{\substack{0 \leq p \leq \bar{p} \\ y \leq y \leq \bar{y}}} \mathbb{E}_{\epsilon, Q} \left[pD(r, p, y, \epsilon) - \omega p + \Phi(D(r, p, y, \epsilon), Q) \right]. \quad (\text{C.1})$$

Step 2. Stage-II rule. Fix (r, p, y, ϵ, Q) and write $D := D(r, p, y, \epsilon)$. By the active-cycle assumption, $D > 0$. Define

$$\Psi(\ell; D, Q) := \omega \hat{v}(\ell) - c(\ell D - Q)^+, \quad \ell \in \mathcal{X}.$$

If $Q/D \geq \bar{x}$, then every feasible $\ell \in \mathcal{X}$ satisfies $\ell \leq Q/D$, so $(\ell D - Q)^+ = 0$ for all feasible ℓ and $\Psi(\ell; D, Q) = \omega \hat{v}(\ell)$. Because $\hat{v}$ is increasing, the maximizer is $\bar{\ell} = \bar{x}$, which agrees with (23).

Hence suppose $Q/D < \bar{x}$. On the no-supplement region $\{\ell \in \mathcal{X} : \ell \leq Q/D\}$, we have $(\ell D - Q)^+ = 0$, so $\Psi(\ell; D, Q) = \omega \hat{v}(\ell)$, which is increasing in ℓ . Thus, whenever this region is nonempty, its best feasible point is $\ell = Q/D$. On the supplementation region $\{\ell \in \mathcal{X} : \ell \geq Q/D\}$, we have

$$\Psi(\ell; D, Q) = \omega \hat{v}(\ell) - cD\ell + cQ.$$

The function $g_D(\ell) := \omega \hat{v}(\ell) - cD\ell$ is concave. Let $\ell^\circ(D)$ be its maximal optimizer over $\mathcal{X}$, as defined immediately before Proposition 4. Hence $g_D(\ell)$ is weakly increasing up to $\ell^\circ(D)$ and weakly decreasing thereafter, so its maximal optimizer over the restricted feasible interval

$$\mathcal{X} \cap [Q/D, \infty) = [\max\{\underline{x}, Q/D\}, \bar{x}]$$

is the projection of $\ell^\circ(D)$ onto that interval, namely

$$\left[\max \left\{ \frac{Q}{D}, \ell^\circ(D) \right\} \right]_{\mathcal{X}}.$$

When $Q/D \in \mathcal{X}$, the no-supplement and supplementation objectives coincide at $\ell = Q/D$, so this restricted-region maximizer weakly dominates the best no-supplement point. When $Q/D < \underline{x}$, the no-supplement region is empty. Therefore,

$$\tilde{\ell}(D, Q) = \left[\max \left\{ \frac{Q}{D}, \ell^\circ(D) \right\} \right]_{\mathcal{X}},$$

proving (23).

Step 3. Marginal scarcity cost. At points where Φ is differentiable in D , define the marginal scarcity cost as $m(D, Q) := -\partial_D \Phi(D, Q)$. The three branches follow from the envelope theorem. If $Q < D\ell^\circ(D)$, then the Stage-II optimizer is $\ell = \ell^\circ(D)$ and

$$\Phi(D, Q) = \omega \hat{v}(\ell^\circ(D)) - c(D\ell^\circ(D) - Q),$$

so $\Phi_D(D, Q) = -c\ell^\circ(D)$. If $D\ell^\circ(D) < Q < D\bar{x}$, then the optimizer is $\ell = Q/D$ and $\Phi(D, Q) = \omega \hat{v}(Q/D)$, so

$$\Phi_D(D, Q) = -\omega \hat{v}'\left(\frac{Q}{D}\right) \frac{Q}{D^2}.$$

If $Q > D\bar{x}$, then the optimizer is $\ell = \bar{x}$ and $\Phi(D, Q) = \omega \hat{v}(\bar{x})$, hence $\Phi_D(D, Q) = 0$. Therefore, the marginal scarcity cost is exactly (22).

Step 4. Derive the Stage-I FOCs. Under the maintained regularity in Section 3.3,

$$\frac{\partial D(r, p, y, \epsilon)}{\partial p} = -\beta \mathbf{1}_{\{d(r, p, \epsilon) < y\}}, \quad \frac{\partial D(r, p, y, \epsilon)}{\partial y} = \mathbf{1}_{\{d(r, p, \epsilon) > y\}}.$$

Using $\Phi_D(D, Q) = -m(D, Q)$ in (C.1), the first-order condition with respect to price is

$$0 = \mathbb{E}_{\epsilon, Q} [D(r, p, y, \epsilon) - \omega - \beta(p - m(D(r, p, y, \epsilon), Q)) \mathbf{1}_{\{d(r, p, \epsilon) < y\}}].$$

On the event $\{d(r, p, \epsilon) < y\}$, $D(r, p, y, \epsilon) = d(r, p, \epsilon)$. Evaluating at $(p, y) = (\tilde{p}, \tilde{y})$ gives the price first-order condition. The first-order condition with respect to the cap is

$$0 = \mathbb{E}_{\epsilon, Q} [(p - m(D(r, p, y, \epsilon), Q)) \mathbf{1}_{\{d(r, p, \epsilon) > y\}}].$$

On the event $\{d(r, p, \epsilon) > y\}$, $D(r, p, y, \epsilon) = y$. Evaluating at $(p, y) = (\tilde{p}, \tilde{y})$ gives the cap first-order condition.

At $(p, y) = (\tilde{p}, \tilde{y})$, the cap first-order condition factors into

$$\Pr(d(r, \tilde{p}, \epsilon) > \tilde{y}) (\tilde{p} - \mathbb{E}_Q[m(\tilde{y}, Q)]) = 0,$$

and division by the positive leading probability gives $\tilde{p} = \mathbb{E}_Q[m(\tilde{y}, Q)]$. For comparison, when the cap is fixed at a nonbinding upper bound, $D(r, \tilde{p}, \tilde{y}, \epsilon) = d(r, \tilde{p}, \epsilon)$ and $\mathbf{1}_{\{d(r, \tilde{p}, \epsilon) < \tilde{y}\}} = 1$. The price condition becomes

$$\alpha r - \beta \tilde{p} - \omega - \beta \tilde{p} + \beta \mathbb{E}_{\epsilon, Q}[m(D(r, \tilde{p}, \tilde{y}, \epsilon), Q)] = 0,$$

which rearranges to (25). □

C.2. Proof of Proposition 5

Proof. Let $W := \bar{r} - \underline{r}$, using the invariant analytical rating interval defined in Section 3.3. If $W = 0$, choose $\tilde{J} = J^*$ on the singleton state space and the result is immediate. Hence suppose $W > 0$. First we show that the optimal value function is Lipschitz continuous in the rating state on this interval.

LEMMA C.1. Assume the feasible sets satisfy $p \in [0, \bar{p}]$, $y \in [\underline{y}, \bar{y}]$, and $\ell \in \mathcal{X}$. Then the optimal infinite-horizon value function J^* is Lipschitz continuous in r with constant

$$L := \frac{\alpha(\bar{p} + c\bar{x})}{1 - \gamma(1 - \delta)}. \quad (\text{C.2})$$

In particular, for any $r, r' \in [\underline{r}, \bar{r}]$, there is $|J^*(r') - J^*(r)| \leq L|r' - r|$.

Proof of Lemma C.1. We use value iteration and induction. Define the finite-horizon value functions $J_0^*(r) \equiv 0$ and, for $k \geq 0$,

$$J_{k+1}^*(r) = \max_{\substack{0 \leq p \leq \bar{p} \\ \underline{y} \leq y \leq \bar{y}}} \mathbb{E}_\epsilon \left[pD(r, p, y, \epsilon) + \mathbb{E}_Q \max_{\ell \in \mathcal{X}} \left\{ \gamma J_k^*(\delta(\hat{v}(\ell) - p) + (1 - \delta)r) - c(\ell D(r, p, y, \epsilon) - Q)^+ \right\} \right].$$

By the contraction property of the Bellman operator, $J_k^*(r) \rightarrow J^*(r)$ as $k \rightarrow \infty$ for each r . We claim that for all $k \geq 0$ and all $r, r' \in [\underline{r}, \bar{r}]$,

$$|J_k^*(r') - J_k^*(r)| \leq L|r' - r|. \quad (\text{C.3})$$

The base case $k = 0$ is immediate since J_0^* is constant.

Fix $k \geq 0$ and assume (C.3) holds for J_k^* . Let $r' > r$ and denote $\Delta r := r' - r > 0$. Let $(p_k^*(r'), y_k^*(r'))$ be an optimal Stage-I pair attaining $J_{k+1}^*(r')$. For each realization of (ϵ, Q) , let $\ell_k^*(r', p_k^*(r'), y_k^*(r'), \epsilon, Q)$ be a measurable maximizer of the inner problem at state r' and pair $(p_k^*(r'), y_k^*(r'))$. For notational brevity, write $\ell_k^*(r', \epsilon, Q) := \ell_k^*(r', p_k^*(r'), y_k^*(r'), \epsilon, Q)$.

Since $J_{k+1}^*(r)$ is the supremum over all admissible controls, it is at least the value obtained by using the same controls $(p_k^*(r'), y_k^*(r'), \ell_k^*(r', \epsilon, Q))$ at state r . Define

$$D'_\epsilon := D(r', p_k^*(r'), y_k^*(r'), \epsilon), \quad D_\epsilon := D(r, p_k^*(r'), y_k^*(r'), \epsilon).$$

Therefore, defining

$$x_{r'} := \delta(\hat{v}(\ell_k^*(r', \epsilon, Q)) - p_k^*(r')) + (1 - \delta)r' \quad \text{and} \quad x_r := \delta(\hat{v}(\ell_k^*(r', \epsilon, Q)) - p_k^*(r')) + (1 - \delta)r,$$

we obtain

$$\begin{aligned} J_{k+1}^*(r') - J_{k+1}^*(r) &\leq \mathbb{E}_\epsilon \left[p_k^*(r')D'_\epsilon - p_k^*(r')D_\epsilon \right. \\ &\quad \left. + \mathbb{E}_Q \left\{ \gamma(J_k^*(x_{r'}) - J_k^*(x_r)) - c \left[(\ell_k^*(r', \epsilon, Q)D'_\epsilon - Q)^+ - (\ell_k^*(r', \epsilon, Q)D_\epsilon - Q)^+ \right] \right\} \right]. \end{aligned}$$

We bound the three differences inside the expectations. First, since $D(r, p, y, \epsilon) = \min\{\alpha r - \beta p + \epsilon, y\}$ and $z \mapsto \min\{z, y\}$ is nondecreasing and 1-Lipschitz,

$$0 \leq D(r', p, y, \epsilon) - D(r, p, y, \epsilon) \leq \alpha \Delta r.$$

Thus

$$p_k^*(r') \left(D(r', p_k^*(r'), y_k^*(r'), \epsilon) - D(r, p_k^*(r'), y_k^*(r'), \epsilon) \right) \leq \alpha p_k^*(r') \Delta r.$$

Second, the next-period rating differs only through the term $(1 - \delta)r$, so $x_{r'} - x_r = (1 - \delta)\Delta r$. By the induction hypothesis (C.3),

$$\gamma(J_k^*(x_{r'}) - J_k^*(x_r)) \leq \gamma L(1 - \delta)\Delta r.$$

Finally, since the positive-part operator is 1-Lipschitz,

$$\begin{aligned} & c(\ell_k^*(r', \epsilon, Q)D(r, p_k^*(r'), y_k^*(r'), \epsilon) - Q)^+ - c(\ell_k^*(r', \epsilon, Q)D(r', p_k^*(r'), y_k^*(r'), \epsilon) - Q)^+ \\ & \leq c \left| \ell_k^*(r', \epsilon, Q) \left(D(r, p_k^*(r'), y_k^*(r'), \epsilon) - D(r', p_k^*(r'), y_k^*(r'), \epsilon) \right) \right| \\ & \leq \alpha c \ell_k^*(r', \epsilon, Q) \Delta r \leq \alpha c \bar{x} \Delta r, \end{aligned}$$

where the last inequality uses $\ell_k^*(r', \epsilon, Q) \in \mathcal{X}$. Combining the bounds and taking expectations yields

$$\begin{aligned} J_{k+1}^*(r') - J_{k+1}^*(r) & \leq \alpha p_k^*(r') \Delta r + \gamma L(1 - \delta) \Delta r + \alpha c \bar{x} \Delta r \\ & \leq (\alpha \bar{p} + \gamma L(1 - \delta) + \alpha c \bar{x}) \Delta r = L \Delta r, \end{aligned}$$

where the last equality uses (C.2). This proves (C.3) for $r' > r$; the case $r' < r$ follows by symmetry. Letting $k \rightarrow \infty$ and using $J_k^*(r) \rightarrow J^*(r)$ for each r , we obtain $|J^*(r') - J^*(r)| \leq L|r' - r|$, which completes the proof. $\square$

With this Lipschitz bound in hand, the rest of the proof has two ingredients: we first construct an affine surrogate $\tilde{J}$ with small uniform error on $[\underline{r}, \bar{r}]$, and then translate that approximation error into a loss bound for the greedy policy induced by $\tilde{J}$.

Because $p \in [0, \bar{p}]$ and $\ell \in \mathcal{X}$, the one-period satisfaction $\hat{v}(\ell) - p$ lies in $[\underline{r}, \bar{r}]$. Therefore, whenever $r \in [\underline{r}, \bar{r}]$, the next rating

$$(1 - \delta)r + \delta(\hat{v}(\ell) - p)$$

is a convex combination of two points in $[\underline{r}, \bar{r}]$ and remains in the same interval. Hence $[\underline{r}, \bar{r}]$ is invariant under all admissible controls.

Choose $r_L := \underline{r} + \frac{W}{6}$ and $r_U := \bar{r} - \frac{W}{6}$, and let $\tilde{J}(r) = \tilde{J}_0 + kr$ be the chord through $(r_L, J^*(r_L))$ and $(r_U, J^*(r_U))$, where

$$k := \frac{J^*(r_U) - J^*(r_L)}{r_U - r_L} \quad \text{and} \quad \tilde{J}_0 := J^*(r_L) - kr_L.$$

This choice leaves endpoint strips of width $W/6$ and a middle interval of length $2W/3$, which balances the two approximation bounds below. By Proposition 1 and Lemma C.1,

$$0 \leq k = \frac{J^*(r_U) - J^*(r_L)}{r_U - r_L} \leq L.$$

Let $\tilde{\pi} = (\tilde{p}, \tilde{y}, \tilde{\ell})$ denote the greedy policy induced by $\tilde{J}$, and define $\varepsilon_J := \max_{r \in [\underline{r}, \bar{r}]} |J^*(r) - \tilde{J}(r)|$.

Step 1 (bounding ε_J). For any $r \in [r_L, r_U]$, write

$$r = (1 - s)r_L + sr_U, \quad \text{where} \quad s := \frac{r - r_L}{r_U - r_L} \in [0, 1].$$

Because $\tilde{J}$ is the chord through $(r_L, J^*(r_L))$ and $(r_U, J^*(r_U))$, there is $\tilde{J}(r) = (1 - s)J^*(r_L) + sJ^*(r_U)$. Therefore,

$$\begin{aligned} |J^*(r) - \tilde{J}(r)| & = |(1 - s)(J^*(r) - J^*(r_L)) - s(J^*(r_U) - J^*(r))| \\ & \leq (1 - s)|J^*(r) - J^*(r_L)| + s|J^*(r_U) - J^*(r)| \\ & \leq (1 - s)L(r - r_L) + sL(r_U - r) \\ & = 2Ls(1 - s)(r_U - r_L) \leq \frac{L}{2}(r_U - r_L) = \frac{LW}{3}. \end{aligned}$$

If $r \notin [r_L, r_U]$, let a denote the nearest endpoint of $[r_L, r_U]$ (so $a = r_L$ for $r \leq r_L$ and $a = r_U$ for $r \geq r_U$). Since $\tilde{J}(a) = J^*(a)$,

$$|J^*(r) - \tilde{J}(r)| \leq |J^*(r) - J^*(a)| + |\tilde{J}(r) - \tilde{J}(a)| \leq (L + |k|) |r - a| \leq 2L|r - a|.$$

Each outer strip has width $W/6$, so $|r - a| \leq W/6$ and therefore again $|J^*(r) - \tilde{J}(r)| \leq LW/3$. Hence $\varepsilon_J \leq \frac{LW}{3}$.

Step 2 (policy loss bound in terms of ε_J). We next show that

$$J^*(r) - J^{\tilde{\pi}}(r) \leq \frac{2\gamma}{1-\gamma} \varepsilon_J \text{ for all } r \in [\underline{r}, \bar{r}].$$

For any stationary policy $\pi = (p, y, \ell)$ and any bounded continuation value function $J(\cdot)$ on $[\underline{r}, \bar{r}]$, write $D_\pi(r, \epsilon) := D(r, p(r), y(r), \epsilon)$ and define the Bellman operator

$$(\mathcal{T}^\pi J)(r) := \mathbb{E}_\epsilon \left[p(r) D_\pi(r, \epsilon) + \mathbb{E}_Q \left\{ \gamma J(r_\pi^+(r, \epsilon, Q)) - c(\ell(r, p(r), D_\pi(r, \epsilon), Q)) D_\pi(r, \epsilon) - Q \right\}^+ \right],$$

where

$$r_\pi^+(r, \epsilon, Q) := \delta(\hat{v}(\ell(r, p(r), D_\pi(r, \epsilon), Q)) - p(r)) + (1 - \delta)r.$$

Throughout this step, we work on the feasible rating interval $[\underline{r}, \bar{r}]$, so $r_\pi^+(r, \epsilon, Q) \in [\underline{r}, \bar{r}]$ whenever $r \in [\underline{r}, \bar{r}]$.

By definition, the value under policy π is the fixed point of $\mathcal{T}^\pi$, i.e., $J^\pi = \mathcal{T}^\pi J^\pi$. Let π^* be an optimal stationary policy, so $J^* = J^{\pi^*} = \mathcal{T}^{\pi^*} J^*$, and define $M := \sup_{x \in [\underline{r}, \bar{r}]} (J^*(x) - J^{\tilde{\pi}}(x)) \geq 0$.

For any two bounded functions J_1, J_2 and any $r \in [\underline{r}, \bar{r}]$,

$$|(\mathcal{T}^\pi J_1)(r) - (\mathcal{T}^\pi J_2)(r)| \leq \gamma \sup_{x \in [\underline{r}, \bar{r}]} |J_1(x) - J_2(x)|.$$

Fix $r \in [\underline{r}, \bar{r}]$. Because $\tilde{\pi}$ is greedy with respect to $\tilde{J}$, the optimal policy π^* is a feasible comparison, so $(\mathcal{T}^{\tilde{\pi}} \tilde{J})(r) \geq (\mathcal{T}^{\pi^*} \tilde{J})(r)$. Therefore,

$$J^*(r) - J^{\tilde{\pi}}(r) = (\mathcal{T}^{\pi^*} J^*)(r) - (\mathcal{T}^{\tilde{\pi}} J^{\tilde{\pi}})(r) \leq (\mathcal{T}^{\pi^*} J^*)(r) - (\mathcal{T}^{\pi^*} \tilde{J})(r) + (\mathcal{T}^{\tilde{\pi}} \tilde{J})(r) - (\mathcal{T}^{\tilde{\pi}} J^{\tilde{\pi}})(r).$$

The first term is at most $\gamma \varepsilon_J$ by the contraction bound. For the second term, by definition of $\mathcal{T}^{\tilde{\pi}}$,

$$(\mathcal{T}^{\tilde{\pi}} \tilde{J})(r) - (\mathcal{T}^{\tilde{\pi}} J^{\tilde{\pi}})(r) = \gamma \mathbb{E} \left[\tilde{J}(r_\pi^+(r, \epsilon, Q)) - J^{\tilde{\pi}}(r_\pi^+(r, \epsilon, Q)) \right] \leq \gamma(\varepsilon_J + M),$$

because, for every $x \in [\underline{r}, \bar{r}]$,

$$\tilde{J}(x) - J^{\tilde{\pi}}(x) \leq |\tilde{J}(x) - J^*(x)| + (J^*(x) - J^{\tilde{\pi}}(x)) \leq \varepsilon_J + M.$$

Hence

$$J^*(r) - J^{\tilde{\pi}}(r) \leq 2\gamma \varepsilon_J + \gamma M \quad \forall r \in [\underline{r}, \bar{r}].$$

Taking the supremum over r on the left-hand side gives $M \leq 2\gamma \varepsilon_J + \gamma M$, hence $M \leq \frac{2\gamma}{1-\gamma} \varepsilon_J$. Finally, since $J^*(r) - J^{\tilde{\pi}}(r) \leq M$ for every r ,

$$J^*(r) - J^{\tilde{\pi}}(r) \leq \frac{2\gamma}{1-\gamma} \varepsilon_J \leq \frac{2\gamma}{1-\gamma} \cdot \frac{LW}{3} = \frac{2\gamma \alpha(\bar{p} + c\bar{x})}{3(1-\gamma)[1-\gamma(1-\delta)]} [\bar{p} + \hat{v}(\bar{x}) - \hat{v}(x)], \quad \forall r \in [\underline{r}, \bar{r}].$$

□

C.3. Proof of Proposition 6

Proof. Define the optimal Bellman operator $\mathcal{T}$ by

$$\mathcal{T}J(r) := \max_{\substack{p \in [0, \bar{p}] \\ y \in [\underline{y}, \bar{y}]}} \mathbb{E}_{\epsilon, Q} \left\{ pD(r, p, y, \epsilon) + \max_{\ell \in \mathcal{X}} [\gamma J(r^+(r, p, \ell)) - c(\ell D(r, p, y, \epsilon) - Q)^+] \right\},$$

where $r^+(r, p, \ell) := \delta(\hat{v}(\ell) - p) + (1 - \delta)r$, $\bar{d}(r, p) := \alpha r - \beta p$, $\bar{D}(r, p, y) := \min\{\bar{d}(r, p), y\}$, and $\bar{Q} := \mathbb{E}[Q]$. Define the fluid Bellman operator $\hat{\mathcal{T}}$ by

$$\hat{\mathcal{T}}J(r) := \max_{\substack{p \in [0, \bar{p}], y \in [\underline{y}, \bar{y}] \\ \ell \in \mathcal{X}}} \left\{ p\bar{D}(r, p, y) - c(\ell \bar{D}(r, p, y) - \bar{Q})^+ + \gamma J(r^+(r, p, \ell)) \right\}.$$

For the DFA policy $\hat{\pi} = \{\hat{p}(\cdot), \hat{y}(\cdot), \hat{\ell}(\cdot)\}$, define the true-system and fluid policy-evaluation operators by

$$(\mathcal{T}^{\hat{\pi}}J)(r) := \mathbb{E}_{\epsilon, Q} \left[\hat{p}(r)D(r, \hat{p}(r), \hat{y}(r), \epsilon) - c(\hat{\ell}(r)D(r, \hat{p}(r), \hat{y}(r), \epsilon) - Q)^+ + \gamma J(r^+(r, \hat{p}(r), \hat{\ell}(r))) \right],$$

and

$$(\hat{\mathcal{T}}^{\hat{\pi}}J)(r) := \hat{p}(r)\bar{D}(r, \hat{p}(r), \hat{y}(r)) - c(\hat{\ell}(r)\bar{D}(r, \hat{p}(r), \hat{y}(r)) - \bar{Q})^+ + \gamma J(r^+(r, \hat{p}(r), \hat{\ell}(r))).$$

Fix an admissible rating $r, p \in [0, \bar{p}]$, $y \in [\underline{y}, \bar{y}]$, and any function J for which the displayed quantities are finite. For each realization (ϵ, Q) , define

$$\Phi_J(r, p, y, \epsilon, Q) := \max_{\ell \in \mathcal{X}} \left\{ \gamma J(r^+(r, p, \ell)) - c[\ell D(r, p, y, \epsilon) - Q]^+ \right\}.$$

Using $|x^+ - y^+| \leq |x - y|$ and $|\max_x f(x) - \max_x g(x)| \leq \max_x |f(x) - g(x)|$, we obtain

$$\begin{aligned} |\Phi_J(r, p, y, \epsilon, Q) - \Phi_J(r, p, y, 0, \bar{Q})| &\leq \max_{\ell \in \mathcal{X}} \left| c(\ell D(r, p, y, \epsilon) - Q)^+ - c(\ell \bar{D}(r, p, y) - \bar{Q})^+ \right| \\ &\leq c \cdot \max_{\ell \in \mathcal{X}} |\ell(D(r, p, y, \epsilon) - \bar{D}(r, p, y)) - (Q - \bar{Q})| \\ &\leq c\bar{x}|\epsilon| + c|Q - \bar{Q}|. \end{aligned}$$

The last inequality uses

$$|D(r, p, y, \epsilon) - \bar{D}(r, p, y)| = |\min\{\bar{d}(r, p) + \epsilon, y\} - \min\{\bar{d}(r, p), y\}| \leq |\epsilon|.$$

Taking expectation over (ϵ, Q) yields

$$\left| \mathbb{E}_{\epsilon, Q} \Phi_J(r, p, y, \epsilon, Q) - \Phi_J(r, p, y, 0, \bar{Q}) \right| \leq c\bar{x} \mathbb{E}|\epsilon| + c \mathbb{E}|Q - \bar{Q}|. \quad (\text{C.4})$$

The revenue term satisfies $|pD(r, p, y, \epsilon) - p\bar{D}(r, p, y)| \leq \bar{p}|\epsilon|$. Combining this bound with (C.4) and $|\max_x f(x) - \max_x g(x)| \leq \max_x |f(x) - g(x)|$ gives the one-step perturbation bound

$$|\mathcal{T}J(r) - \hat{\mathcal{T}}J(r)| \leq \eta := (\bar{p} + c\bar{x})\mathbb{E}|\epsilon| + c\mathbb{E}|Q - \bar{Q}|, \quad \text{for all admissible } r. \quad (\text{C.5})$$

By the same argument,

$$|\mathcal{T}^{\hat{\pi}}J(r) - \hat{\mathcal{T}}^{\hat{\pi}}J(r)| \leq \eta, \quad \text{for all admissible } r. \quad (\text{C.6})$$

We now work with finite-horizon value functions. Define

$$J_0^*(r) = \hat{J}_0(r) = 0, \quad J_{n+1}^* := \mathcal{T}J_n^*, \quad \hat{J}_{n+1} := \hat{\mathcal{T}}\hat{J}_n,$$

and similarly

$$J_0^{\hat{\pi}}(r) = \hat{J}_0^{\hat{\pi}}(r) = 0, \quad J_{n+1}^{\hat{\pi}} := \mathcal{T}^{\hat{\pi}} J_n^{\hat{\pi}}, \quad \hat{J}_{n+1}^{\hat{\pi}} := \hat{\mathcal{T}}^{\hat{\pi}} \hat{J}_n^{\hat{\pi}}.$$

Let $M_n := \sup_{r \text{ admissible}} |J_n^*(r) - \hat{J}_n(r)|$. We claim that $M_n \leq \eta \sum_{t=0}^{n-1} \gamma^t = \eta \frac{1-\gamma^n}{1-\gamma}$. This is true for $n = 0$. If it holds for n , then for any admissible r ,

$$|J_{n+1}^*(r) - \hat{J}_{n+1}(r)| = |\mathcal{T} J_n^*(r) - \hat{\mathcal{T}} \hat{J}_n(r)| \leq |\mathcal{T} J_n^*(r) - \hat{\mathcal{T}} J_n^*(r)| + |\hat{\mathcal{T}} J_n^*(r) - \hat{\mathcal{T}} \hat{J}_n(r)| \leq \eta + \gamma M_n.$$

Therefore

$$M_{n+1} \leq \eta + \gamma M_n \leq \eta \sum_{t=0}^n \gamma^t.$$

Hence

$$|J_n^*(r) - \hat{J}_n(r)| \leq \eta \frac{1-\gamma^n}{1-\gamma}, \quad \text{for all admissible } r. \quad (\text{C.7})$$

Exactly the same induction, now using (C.6), yields

$$|J_n^{\hat{\pi}}(r) - \hat{J}_n^{\hat{\pi}}(r)| \leq \eta \frac{1-\gamma^n}{1-\gamma}, \quad \text{for all admissible } r. \quad (\text{C.8})$$

It remains to pass to the infinite-horizon limit. For any feasible policy π and any admissible initial state r , the reputation process satisfies $r_{t+1} \leq (1-\delta)r_t + \delta\hat{v}(\bar{x})$, hence $r_t \leq \max\{r, \hat{v}(\bar{x})\}$ for all $t \geq 0$. Since prices and bag values are bounded and accepted demand is positive and bounded above by $\bar{y}$ under the active-cycle assumption, the one-period reward under any feasible policy satisfies

$$|R_t^\pi| \leq (\bar{p} + c\bar{x})D_t^\pi + c|Q_t| \leq (\bar{p} + c\bar{x})\bar{y} + c|Q_t|.$$

The right-hand side has finite expectation, uniformly in t . Therefore discounted tails vanish, and the finite-horizon values converge pointwise: $J_n^*(r) \rightarrow J^*(r)$, $\hat{J}_n(r) \rightarrow \hat{J}(r)$, $J_n^{\hat{\pi}}(r) \rightarrow J^{\hat{\pi}}(r)$, and $\hat{J}_n^{\hat{\pi}}(r) \rightarrow \hat{J}^{\hat{\pi}}(r)$.

Letting $n \rightarrow \infty$ in (C.7) and (C.8) gives $|J^*(r) - \hat{J}(r)| \leq \frac{\eta}{1-\gamma}$ and $|J^{\hat{\pi}}(r) - \hat{J}^{\hat{\pi}}(r)| \leq \frac{\eta}{1-\gamma}$ for all admissible r . Since $\hat{\pi}$ is optimal for the fluid problem, $\hat{J}^{\hat{\pi}} = \hat{J}$. Therefore, for any admissible r ,

$$0 \leq J^*(r) - J^{\hat{\pi}}(r) \leq |J^*(r) - \hat{J}(r)| + |J^{\hat{\pi}}(r) - \hat{J}(r)| \leq \frac{2\eta}{1-\gamma}.$$

Finally, by Cauchy–Schwarz,

$$\mathbb{E}|\epsilon| \leq \sqrt{\mathbb{E}[\epsilon^2]} = \sqrt{\text{Var}(\epsilon)}, \quad \mathbb{E}|Q - \bar{Q}| \leq \sqrt{\mathbb{E}[(Q - \bar{Q})^2]} = \sqrt{\text{Var}(Q)}.$$

Hence $\eta \leq (\bar{p} + c\bar{x})\sqrt{\text{Var}(\epsilon)} + c\sqrt{\text{Var}(Q)}$. Substituting this into the previous inequality completes the proof. $\square$

C.4. Proof of Proposition 7

Proof. We first show that the reservation cap need not bind at the interior deterministic steady state characterized in the proposition. Suppose, to the contrary, that $\hat{y} < \bar{d}(\hat{r}, \hat{p})$ and $0 < \hat{p} < \bar{p}$. Let $g = \bar{d}(\hat{r}, \hat{p}) - \hat{y} > 0$. Increase the price by a small $\eta > 0$, holding $\hat{y}$ and $\hat{\ell}$ fixed, where $\eta < \frac{g}{\alpha+\beta}$. Under this perturbation, the deterministic rating path remains between $\hat{r}$ and $\hat{r} - \eta$. Hence latent demand remains above $\hat{y}$, so accepted demand remains equal to $\hat{y}$. Supplementation cost and delivered bag value are unchanged, while per-period revenue increases by $\eta\hat{y}$, contradicting optimality. Therefore an interior deterministic steady state must satisfy $\hat{y} \geq \bar{d}(\hat{r}, \hat{p})$. We can therefore select any feasible

nonbinding cap and proceed with accepted demand $D = \bar{d}(r, p)$. Recall the deterministic model with latent requests $\bar{d}(r, p) = \alpha r - \beta p$, accepted demand $\bar{D}(r, p, y) = \min\{\bar{d}(r, p), y\}$, and reputation update

$$r_{t+1} = (1 - \delta)r_t + \delta(\hat{v}(\ell_t) - p_t), \quad \delta \in (0, 1). \quad (\text{C.9})$$

For the local perturbation argument, select an equivalent cap with slack, $\hat{y} > \bar{d}(\hat{r}, \hat{p})$; after deriving the first-order conditions, we return to the minimal-cap normalization $\hat{y} = \bar{d}(\hat{r}, \hat{p})$. In a neighborhood of this slack-cap representation, accepted demand equals latent requests. Let the local one-period payoff be

$$\Pi(p, \ell, r) := p \bar{d}(r, p) - c(\ell \bar{d}(r, p) - \bar{Q})^+. \quad (\text{C.10})$$

We first derive the steady-state identities in (28)–(29). We then add a short existence/uniqueness/convergence argument under an explicit interiority condition, and finally summarize the non-interior regimes.

At a steady state, $(\hat{r}, \hat{p}, \hat{y}, \hat{\ell})$ satisfies $\hat{r} = (1 - \delta)\hat{r} + \delta(\hat{v}(\hat{\ell}) - \hat{p})$, which rearranges to the first equality in (28): $\hat{r} = \hat{v}(\hat{\ell}) - \hat{p}$. Let $\hat{J}$ denote the DFA value function in (26). Since $T(r, p, \ell) := (1 - \delta)r + \delta(\hat{v}(\ell) - p)$ satisfies $T(\hat{r}, \hat{p}, \hat{\ell}) = \hat{r}$, the Bellman equation at $\hat{r}$ implies

$$\hat{J}(\hat{r}) = \Pi(\hat{p}, \hat{\ell}, \hat{r}) + \gamma \hat{J}(\hat{r}) \quad \Rightarrow \quad \hat{J}(\hat{r}) = \Pi(\hat{p}, \hat{\ell}, \hat{r}) / (1 - \gamma). \quad (\text{C.11})$$

The proof uses two-period perturbations that return reputation to its steady-state level at $t = 2$, so they have no effect on later payoffs.

Next, fix $\ell_t \equiv \hat{\ell}$ and perturb price in periods $t = 0, 1$ so that $r_2 = \hat{r}$. For ε small, consider the feasible policy

$$(p_0, \ell_0) = (\hat{p} + \varepsilon, \hat{\ell}), \quad (p_1, \ell_1) = (\hat{p} - (1 - \delta)\varepsilon, \hat{\ell}), \quad (p_t, \ell_t) = (\hat{p}, \hat{\ell}), \quad t \geq 2.$$

Using (28) and (C.9), the induced reputation path is

$$r_0 = \hat{r}, \quad r_1 = \hat{r} - \delta\varepsilon, \quad r_2 = \hat{r}, \quad r_t = \hat{r} \quad (t > 2).$$

Let $V_p(\varepsilon)$ denote the value from $\hat{r}$ under this perturbed policy. Because $r_t = \hat{r}$ for all $t \geq 2$,

$$V_p(\varepsilon) = \Pi(\hat{p} + \varepsilon, \hat{\ell}, \hat{r}) + \gamma \Pi(\hat{p} - (1 - \delta)\varepsilon, \hat{\ell}, \hat{r} - \delta\varepsilon) + \gamma^2 \hat{J}(\hat{r}).$$

Optimality of $(\hat{p}, \hat{\ell})$ at $\hat{r}$ implies $\hat{J}(\hat{r}) \geq V_p(\varepsilon)$. Using (C.11) to eliminate $\hat{J}(\hat{r})$ gives, for all sufficiently small ε ,

$$0 \geq \left[\Pi(\hat{p} + \varepsilon, \hat{\ell}, \hat{r}) - \Pi(\hat{p}, \hat{\ell}, \hat{r}) \right] + \gamma \left[\Pi(\hat{p} - (1 - \delta)\varepsilon, \hat{\ell}, \hat{r} - \delta\varepsilon) - \Pi(\hat{p}, \hat{\ell}, \hat{r}) \right]. \quad (\text{C.12})$$

Define $F_p(\varepsilon)$ as the right-hand side of (C.12). Then $F_p(0) = 0$ and, by interiority of $\hat{p}$, we have $F_p(\varepsilon) \leq 0$ for ε of both signs in a neighborhood of 0. Because the active-supplementation condition gives $\hat{\ell} \bar{d}(\hat{r}, \hat{p}) > \bar{Q}$, the positive-part term is locally linear around $(\hat{p}, \hat{\ell}, \hat{r})$ and Π is differentiable there. Hence $F'_p(0) = 0$, which yields

$$0 = \partial_p \Pi(\hat{p}, \hat{\ell}, \hat{r}) + \gamma \left(-(1 - \delta) \partial_p \Pi(\hat{p}, \hat{\ell}, \hat{r}) - \delta \partial_r \Pi(\hat{p}, \hat{\ell}, \hat{r}) \right).$$

Therefore

$$(1 - \gamma(1 - \delta)) \partial_p \Pi(\hat{p}, \hat{\ell}, \hat{r}) = \gamma \delta \partial_r \Pi(\hat{p}, \hat{\ell}, \hat{r}). \quad (\text{C.13})$$

Next, let

$$\Delta v(\varepsilon) := \hat{v}(\hat{\ell} + \varepsilon) - \hat{v}(\hat{\ell})$$

and consider the two-period perturbation

$$\begin{aligned}(p_0, \ell_0) &= (\hat{p}, \hat{\ell} + \varepsilon), \\ (p_1, \ell_1) &= (\hat{p} + (1 - \delta)\Delta v(\varepsilon), \hat{\ell}), \\ (p_t, \ell_t) &= (\hat{p}, \hat{\ell}), \quad t \geq 2.\end{aligned}$$

Using (28) and (C.9), the induced reputation path is

$$r_0 = \hat{r}, \quad r_1 = \hat{r} + \delta\Delta v(\varepsilon), \quad r_t = \hat{r} \quad (t \geq 2).$$

Optimality, interiority, and differentiability of $\hat{v}$ at $\hat{\ell}$ therefore give

$$0 = \partial_{\ell}\Pi(\hat{p}, \hat{\ell}, \hat{r}) + \gamma\hat{v}'(\hat{\ell}) \left[(1 - \delta)\partial_p\Pi(\hat{p}, \hat{\ell}, \hat{r}) + \delta\partial_r\Pi(\hat{p}, \hat{\ell}, \hat{r}) \right].$$

Equation (C.13) implies that the bracketed term multiplied by γ equals $\partial_p\Pi(\hat{p}, \hat{\ell}, \hat{r})$. Hence

$$\partial_{\ell}\Pi(\hat{p}, \hat{\ell}, \hat{r}) + \hat{v}'(\hat{\ell})\partial_p\Pi(\hat{p}, \hat{\ell}, \hat{r}) = 0. \quad (\text{C.14})$$

Because $\hat{\ell}\bar{d}(\hat{r}, \hat{p}) > \bar{Q}$, locally

$$\Pi(p, \ell, r) = (p - c\ell)(\alpha r - \beta p) + c\bar{Q}.$$

Hence $\partial_p\Pi = \alpha r - 2\beta p + \beta c\ell$, $\partial_r\Pi = \alpha(p - c\ell)$, and $\partial_{\ell}\Pi = -c(\alpha r - \beta p)$.

Substituting these into (C.13) and (C.14) gives

$$(1 - \gamma(1 - \delta))(\alpha\hat{r} - 2\beta\hat{p} + \beta c\hat{\ell}) = \gamma\delta\alpha(\hat{p} - c\hat{\ell}), \quad (\text{C.15})$$

$$\hat{v}'(\hat{\ell})(\alpha\hat{r} - 2\beta\hat{p} + \beta c\hat{\ell}) = c(\alpha\hat{r} - \beta\hat{p}). \quad (\text{C.16})$$

Using $\kappa = \gamma\delta/[1 - \gamma(1 - \delta)]$, (C.15) can be written as

$$\alpha\hat{r} - 2\beta\hat{p} + \beta c\hat{\ell} = \alpha\kappa(\hat{p} - c\hat{\ell}) \iff \bar{d}(\hat{r}, \hat{p}) = (\beta + \alpha\kappa)(\hat{p} - c\hat{\ell}).$$

This is the third equality in (28). Because demand is strictly positive in the interior active regime, it also implies $\hat{p} > c\hat{\ell}$.

Substituting this identity into (C.16) gives $\hat{v}'(\hat{\ell})\alpha\kappa(\hat{p} - c\hat{\ell}) = c(\beta + \alpha\kappa)(\hat{p} - c\hat{\ell})$, hence

$$\hat{v}'(\hat{\ell}) = c \left(1 + \frac{\beta}{\alpha\kappa} \right),$$

which is the second equality in (28). Finally, combining $\hat{r} = \hat{v}(\hat{\ell}) - \hat{p}$ with $\bar{d}(\hat{r}, \hat{p}) = (\beta + \alpha\kappa)(\hat{p} - c\hat{\ell})$ yields (29). The perturbation argument at the start of the proof shows that any interior deterministic steady state can be represented with a nonbinding cap. Hence any feasible cap $\hat{y} \geq \bar{d}(\hat{r}, \hat{p})$ is payoff-equivalent in this deterministic steady state. When $\bar{d}(\hat{r}, \hat{p}) \geq \underline{y}$, selecting $\hat{y} = \bar{d}(\hat{r}, \hat{p})$ gives the minimal-cap normalization.

To characterize solutions of the necessary conditions, assume additionally that $\hat{v}$ is differentiable on $\mathcal{X}$ and that $\hat{v}'$ is continuous and strictly decreasing. Define $\chi := c \left(1 + \frac{\beta}{\alpha\kappa} \right)$. If

$$\hat{v}'(\bar{x}) < \chi < \hat{v}'(\underline{x}), \quad (\text{C.17})$$

where the endpoint derivatives are understood one-sidedly, then by continuity and strict monotonicity of $\hat{v}'$ there exists a unique $\hat{\ell} \in (\underline{x}, \bar{x})$ such that $\hat{v}'(\hat{\ell}) = \chi$. Define $\hat{p}$ by (29) and $\hat{r} = \hat{v}(\hat{\ell}) - \hat{p}$. If, in addition, $0 < \hat{p} < \bar{p}$, $0 <$

$\bar{d}(\hat{r}, \hat{p}) < \bar{y}$, and $\hat{\ell} \bar{d}(\hat{r}, \hat{p}) > \bar{Q}$, then $(\hat{r}, \hat{p}, \hat{y}, \hat{\ell})$, with any feasible nonbinding cap $\hat{y} \geq \bar{d}(\hat{r}, \hat{p})$, is the unique interior candidate satisfying (28)–(29) up to the payoff-equivalent cap normalization. These conditions establish uniqueness among interior solutions to the steady-state necessary conditions; they do not by themselves provide a global Bellman verification.

Moreover, once this pair $(\hat{p}, \hat{\ell})$ is fixed, the induced deterministic rating dynamics are

$$r_{t+1} = (1 - \delta)r_t + \delta(\hat{v}(\hat{\ell}) - \hat{p}) = (1 - \delta)r_t + \delta\hat{r}.$$

Hence, for any initial rating r_0 , $r_t = \hat{r} + (1 - \delta)^t(r_0 - \hat{r}) \rightarrow \hat{r}$. Therefore the deterministic trajectory generated by repeatedly applying the steady-state pair $(\hat{p}, \hat{\ell})$ and any equivalent nonbinding cap converges globally to $(\hat{r}, \hat{p}, \hat{y}, \hat{\ell})$.

Finally, when (C.17) fails, the necessary conditions instead identify a boundary candidate in the bag-value dimension: if $\chi \geq \hat{v}'(\underline{x})$, then the quality choice is pushed to the lower boundary $\hat{\ell} = \underline{x}$; if $\chi \leq \hat{v}'(\bar{x})$, then it is pushed to the upper boundary $\hat{\ell} = \bar{x}$. Likewise, if the price reconstructed from (29) violates $0 < p < \bar{p}$, the associated candidate lies at $p = 0$ or $p = \bar{p}$. $\square$

C.5. Proof of Corollary 3

Proof. The additional conditions in Corollary 3 permit local implicit differentiation at the unique interior solution.

(i) **Uniqueness and comparative statics of $\hat{\ell}$.** From Proposition 7, $\hat{v}'(\hat{\ell}) = \chi(\alpha, \beta, \kappa, c)$, where $\chi(\alpha, \beta, \kappa, c) := c(1 + \frac{\beta}{\alpha\kappa})$. The additional condition in Corollary 3 gives a unique interior solution, and $\hat{v}''(\hat{\ell}) < 0$ makes it locally regular.

For comparative statics, implicitly differentiate $\hat{v}'(\hat{\ell}) = \chi(\alpha, \beta, \kappa, c)$ with respect to $\theta \in \{\kappa, \alpha, \beta, c\}$:

$$\hat{v}''(\hat{\ell}) \frac{\partial \hat{\ell}}{\partial \theta} = \frac{\partial \chi}{\partial \theta}.$$

Since $\hat{v}''(\hat{\ell}) < 0$, the sign of $\partial \hat{\ell} / \partial \theta$ is the opposite of the sign of $\partial \chi / \partial \theta$. Direct calculation yields

$$\frac{\partial \chi}{\partial \kappa} = -\frac{c\beta}{\alpha\kappa^2} < 0, \quad \frac{\partial \chi}{\partial \alpha} = -\frac{c\beta}{\alpha^2\kappa} < 0, \quad \frac{\partial \chi}{\partial \beta} = \frac{c}{\alpha\kappa} > 0, \quad \frac{\partial \chi}{\partial c} = 1 + \frac{\beta}{\alpha\kappa} > 0,$$

which implies that $\hat{\ell}$ is strictly increasing in κ and α , and strictly decreasing in β and c .

(ii) **Monotonicity of $\hat{r}$ in κ .** Using (29) and $\hat{r} = \hat{v}(\hat{\ell}) - \hat{p}$, define

$$\hat{r}(\ell, \kappa) := \frac{(\alpha\kappa + 2\beta)\hat{v}(\ell) - c(\beta + \alpha\kappa)\ell}{\alpha(1 + \kappa) + 2\beta}.$$

Along the steady state, where $\hat{\ell}$ varies with κ , $\frac{d\hat{r}}{d\kappa} = \frac{\partial \hat{r}}{\partial \kappa} + \frac{\partial \hat{r}}{\partial \ell} \cdot \frac{d\hat{\ell}}{d\kappa}$. From part (i), $d\hat{\ell}/d\kappa > 0$. A direct derivative computation gives

$$\frac{\partial \hat{r}}{\partial \kappa} = \frac{\alpha(\alpha\hat{v}(\ell) - c(\alpha + \beta)\ell)}{(\alpha(1 + \kappa) + 2\beta)^2}.$$

Evaluating at $\ell = \hat{\ell}$ and using (29), we have $\alpha\hat{v}(\hat{\ell}) - c(\alpha + \beta)\hat{\ell} = (\alpha(1 + \kappa) + 2\beta)(\hat{p} - c\hat{\ell})$. Therefore, $\bar{d}(\hat{r}, \hat{p}) = \hat{D} > 0$ implies $\hat{p} > c\hat{\ell}$, hence $\partial \hat{r} / \partial \kappa > 0$ at $\ell = \hat{\ell}$. Moreover,

$$\frac{\partial \hat{r}}{\partial \ell} = \frac{(\alpha\kappa + 2\beta)\hat{v}'(\ell) - c(\beta + \alpha\kappa)}{\alpha(1 + \kappa) + 2\beta},$$

and substituting $\hat{v}'(\hat{\ell}) = c(1 + \beta/(\alpha\kappa))$ from (28) yields

$$\frac{\partial \hat{r}}{\partial \ell} \Big|_{\ell=\hat{\ell}} = \frac{2c\beta(1 + \beta/(\alpha\kappa))}{\alpha(1 + \kappa) + 2\beta} > 0.$$

Therefore both terms in $d\hat{r}/d\kappa$ are positive, and $\hat{r}$ is strictly increasing in κ .

(iii) **Implications for γ and δ .** Since $\kappa = \gamma\delta/[1 - \gamma(1 - \delta)]$, we have

$$\frac{\partial \kappa}{\partial \gamma} = \frac{\delta}{(1 - \gamma(1 - \delta))^2} > 0 \quad \text{and} \quad \frac{\partial \kappa}{\partial \delta} = \frac{\gamma(1 - \gamma)}{(1 - \gamma(1 - \delta))^2} > 0.$$

Thus the monotonicity of $\hat{\ell}$ and $\hat{r}$ in γ and δ follows from parts (i)–(ii). $\square$

Appendix D: Numerical Details

Numerical grids and stochastic environments. The main parameter grid is $\alpha, \beta \in \{1.8, 1.9, 2.0, 2.1, 2.2\}$ and $c \in \{0.75, 0.80, 0.85, 0.90, 0.95\}$, with $\eta = 0.60$, $\gamma = 0.90$, and $\delta = 0.10$. The action ranges are $\mathcal{X} = [1, 5]$, $p \in [0, 5]$, and $y \in [0.25, 5]$, with $h_0 = 3.5$. The dynamic program is solved on 61 rating grid points on $[0, 5]$, 61 price grid points, 41 reservation-cap grid points, and 81 bag-value grid points. Thus the baseline DP search is over (r, p, y, ℓ) , and the extension with endogenous promised value adds h . Modified policy iteration uses tolerance 10^{-5} , at most 260 iterations, 7-point equiprobable quantile discretizations for (ϵ, Q) , and linear interpolation on the rating grid. The DP solver and simulator project candidate next-period ratings to $[0, 5]$; this safeguard did not bind over the saved fixed- and endogenous-promise policy grids. Each policy is evaluated from ten evenly spaced initial ratings on $[3, 5]$, with $T = 2,000$, a 200-period burn-in for long-run operating measures, and eight common-random-number replications. After averaging replications and initial states within each environment, scenario-bootstrap intervals resample environment identifiers 2,000 times.

Reported profit is the normalized discounted average

$$\bar{\Pi}^\gamma = \frac{\sum_{t=0}^{T-1} \gamma^t \Pi_t}{\sum_{t=0}^{T-1} \gamma^t}, \quad \Pi_t = p_t D_t - c(D_t \ell_t - Q_t)^+.$$

The DP solver, numerical LVA search, and simulator all use $D_t = \min\{(\alpha r_t - \beta p_t + \epsilon_t)^+, y_t\}$. When $D_t = 0$, no Stage-II allocation is made and the rating remains unchanged; the period contributes zero profit and household waste, and Q_t units of in-store waste. These numerical rules allow no-sale states outside the positive-demand region maintained in the analytical model.

For each deterministic grid point in the main policy comparison, we draw five stochastic environments. The demand shock follows a mean-zero truncated normal distribution with support $[-3\sigma_\epsilon, 3\sigma_\epsilon]$, where $\sigma_\epsilon \sim U[0.35, 0.65]$. End-of-cycle surplus follows a scaled beta distribution on $[Q_{\min}, Q_{\max}]$, where $Q_{\min} \sim U[1.5, 2.5]$, $Q_{\max} \sim U[4.5, 5.5]$, and the beta shape parameters are drawn from $U[1.5, 4.0]$. The same quantile discretization is used when solving the DP and when evaluating approximate policies.

Policy and extension details. The oracle-calibrated LVA estimates the affine slope parameter k by least-squares fitting to the DP value function on the rating grid. This design isolates approximation-form error and does not test how k would be learned out of sample. Outside the extension with endogenous promised value, when delivered bag value exceeds household regular needs $\bar{\ell}_{\text{need}} = 1.1 \ell_{\text{ref}}$, the excess amount is wasted with probability weight $\xi = 0.3$. That extension instead keeps this need benchmark fixed at $1.1h_0$, as specified below.

For the selective-rating extension, we use five stochastic environments at $\alpha = \beta = 2$ and $c = 0.95$. For each target mean on the 81-point bag-value grid, we enumerate every feasible one- and two-point support and select the distribution that maximizes the rating-weighted satisfaction ratio in (31); we call this the rating-weighted optimal allocation. When $\zeta = 0$, this procedure recovers the ordinary envelope allocation.

The extension with endogenous promised value keeps $(\lambda_+, \lambda_-) = (1.5, 3.0)$ and compares the fixed benchmark $h_t = h_0 = 3.5$ with endogenous choices from $\mathcal{H} = \{3.0, 3.5, \dots, 5.0\}$. Both regimes obey $2p_t \leq h_t \leq 4p_t$ and use the centered demand adjustment $\rho(h_t - h_0)$. Within each matched environment, the quality-adjusted-value support $\mathcal{X} = [1, 5]$, the surplus distribution, and c are held fixed across h ; the customer-valuation function and its envelope are recomputed for each promise. Baseline and Envelope retain their fixed-promise definitions; DP searches jointly over (p, y, h) at every rating-grid point, while LVA searches over the same actions under its separately fitted affine continuation-value approximation. We use $\rho \in \{0.50, 0.55, 0.60\}$ in 100 independently generated stochastic environments at $\alpha = \beta = 2$ and $c = 0.95$, with matched random-number streams across policies. For each endogenous policy, scenario-level metrics are averaged equally across the three ρ values before resampling the 100 scenario identifiers; the fixed result is retained once. Household regular needs remain fixed at $\bar{\ell}_{\text{need}} = 1.1h_0$ rather than moving mechanically with the chosen promise.

Because the price–promise band does not ensure the sufficient profitability condition in Assumption 1, we checked the computed DP value function directly. We classified it as nondecreasing when every adjacent difference on the 61-point rating grid was at least -10^{-10} ; all 300 solves (100 environments at each of the three ρ values) passed, with a maximum downward change of zero. The largest final sup-norm change between value iterations was 9.96×10^{-6} , below the solver tolerance 10^{-5} . Thus the envelope reduction used by DP remains valid in the endogenous-promise extension.

Responsiveness-index comparative statics. To extend the DFA comparative statics in Corollary 3 to the stochastic model, we vary $\kappa = \gamma\delta/(1 - \gamma(1 - \delta))$ through two channels:

- (a) *Discount-factor channel*: $\gamma \in [0.50, 0.90]$ with $\delta = 0.10$.
- (b) *Rating-update channel*: $\delta \in [0.03, 0.10]$ with $\gamma = 0.90$.

All other parameters stay at baseline ($\alpha = 2$, $\beta = 2$, $c = 0.95$). Each channel uses an 8-point equispaced grid, with the baseline value at the upper bound, yielding κ up to 0.47. All tested points satisfy $(1 - \delta)\beta \geq \delta\alpha$ and converge. At each grid point, we solve the DP with the numerical settings above and simulate using the protocol in Section 6. Post-burn-in averages of ℓ_t , r_t , and p_t serve as stochastic analogues of ℓ^* , r^* , and p^* .

Figure D.1 Comparative statics for the responsiveness index κ , varied through γ and δ .

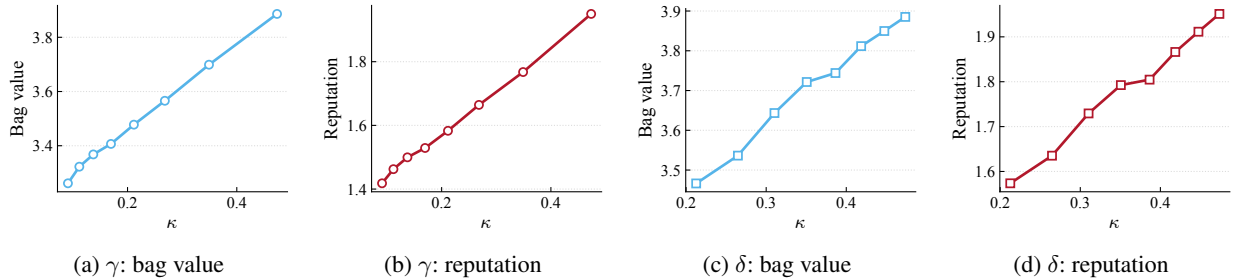


Figure D.1 shows that increasing κ through either γ or δ raises long-run bag value and reputation, matching the deterministic force in Corollary 3 and supporting κ as a useful summary statistic for reputation incentives in the stochastic model.