Supplementary Appendix

Learning to Choose with Categories:
Theory and an Application to Algorithmic Prominence

Philippe Jehiel Aviman Satpathy

November 9, 2026

This appendix supplies proofs for the theorems, propositions, lemma, and corollaries in the manuscript.

Maintained primitives and convention: There are finitely many alternatives $a \in \mathcal{A}$ with stationary independent reward laws R_a (with finite second moments uniformly bounded by M_2), and a fixed partition into similarity classes $\mathcal{S}$. Menus are drawn independently over time, independently of rewards. Within a chosen class, the tie-breaking kernel is fixed and exogenous. Its reduction gives a nonempty finite supported class-menu set Ω^+, probabilities $p(\omega) > 0$ summing to one, and conditional class-menu means $\pi_s(\omega)$. We discard classes with zero availability: $P_s = \sum_{\omega \ni s} p(\omega) > 0$ for every $s \in \mathcal{S}$. At each fixed finite sensitivity $\beta \geq 0$, write

$$\sigma^s_\omega(v;\beta) = \frac{\exp(\beta v_s)}{\sum_{r\in\omega}\exp(\beta v_r)}, \qquad \Lambda_s(v;\beta) = \sum_{\omega \ni s} p(\omega)\sigma^s_\omega(v;\beta),$$

with $\sigma^s_\omega = 0$ for $s \notin \omega$. Let $N_k(s)$ count class visits before period k, let $\kappa_s > 0$, and let $\chi = \sum_s \kappa_s$. The maintained recursion is

$$v_{k+1}(s) = v_k(s) + \frac{\mathbf{1}\{s_k = s\}}{\kappa_s + N_k(s) + 1}\big(r_k - v_k(s)\big), \qquad x_{k,s} = \frac{\kappa_s + N_k(s)}{\chi + k}.$$

Sensitivity remains fixed along a learning path; limits as $\beta \to \infty$ compare the long-run behavior of different fixed environments (except for the impossibility result in the section on welfare analysis where we explicitly permit time-varying sensitivity). Statements about stochastic periodic limit sets concern the harmonic learning clock.

Proof guide: The following table identifies every formal claim in the paper. Supporting claims are proved in the adjacent subsections, including attainability, the VE continuum, the logit axioms, the two-class illustrations, and cycling minimality.

Main-text statement	Appendix proof
Continuous-time approximation	Propositions A.4, A.5
Equilibrium and local-stability correspondence	Proposition A.7
Structure of SVE and high-sensitivity limits	Theorem B.1
Strict VE characterization and local stability	Theorem B.6
Strict-equilibrium coverage corollary	Corollary B.7
Constant competitive-menu payoffs	Theorem C.1
Robustness to payoff heterogeneity	Corollary C.3

Convergence with two classes	Corollary C.4
Convergence under nonpositive payoff-crowding covariance	Theorem D.2
High singleton shifts and Limiting Indifference	Theorem D.4
Total indifference under monotone menu probabilities	Proposition D.5
Additional primitive conditions for corrective feedback	Proposition D.6
Low singleton shifts and strict multiplicity	Theorem D.11
Every strict ranking can be sustained	Proposition D.12
Nearly equal singleton probabilities	Proposition D.13
Persistent cycling	Theorem E.2, Proposition E.5
Common-score oracle benchmark and efficiency	Proposition F.1
Efficient scores in own-payoff intervals	Proposition F.2
Path-based welfare bounds	Lemma F.3, Proposition F.4, Corollary F.5
Singleton welfare and the wine illustrations	Lemma F.6, equation (F.11)
Limits of payoff aggregation	Proposition F.8, Theorem F.9
Learning under stationary prominence	Proposition G.1, Lemma G.2
Two-state illustration and its welfare gap	Lemma G.3
Optimal deterministic prominence (three states)	Proposition G.4
Optimal randomized prominence	Proposition G.5
Approximately optimal prominence and welfare	Corollary G.6
Heterogeneous consumer partitions	Proposition G.7, Corollary G.8
Prominence caps and welfare	Propositions G.9, G.10, G.11
General policy reduction and randomization bounds	Lemma G.12, Corollary G.13
Optimal costly prominence and the example	Propositions G.15, G.16
Bayesian learning: equilibrium, strict stability, and purification	Proposition H.1
Convergence in balanced binary constant-sum decision trees	Proposition I.1

Contents

A The stochastic learning process and its mean field ODE limit **3**

B Equilibrium structure and strict-equilibrium refinement **10**

C Constant competitive-menu payoffs and two-class convergence **18**

D Corrective feedback and singleton-payoff shifts **25**

E The three-menu cycling construction **36**

F Coarse representation and the welfare cost of learning **51**

G Algorithmic prominence and consumer learning **59**

H Bayesian learning and coarse misspecification **86**

I Convergence in balanced binary constant-sum decision trees **90**

A The stochastic learning process and its mean field ODE limit

This section proves the two propositions in the model section and the elementary benchmark claims. Throughout, the sensitivity β is fixed and finite. Write $m = |\mathcal{S}|$, omit classes that have zero availability, and let

$$P_s = \sum_{\omega \ni s} p(\omega) > 0, \qquad R_s(v) = \sum_{\omega \ni s} p(\omega)\sigma_\omega^s(v)\pi_s(\omega), \qquad g_s(v) = \frac{R_s(v)}{\Lambda_s(v)}.$$

Here R_s denotes an unconditional expected reward contribution of a class, not a primitive reward distribution. Let $\chi = \sum_s \kappa_s$ and $\gamma_{k+1} = (\chi + k + 1)^{-1}$. All initial valuations are finite, and $\kappa_s > 0$. An interior of a probability simplex always means its *relative* interior.

A.1 Reduction and deterministic preliminary facts

Lemma A.1 (Reduction to class menus)**.** *Fix a within-class tie-breaking kernel $\Upsilon(a \mid s, \psi)$, independent of time, past observations, and valuations, with support in $s \cap \psi$. Define*

$$\mu_s(\psi) = \sum_{a \in s \cap \psi} \Upsilon(a \mid s, \psi)\mu_a, \quad p(\omega) = \sum_{\psi : \rho[\psi] = \omega} f(\psi), \quad \pi_s(\omega) = \frac{\sum_{\psi : \rho[\psi] = \omega} f(\psi)\mu_s(\psi)}{p(\omega)}$$

for $p(\omega) > 0$. These reduced primitives preserve Λ, R, and g, and therefore preserve both mean-field systems. Uniform within-class selection is the special case $\Upsilon(a \mid s, \psi) = |s \cap \psi|^{-1}$.

Proof. The class choice probability is the same in all objective menus inducing ω. For such a menu and $s \in \omega$, Bayes' law gives

$$\Pr(\psi \mid \omega, s \text{ chosen}, v) = \frac{f(\psi)\sigma_\omega^s(v)}{p(\omega)\sigma_\omega^s(v)} = \frac{f(\psi)}{p(\omega)}.$$

Consequently the conditional reward mean is $\pi_s(\omega)$. Summing over class menus gives the stated expressions for Λ and R. In fact the conditional reward law is the fixed mixture

$$\widetilde{R}_{s,\omega} = \sum_{\psi : \rho[\psi] = \omega} \frac{f(\psi)}{p(\omega)} \sum_{a \in s \cap \psi} \Upsilon(a \mid s, \psi) R_a.$$

Using this mixture preserves the law of the observed class-menu, class, and reward process, conditional on its history. □

Let $\underline{\pi}_s = \min_{\omega \ni s:\, p(\omega) > 0} \pi_s(\omega)$, and $\overline{\pi}_s = \max_{\omega \ni s:\, p(\omega) > 0} \pi_s(\omega)$.

Lemma A.2 (Well-posedness, boundedness, and positive frequencies)**.** *The normalized system $\dot{v} = g(v) - v$ and the joint asynchronous system*

$$\dot{v}_s = \frac{\Lambda_s(v)}{x_s}(g_s(v) - v_s), \qquad \dot{x}_s = \Lambda_s(v) - x_s \tag{A.1}$$

have unique global forward solutions from every finite valuation vector and, for the joint asynchronous system, every $x(0) \in \operatorname{ri} \Delta(\mathcal{S})$. Each valuation remains in

$$[\min\{v_s(0), \underline{\pi}_s\}, \max\{v_s(0), \overline{\pi}_s\}].$$

Along each joint trajectory, $\inf_{t \geq 0} x_s(t) > 0$. Every equilibrium valuation belongs to $\prod_s [\underline{\pi}_s, \overline{\pi}_s]$, and at least one equilibrium exists.

Proof. For finite β, all available-class logits and all Λ_s are strictly positive, and g, Λ are smooth. Since $g_s(v)$ is a convex combination of $\pi_s(\omega)$, the valuation vector field points inward on the indicated coordinate intervals. Thus the valuation box is forward invariant, for both systems. On this compact box there is $\ell_s > 0$ with $\Lambda_s(v) \geq \ell_s$. The frequency equation gives

$$x_s(t) = e^{-t}x_s(0) + \int_0^t e^{-(t-u)}\Lambda_s(v(u))\,du \geq e^{-t}x_s(0) + (1 - e^{-t})\ell_s. \tag{A.2}$$

Moreover $\sum_s \dot{x}_s = 1 - \sum_s x_s$, so the simplex is invariant. The solution stays in a compact subset of the domain of its smooth vector field, which proves global existence. Equilibrium valuations satisfy $v = g(v)$ and hence belong to the payoff box. The continuous map g maps that nonempty compact convex box into itself, so Brouwer's theorem yields a fixed point, including when some coordinate intervals are singletons. Together with $x = \Lambda(v)$ it is a joint equilibrium. □

The auxiliary coordinates

$$y_s = x_s v_s \tag{A.3}$$

will be useful repeatedly. On $x_s > 0$, they are a smooth change of coordinates, and the joint system becomes

$$\dot{x}_s = \Lambda_s(y/x) - x_s, \qquad \dot{y}_s = R_s(y/x) - y_s, \tag{A.4}$$

where division is coordinatewise. This follows by differentiating $y_s = x_s v_s$; the terms involving $x_s v_s$ cancel exactly.

A.2 Adaptive sampling and the second-moment assumption

Let $\mathcal{F}_k$ contain the entire history before period k's realized menu and choice, including initial conditions and past objective menus, chosen alternatives, and observed rewards. The current valuation and visit counts are $\mathcal{F}_k$-measurable. Let $\mathcal{G}_k$ augment $\mathcal{F}_k$ by the current menu and the random choices, before the reward is observed. The independence assumptions give

$$\mathbb{E}[r_k \mid \mathcal{G}_k] = \mu_{a_k}, \qquad \mathbb{E}[r_k^2 \mid \mathcal{G}_k] \leq M_2. \tag{A.5}$$

The larger filtration is used only for the proof; it grants Alice no additional information.

Lemma A.3 (A visit-count martingale law)**.** *Let $I_{k,s} = \mathbf{1}\{s_k = s\}$ and $\xi_{k,s} = I_{k,s}(r_k - \mu_{a_k})$. Almost surely, for every class s, the series*

$$\sum_{k\geq 0} \frac{\xi_{k,s}}{\kappa_s + N_k(s) + 1} \tag{A.6}$$

converges. On $\{N_k(s) \to \infty\}$,

$$\frac{\sum_{j<k} \xi_{j,s}}{\kappa_s + N_k(s)} \longrightarrow 0. \tag{A.7}$$

Proof. By iterated conditioning in (A.5), $\mathbb{E}[\xi_{k,s} \mid \mathcal{F}_k] = 0$ and

$$\mathbb{E}[\xi_{k,s}^2 \mid \mathcal{F}_k] \leq M_2\,\mathbb{E}[I_{k,s} \mid \mathcal{F}_k].$$

The denominator in (A.6) is predictable. Furthermore, by monotone convergence and by counting the successive actual visits to s,

$$\sum_{k\geq 0} \mathbb{E}\left[\frac{\xi_{k,s}^2}{(\kappa_s + N_k(s) + 1)^2}\right] \leq M_2 \mathbb{E}\sum_{k\geq 0} \frac{I_{k,s}}{(\kappa_s + N_k(s) + 1)^2} \leq M_2 \sum_{j\geq 1} \frac{1}{(\kappa_s + j)^2} < \infty.$$

The partial sums are therefore an L^2-bounded martingale and converge almost surely. On a sample path with infinitely many visits, discard the zero terms and index the remaining terms by visit number. Kronecker's lemma with denominators $\kappa_s + j$ proves (A.7). No positive lower bound on sampling frequencies is used in this lemma. □

Proposition A.4 (Boundedness and infinite sampling)**.** *Under the maintained reward assumptions,*

$$\sup_k \|v_k\| < \infty, \qquad \liminf_{k\to\infty} \min_s x_{k,s} > 0 \quad \textit{almost surely,} \qquad x_{k,s} = \frac{\kappa_s + N_k(s)}{\chi + k}.$$

In particular every class is visited infinitely often, and its effective learning rates have divergent sum and square-summable squares.

Proof. Induction in the update rule gives the exact identity

$$v_k(s) = \frac{\kappa_s v_0(s) + \sum_{j<k} I_{j,s} r_j}{\kappa_s + N_k(s)}. \tag{A.8}$$

Replace each reward by $\mu_{a_j} + (r_j - \mu_{a_j})$. On a path with infinitely many visits, the contribution of the second term tends to zero by Lemma A.3; the first term is a convex combination of finitely many finite means and the initial value. Hence the valuation sequence for that class is bounded. On a path with finitely many visits it is eventually constant and has only finitely many finite values. Taking the intersection of these probability-one events over the finite class set proves boundedness before invoking continued sampling. The bounded martingale differences $I_{k,s} - \Lambda_s(v_k)$ satisfy

$$\frac{N_k(s)}{k} - \frac{1}{k}\sum_{j<k} \Lambda_s(v_j) \longrightarrow 0 \quad \text{almost surely.} \tag{A.9}$$

For completeness, convergence of $\sum_k [I_{k,s} - \Lambda_s(v_k)]/(k+1)$ follows from square summability of its variances, and Kronecker's lemma gives the assertion. On any path with $\sup_k \|v_k\|_\infty \le B < \infty$,

$$\Lambda_s(v_k) \ge \frac{P_s}{m} e^{-2\beta B} > 0.$$

Combining this pathwise bound with (A.9) proves the positive lower limiting frequency and thus infinitely many visits. Finally, the nonzero effective gains are precisely $(\kappa_s + j)^{-1}$, $j = 1, 2, \ldots$. Their first powers sum to infinity and their squares have finite sum. □

Because the primitive family is finite and stationary, the maintained second-moment bound also implies uniform integrability of the squared primitive rewards.

A.3 Proof of the continuous-time approximation

Proposition A.5 (Asymptotic pseudo-trajectory)**.** *Let $\tau_0 = 0$, $\tau_k = \sum_{j<k} \gamma_{j+1}$, and let $Z(t)$ be the affine interpolation of (v_k, x_k) at these times. Almost surely, for every $T > 0$,*

$$\lim_{t\to\infty} \sup_{0\le h\le T} \|Z(t+h) - \Phi_h(Z(t))\| = 0,$$

where Φ is the flow of (A.1). The set of subsequential limits is nonempty, compact, connected, invariant, and internally chain transitive for this flow.

Proof. Define $y_{k,s} = x_{k,s} v_k(s)$. By (A.8), the following identities are exact:

$$\begin{aligned} x_{k+1,s} - x_{k,s} &= \gamma_{k+1}(I_{k,s} - x_{k,s}), \\ y_{k+1,s} - y_{k,s} &= \gamma_{k+1}(I_{k,s} r_k - y_{k,s}). \end{aligned} \tag{A.10}$$

Write $W_k = (x_k, y_k)$, and let F be the vector field in (A.4). Then

$$W_{k+1} - W_k = \gamma_{k+1}\{F(W_k) + U_{k+1}\}, \qquad U_{k+1} = \begin{pmatrix} I_k - \Lambda(v_k) \\ I_k r_k - R(v_k) \end{pmatrix},$$

where I_k is the chosen-class unit vector. The conditional expectation of U_{k+1} is zero. Since $\|I_k\|^2 = 1$ and $\|I_k r_k\|^2 = r_k^2$, its conditional second moment is bounded by $1 + M_2$ (or any larger fixed constant). Thus

$$\sum_k \gamma_{k+1} U_{k+1} \quad \text{converges almost surely in every coordinate.} \tag{A.11}$$

This uses $\sum_k \gamma_{k+1}^2 < \infty$. By Proposition A.4, W_k eventually lies in a compact subset of the open set $\{(x, y) : x_s > 0\}$. On a slightly larger compact set F is bounded and Lipschitz. The affine interpolation $\overline{W}$ therefore satisfies, on every fixed future time window, the integral equation for F up to an error tending uniformly to zero. Here are the two ingredients. First, every tail partial sum of the noise series in (A.11) is uniformly small; the partial first and last interpolation intervals are also small because $\gamma_{k+1} U_{k+1} \to 0$. Second, $W_{k+1} - W_k \to 0$, and the error from replacing $F(\overline{W}(u))$ by the piecewise constant $F(W_k)$ is bounded by the Lipschitz constant times the largest mesh increment, multiplied by the window length. Gronwall's inequality now gives the asymptotic-pseudo-trajectory property for (A.4). This is also the criterion in Benaïm (1999, Proposition 4.1). The smooth coordinate map $Q(x, y) = (y/x, x)$ is uniformly Lipschitz on the eventual compact set and maps the flow of (A.4) to that of (A.1). Consequently $Q(\overline{W})$ is an asymptotic pseudo-trajectory for the joint flow. Its distance from the affine interpolation Z tends to zero: on each mesh interval both are uniformly close to the common left endpoint, since the mesh increments tend to zero. Hence Z has the same property and the same limit set. Precompactness and Benaïm (1999, Theorem 5.7) give internal chain transitivity and the stated compactness, connectedness, and invariance properties. □

Corollary A.6 (Consequences and their scope)**.** *If the joint flow has a globally asymptotically stable equilibrium, the stochastic process converges to it almost surely. If v_k converges almost surely on an event, its limit is an SVE on that event and $x_k \to \Lambda(\lim_k v_k)$.*

Proof. A compact internally chain transitive set intersecting the basin of an attractor lies in that attractor; apply this standard attractor property to the singleton global attractor and Proposition A.5 (equivalently, use Benaïm (1999, Theorem 6.9)). Alternatively, stability and uniform attraction on a compact neighborhood of the sample path imply trapping of every sufficiently late small perturbation. If $v_k \to v^*$, equation (A.9) and continuity imply $x_k \to \Lambda(v^*)$. The resulting singleton limit set of an asymptotic pseudo-trajectory is invariant, so it must be an equilibrium, which implies $g(v^*) = v^*$. □

A.4 Equilibria, linearization, and elementary benchmarks

Proposition A.7 (Exact local correspondence)**.** *The equilibrium valuation vectors of the joint asynchronous and normalized systems coincide. At an equilibrium $(v^*, \Lambda(v^*))$, the joint Jacobian on $\mathbb{R}^m \times T\Delta$ has the eigenvalues of $Dg(v^*) - I_m$, together with $m - 1$ additional eigenvalues equal to -1, counting algebraic multiplicities. Thus, hyperbolicity and the number of unstable directions agree. For hyperbolic equilibria, local asymptotic stability under one system is equivalent to local exponential stability under both systems.*

Proof. The frequency equation at an equilibrium forces $x = \Lambda(v) > 0$; the valuation equation then forces $v = g(v)$. Conversely these conditions give an equilibrium. Let $a_s(v, x) = \Lambda_s(v)/x_s$. At an equilibrium $a_s = 1$, and the factor $g_s(v) - v_s$ multiplying every derivative of a_s vanishes. Therefore the full linearization is

$$\begin{pmatrix} Dg(v^*) - I_m & 0 \\ D\Lambda(v^*) & -I_{T\Delta} \end{pmatrix}, \qquad T\Delta = \{u \in \mathbb{R}^m : \sum_s u_s = 0\}.$$

The lower left block has image in $T\Delta$ because $\sum_s \Lambda_s \equiv 1$. The characteristic polynomial is the product of the two diagonal-block characteristic polynomials. The claims follow from the linearization theorem for a hyperbolic equilibrium. In particular there are $m - 1$, rather than m, additional frequency eigenvalues on the simplex state space. □

Proposition A.8 (Constant-target benchmarks)**.** *Suppose $g(v) \equiv c$. The normalized and joint asynchronous systems each have a unique globally asymptotically stable equilibrium, with valuation c and, in the joint system, frequency $\Lambda(c)$. The stochastic harmonic-visit process converges to these values almost surely. This applies to pure exploration, the finest partition, full class availability, and homogeneous within-class expected payoffs.*

Proof. The normalized solution is $v(t) = c + (v(0) - c)e^{-t}$. In the joint system, (A.4) gives

$$\frac{d}{dt}(y_s - c_s x_s) = -(y_s - c_s x_s), \qquad v_s(t) - c_s = \frac{x_s(0)}{x_s(t)}(v_s(0) - c_s)e^{-t}.$$

Lemma A.2 bounds $x_s(t)$ away from zero, so $v(t) \to c$. Formula (A.2) then implies $x(t) \to \Lambda(c)$. The unique equilibrium is locally exponentially stable by Proposition A.7, because $Dg = 0$, and hence is globally asymptotically stable. The stochastic conclusion follows from Corollary A.6.

When $\beta = 0$, the constant target is

$$c_s = \frac{\sum_{\omega \ni s} p(\omega)|\omega|^{-1}\pi_s(\omega)}{\sum_{\omega \ni s} p(\omega)|\omega|^{-1}}.$$

For the finest partition it is the corresponding alternative mean μ_a. If $p(\mathcal{S}) = 1$, the common logit factor cancels in the ratio defining g_s, leaving $c_s = \pi_s(\mathcal{S})$. Finally, homogeneous means $\mu_a = c_s$ within class s imply $\pi_s(\omega) = c_s$ whenever $s \in \omega$. These arguments establish all four benchmarks. □

A.5 Inverse-propensity normalization with estimated propensities

The normalized dynamics can be asymptotically implemented without knowing the menu distribution or the current aggregate propensities. Alice estimates selection probabilities on a faster timescale than she updates valuations, following the two-timescale stochastic-approximation principle of Borkar (1997). Initialize $\widehat{\Lambda}_0 \in \operatorname{ri}\Delta(\mathcal{S})$ and update *every* class's estimate each period:

$$\begin{aligned} \widehat{\Lambda}_{k+1,s} &= \widehat{\Lambda}_{k,s} + b_k(I_{k,s} - \widehat{\Lambda}_{k,s}), \\ v_{k+1}(s) &= v_k(s) + \alpha_k \frac{I_{k,s}}{d_{k,s}}(r_k - v_k(s)), \qquad d_{k,s} = \max\{\widehat{\Lambda}_{k,s}, \varepsilon_k\}. \end{aligned} \tag{A.12}$$

The valuation update uses the estimate available *before* the current choice. The deterministic floor $\varepsilon_k \downarrow 0$ prevents small initial estimates from generating large inverse weights. One concrete choice of gains is

$$\alpha_k = (k+2)^{-1}, \qquad b_k = (k+2)^{-2/3}, \qquad \varepsilon_k = (k+2)^{-1/6}. \tag{A.13}$$

In particular, $\alpha_k/b_k \to 0$: the estimator adjusts faster than valuations. For a frozen valuation vector v, its mean equation is $\dot{\ell} = \Lambda(v) - \ell$, whose unique globally attracting equilibrium is $\Lambda(v)$. The following result establishes tracking, including when valuations themselves do not converge.

Proposition A.9 (Learning the normalizing propensities)**.** *Under the maintained assumptions and the gains* (A.13)*, the recursion* (A.12) *satisfies, almost surely,*

$$\sup_k \|v_k\| < \infty, \qquad \widehat{\Lambda}_{k,s} - \Lambda_s(v_k) \longrightarrow 0 \quad \textit{for every } s.$$

The denominator floor is inactive from some finite random time onward. The affine interpolation of v_k on the clock $\tau_k = \sum_{j<k} \alpha_j$ is an asymptotic pseudo-trajectory of $\dot{v} = g(v) - v$.

Proof. The gains satisfy

$$\sum_k \alpha_k = \sum_k b_k = \infty, \qquad \sum_k (\alpha_k/\varepsilon_k)^2 + \sum_k b_k^2 < \infty, \qquad \frac{\alpha_k}{b_k \varepsilon_k} \longrightarrow 0.$$

Also $0 < b_k < 1$, so the estimates remain in the simplex interior, and $c_{k,s} := \alpha_k I_{k,s}/d_{k,s} \in [0,1]$. By (A.5), the martingale

$$M_{n,s} = \sum_{k<n} c_{k,s}(r_k - \mu_{a_k})$$

has summable increment variances, bounded by $M_2 \sum_k (\alpha_k/\varepsilon_k)^2$. It therefore converges almost surely. Put $u_k = v_k - M_k$. Coordinatewise,

$$u_{k+1,s} = (1 - c_{k,s})u_{k,s} + c_{k,s}(\mu_{a_k} - M_{k,s}).$$

The targets on the right are bounded on each such sample path, since the primitive means form a finite set and M_k converges. Convex averaging therefore bounds u_k, and hence v_k.

We next prove tracking directly. Write $M_\infty = \lim_k M_k$ and define $L_k = \Lambda(u_k + M_\infty)$. Finite-sensitivity logit propensities are globally Lipschitz. Since $u_{k+1,s} - u_{k,s} = c_{k,s}(\mu_{a_k} - v_k(s))$, boundedness gives

$$L_k - \Lambda(v_k) \longrightarrow 0, \qquad \|L_{k+1} - L_k\| \le C\alpha_k/\varepsilon_k = o(b_k)$$

for a finite path-dependent constant C. Thus $e_k = \widehat{\Lambda}_k - L_k$ obeys

$$e_{k+1} = (1 - b_k)e_k + b_k(\zeta_{k+1} + \rho_k), \qquad \zeta_{k+1} = I_k - \Lambda(v_k), \qquad \rho_k \longrightarrow 0,$$

where explicitly $\rho_k = \Lambda(v_k) - L_k - (L_{k+1} - L_k)/b_k$. The series $\sum_k b_k \zeta_{k+1}$ converges almost surely, since ζ_{k+1} are bounded martingale differences. Variation of constants now gives $e_k \to 0$: products of the factors $1 - b_k$ tend to zero, the weighted averages of ρ_k vanish, and summation by parts makes the contribution of the convergent martingale series vanish. Consequently $\widehat{\Lambda}_k - \Lambda(v_k) \to 0$.

On every bounded valuation path, positive availability and finite β imply $\inf_k \Lambda_s(v_k) > 0$. Tracking therefore bounds the estimated propensities away from zero eventually, and the vanishing floor becomes inactive. Conditioning before the current menu gives

$$\mathbb{E}[v_{k+1}(s) - v_k(s) \mid \mathcal{F}_k] = \alpha_k \frac{\Lambda_s(v_k)}{d_{k,s}}(g_s(v_k) - v_k(s)). \tag{A.14}$$

The ratio tends to one. After localization to bounded valuation sets, the centered valuation increments have summable variances, bounded by a constant times $\sum_k (\alpha_k/\varepsilon_k)^2$. Their martingale series converges. The drift error relative to $g(v_k) - v_k$ vanishes, so the integral-equation and Gronwall argument in Prop. A.5 proves the asymptotic pseudo-trajectory property on the slow clock. □

This implementation uses only the chosen class, its realized reward, and the learner's current estimates. Faster adjustment lets the propensity estimator track the current policy. Note that the harmonic visit-count benchmark, whose historical frequencies adjust on the same timescale as valuations, has the joint asynchronous mean field (A.1).

A.6 Attainability and positive-probability convergence

Proposition A.10 (Capture of a stable SVE from every finite prior)**.** *Every SVE is attainable by the harmonic-visit process: from any finite initial valuations and positive prior counts, every neighborhood of its sufficient-statistic representation has positive probability of being visited at a deterministic date later than any prescribed date. Every locally asymptotically stable SVE of the joint mean field has positive probability of being the stochastic process's limit, from every such prior.*

Proof. Let v^* be an SVE, $x^* = \Lambda(v^*)$, and $y^* = x^* v^* = R(v^*)$. Freeze the logit policy at v^* and let $H = (e_s, re_s)$ be the one-period outcome vector in the (x, y) coordinates. The primitive reward assumptions give a finite expectation, and $\mathbb{E}[H] = (x^*, y^*) =: w^*$. The expectation of an integrable finite-dimensional random vector lies in the closed convex hull of its support: otherwise a separating linear functional would separate its expectation from all its realizations, which is impossible after taking expectations. Therefore every neighborhood of w^* contains a convex combination of finitely many supported outcome vectors. Approximate its weights by rational numbers, repeat the corresponding finite outcome word, and require each reward to lie in a sufficiently small bounded interval about the selected support point. Choose specific objective menus and implemented alternatives realizing those support points; their probabilities and the probabilities of the reward intervals are positive.

Every finite prescribed word of these events has positive probability under the actual adaptive process. Menus have their fixed positive probabilities, the chosen classes have positive logit probabilities at all finite valuations, and the selected implementation probabilities are positive. On the prescribed bounded reward events all valuations remain finite. Independence of the next reward conditional on the implemented alternative then gives positive probability to the entire word. In view of the exact formula

$$W_n = \frac{\chi W_0 + \sum_{k<n} H_{k+1}}{\chi + n},$$

an arbitrarily long repetition makes the prior negligible and places W_n in the desired neighborhood. This proves attainability, without requiring any individual reward to equal its expectation.

We verify that attainability implies capture under the present second-moment assumption. Let w^* be locally asymptotically stable. Choose a compact neighborhood K in its basin and a larger compact neighborhood K' in the domain $x_s > 0$. Stability and attraction give a time T and a tracking tolerance $\delta > 0$ such that every path starting sufficiently near w^* and remaining within δ of each successive length-T solution segment stays in K', and its values at the ends of these segments stay in K. This is the deterministic trapping argument for an asymptotically stable equilibrium; it follows by choosing K inside a stable neighborhood and using uniform attraction on K, or from Benaïm (1999, Lemma 6.8).

On K' the field is bounded and Lipschitz. For the martingale in (A.11), the conditional maximal inequality gives, at any deterministic date n,

$$\Pr\left(\sup_{l \geq n} \left\| \sum_{k=n}^{l} \gamma_{k+1} U_{k+1} \right\| > \eta \,\middle|\, \mathcal{F}_n \right) \leq \frac{C}{\eta^2} \sum_{k=n}^{\infty} \gamma_{k+1}^2,$$

where C is deterministic. On the complementary event all future window noise sums and individual noise increments are at most 2η. The integral-equation estimate in the proof of Proposition A.5, stopped upon leaving K', consequently bounds every length-T tracking error by $C_T(\eta + \sup_{k\geq n}\gamma_{k+1})$. Choose η small enough and then n large enough to make this less than δ and the displayed exception probability less than $1/2$. Successive trapping prevents exit from K'. On this event the almost-sure APT property and Benaïm (1999, Theorem 6.9) imply $W_k \to w^*$. Thus, conditional on reaching a sufficiently small neighborhood at a large deterministic date, convergence has probability at least $1/2$. Attainability supplies positive probability of that event. The change of coordinates $v = y/x$ yields convergence of the original state. □

A.7 A UCB-style vanishing exploration-bonus extension

Proposition A.11 (Vanishing bonuses in the update)**.** *Keep the same logit choice rule and harmonic gains, but replace the observed update target by $r_k + b_{k,s}$ when class s is selected. Suppose $b_{k,s}$ is $\mathcal{F}_k$-measurable and $\max_s |b_{k,s}| \to 0$ almost surely. Then boundedness, positive lower limiting frequencies, and the joint asymptotic-pseudo-trajectory conclusion continue to hold, with the same limiting ODE.*

Proof. The sample-average identity acquires the additional numerator $\sum_{j<k} I_{j,s} b_{j,s}$. On paths with infinitely many visits, its ratio to $\kappa_s + N_k(s)$ tends to zero by the Cesàro argument on the visit subsequence; on paths with finitely many visits it has a finite limit. The boundedness and continued-sampling proof is otherwise unchanged. In sufficient statistics the conditional drift gains $\Lambda_s(v_k)b_{k,s}$ in its y_s component, a term tending to zero. The additional centered noise is $(I_{k,s} - \Lambda_s(v_k))b_{k,s}$; after localization to bounded bonuses, its squared harmonic increments are summable. Its martingale series therefore converges. A drift remainder tending to zero contributes vanishing integral error on fixed learning-time windows, proving the APT assertion as before. □

B Equilibrium structure and strict-equilibrium refinement

This section proves Theorem 1 of the main text, the characterization and local-stability Theorem 2 for strict VE, and its full-coverage corollary. Throughout, the support $\Omega^+ := \mathrm{supp}(p)$ is finite, $p(\omega) > 0$ for $\omega \in \Omega^+$, and every class occurs in at least one supported menu. Set $n := |\mathcal{S}|$ and $\Omega_s^+ := \{\omega \in \Omega^+ : s \in \omega\}$. We use the following notation, displaying sensitivity when needed:

$$\sigma_\omega^s(v;\beta) = \mathbf{1}_{\{s\in\omega\}} \frac{\exp(\beta v_s)}{\sum_{j\in\omega}\exp(\beta v_j)}, \qquad \Lambda_s(v;\beta) = \sum_{\omega\in\Omega_s^+} p(\omega)\sigma_\omega^s(v;\beta), \tag{B.1}$$

$$w_s(\omega;v,\beta) = \frac{p(\omega)\sigma_\omega^s(v;\beta)}{\Lambda_s(v;\beta)}, \qquad g_s(v;\beta) = \sum_{\omega\in\Omega_s^+} w_s(\omega;v,\beta)\pi_s(\omega). \tag{B.2}$$

Let $m_s := \min_{\omega\in\Omega_s^+}\pi_s(\omega)$, $M_s := \max_{\omega\in\Omega_s^+}\pi_s(\omega)$, and $K := \prod_s [m_s, M_s]$. For clarity, a valuation equilibrium is a pair (v,σ) such that σ_ω is a probability vector supported on $\arg\max_{s\in\omega} v_s$ and, on writing $D_s(\sigma) := \sum_{\omega\in\Omega_s^+} p(\omega)\sigma_\omega^s$, for a class s whose $D_s(\sigma) > 0$,

$$v_s = \frac{1}{D_s(\sigma)} \sum_{\omega\in\Omega_s^+} p(\omega)\sigma_\omega^s \pi_s(\omega) \qquad (s \in \mathcal{S}). \tag{B.3}$$

If $D_s(\sigma) = 0$, this equation imposes no restriction on v_s. Write $\mathcal{VE}$ for the set of valuation vectors admitting such a profile, $\mathcal{V}(\beta) := \{v \in K : v = g(v;\beta)\}$ for the SVE correspondence for a fixed $\beta \in [0,\infty)$, and $\mathcal{V}(\infty)$ for its accumulation set as $\beta \to \infty$.

B.1 The SVE correspondence

Theorem B.1 (Equilibrium structure)**.** *For every fixed tree, the following statements hold.*

1. *For each finite $\beta \geq 0$, $\mathcal{V}(\beta)$ is nonempty, compact, and contained in* $\operatorname{ri} K$*. The correspondence is upper hemicontinuous. Its graph*

$$\mathcal{G} := \{(v, \beta) \in K \times [0, \infty) : v = g(v; \beta)\}$$

has finitely many path-connected components. The unique equilibrium at $\beta = 0$ belongs to the principal component whose projection covers all of $[0, \infty)$.

2. *The set $\mathcal{V}(\infty)$ is nonempty and compact, $\mathcal{V}(\infty) \subseteq \mathcal{VE}$, and*

$$\lim_{\beta \to \infty} \sup_{v \in \mathcal{V}(\beta)} \operatorname{dist}(v, \mathcal{V}(\infty)) = 0.$$

3. *Fix the menu support and vary (p, π) in the relative Euclidean parameter space specified below. On a comeager set of full Lebesgue measure, $\mathcal{G}$ is a one-dimensional real-analytic embedded manifold with its sole boundary point at $\beta = 0$. Except at finitely many sensitivities, all equilibria are hyperbolic and their number is finite and odd with each having a locally unique real-analytic continuation in β. The set $\mathcal{V}(\infty)$ is finite. The component containing the equilibrium at zero sensitivity is eventually a single real-analytic graph and has exactly one high-sensitivity limit.*

In particular, part (iii) implies hyperbolicity of every SVE at all sufficiently large finite sensitivities.

Proof. Existence, compactness, and upper hemicontinuity: At finite sensitivity, $\Lambda_s(v; \beta) > 0$ for every v and class s. The weights in (B.2) are positive and sum to one. Thus g is jointly real analytic in (v, β) and maps $\mathbb{R}^n$ into $\operatorname{ri} K$. Brouwer's theorem, applied on the possibly lower-dimensional compact convex set K, gives a fixed point. The fixed-point set is closed in K, hence compact. If $\beta_j \to \beta < \infty$ and $v_j \in \mathcal{V}(\beta_j)$, compactness of K and continuity of g imply that every limit point belongs to $\mathcal{V}(\beta)$. This proves upper hemicontinuity because K is a common compact codomain. At $\beta = 0$, g is independent of v, so there is exactly one equilibrium v^0 and $D_v(v - g(v; 0)) = I$.

Degree 1: Set $h(v, \beta) := v - g(v; \beta)$ and introduce the enlarged open box

$$Q := \prod_s (m_s - 1, M_s + 1).$$

For any $c \in K$, the homotopy $v - [(1 - t)c + tg(v; \beta)]$, $0 \leq t \leq 1$, has no zero on ∂Q: its bracketed term lies in $K \subset Q$. Homotopy invariance of Brouwer degree consequently gives

$$\deg(h(\cdot, \beta), Q, 0) = 1 \qquad (0 \leq \beta < \infty). \tag{B.4}$$

Whenever every equilibrium is regular, each is isolated. There are then finitely many: an infinite set in the compact zero set would have an accumulation point, which would itself be a non-isolated zero. Their local degrees are $\operatorname{sgn} \det D_v h \in \{-1, 1\}$ and sum to one. Their number is therefore odd. More generally, an index sum is asserted only when the equilibria are isolated.

Definability and continuation from zero sensitivity: The real field with exponentiation is o-minimal, and the formulas in (B.1)–(B.2) use only its definable operations. Thus $\mathcal{G}$ is definable in the o-minimal structure $\mathbb{R}_{\exp}$. We use the standard consequences of o-minimality that definable sets have finitely many connected components, each definably path connected; that the frontier has smaller dimension than a nonempty definable set; and that bounded one-variable definable functions

are eventually monotone and have limits at an endpoint. These facts follow from cell decomposition and monotonicity; see van den Dries, Macintyre, and Marker (1994); van den Dries (1998).

Fix $b > 0$ and let $\mathcal{G}_b := \mathcal{G} \cap (K \times [0, b])$. It is compact and has finitely many connected components. Suppose the component C_0 containing $(v^0, 0)$ did not meet $K \times \{b\}$. Because the components are finitely many disjoint compact sets, there exists an open neighborhood U of C_0 in $Q \times [0, b]$ such that $\overline{U} \cap \mathcal{G}_b = C_0$, $\overline{U}$ misses the top face, and the relative boundary of U contains no zero of h. For $U_t := \{v : (v, t) \in U\}$, the integer $\deg(h(\cdot, t), U_t, 0)$ is locally constant in t. Here is the usual excision argument for the varying slices: at a given t_0, surround the compact set of zeros in U_{t_0} by finitely many open sets with closures inside U_{t_0}; for sufficiently nearby t, these closures remain inside U_t and contain all its zeros. Otherwise a sequence of omitted zeros would converge to a zero on the relative boundary or to an omitted zero in U_{t_0}. Excision then reduces the degree to a fixed open set, where homotopy invariance applies. The same argument covers an empty zero set and the endpoints by one-sided neighborhoods. The degree at zero is $+1$, whereas at b it is zero, a contradiction. Thus C_0 meets the top face. As b is arbitrary, the component of the full graph containing $(v^0, 0)$ projects onto $[0, \infty)$. It is the only component with this property because the fiber at zero is a singleton.

High-sensitivity limits and VE consistency: Compactness of K makes $\mathcal{V}(\infty)$ nonempty. More explicitly it is the nested intersection $\mathcal{V}(\infty) = \bigcap_{b>0} \overline{\bigcup_{\beta \geq b} \mathcal{V}(\beta)}$, so it is compact. Let $\beta_j \to \infty$ and $v^j \in \mathcal{V}(\beta_j)$ converge to v^*. Passing to a subsequence, every menu's logit probability vector converges to a profile σ^*. If $v^*_s < \max_{r \in \omega} v^*_r$, the corresponding gap is bounded below by a positive constant for large j, whence $\sigma^s_\omega(v^j; \beta_j) \to 0$. Thus the limit profile chooses only maximizers. Multiplying the fixed-point equation by $\Lambda_s(v^j; \beta_j)$ and taking limits gives the balance identity $v^*_s D_s(\sigma^*) = \sum_{\omega \in \Omega^+_s} p(\omega) \sigma^{*,s}_\omega \pi_s(\omega)$. For $D_s(\sigma^*) > 0$, division gives (B.3); when $D_s(\sigma^*) = 0$, the identity imposes no restriction. Hence $v^* \in \mathcal{VE}$. If the stated uniform-distance conclusion failed, a sequence of equilibria at sensitivities tending to infinity would stay at a positive distance from $\mathcal{V}(\infty)$; its convergent subsequence in K contradicts the definition of that set. Moreover, $\mathcal{VE}$ is closed: the menu-wise optimality conditions and the balance identities above define a closed set of pairs (v, σ), and the space of choice profiles is compact.

Generic regularity of the entire equilibrium graph: The parameter space is

$$\Theta = \left\{ p \in \mathbb{R}^{\Omega^+} : p(\omega) > 0, \sum_{\omega \in \Omega^+} p(\omega) = 1 \right\} \times \mathbb{R}^{\{(s,\omega) : \omega \in \Omega^+,\ s \in \omega\}}.$$

It carries its relative Euclidean topology and Lebesgue measure in its affine coordinates. Write $d := \dim \Theta$. On the fixed open domain $\mathbb{R}^n \times (0, \infty) \times \Theta$, consider $\Phi(v, \beta, \theta) := v - g(v; \beta, \theta)$. For each s, choose one supported menu ω_s containing s. The derivative of Φ with respect to the n distinct payoff coordinates $\pi_s(\omega_s)$ is diagonal, with entries $-w_s(\omega_s; v, \beta) < 0$. Consequently Φ is a submersion, and $\mathcal{M} := \Phi^{-1}(0)$ is a smooth manifold of dimension $d + 1$.

Let $P : \mathcal{M} \to \Theta$ be parameter projection. A direct linear-algebra argument shows that θ is a regular value of P if and only if

$$\operatorname{rank} D_{(v,\beta)} \Phi(v, \beta, \theta) = n \quad \text{at every zero with } \beta > 0. \tag{B.5}$$

Indeed, write the derivative as $A\xi + B\zeta$. Surjectivity of A lets one solve $A\xi = -B\zeta$ for every parameter direction ζ, which is precisely surjectivity of DP on the tangent space of $\mathcal{M}$. Conversely, that solvability implies $\operatorname{im} B \subseteq \operatorname{im} A$; surjectivity of $[A\ B]$ then forces A to be surjective.

Sard's theorem says that the set of critical values of P has measure zero. It is also meager: exhaust $\mathcal{M}$ by countably many compact sets; the image of the closed critical set intersected with each such compact set is compact and null, hence nowhere dense. Their union is the critical-value

set. Therefore the *same* set of regular parameters is both comeager and of full measure. For each such parameter, (B.5) and the analytic regular-level-set theorem make the positive-sensitivity graph a one-dimensional analytic embedded manifold. At $\beta = 0$, $D_v h = I$, so the analytic implicit-function theorem supplies a boundary chart and no additional boundary points occur. Consequently, the graph has dimension one.

Exceptional sensitivities and finite limit sets: Fix a regular parameter and let $\xi : \mathcal{G} \to [0, \infty)$ be sensitivity projection. On the interior manifold, a value of ξ is regular exactly when $D_v h$ is invertible at every equilibrium in that fiber. To see this, a tangent vector (u, t) obeys $D_v h\, u + \partial_\beta h\, t = 0$. If $D_v h$ is invertible, a tangent vector with $t = 1$ exists. If it is singular, its kernel already gives a nonzero tangent vector with $t = 0$; the tangent space is one-dimensional, so $D\xi = 0$ there. The critical-value set is definable and null by Sard's theorem. A definable null subset of the real line is finite. At every other sensitivity, compactness, (B.4), and the analytic implicit-function theorem give the claimed finiteness, parity, and local branches. Compactify sensitivity by $t = (1 + \beta)^{-1}$ and put

$$\widehat{\mathcal{G}} := \{(v, t) \in K \times (0, 1] : (v, t^{-1} - 1) \in \mathcal{G}\}.$$

This is a one-dimensional definable set. Its frontier $\overline{\widehat{\mathcal{G}}} \setminus \widehat{\mathcal{G}}$ has dimension strictly below one and is therefore finite. The points added at $t = 0$ are exactly $\mathcal{V}(\infty) \times \{0\}$. This proves finiteness of $\mathcal{V}(\infty)$.

The unique limit of the principal component: Sensitivity projection is proper: its inverse image of a compact interval is closed in the compact product of that interval with K. Its restriction to a component is proper as well. The principal component C is a connected one-dimensional manifold with exactly one boundary point. Since its projection is unbounded, it is non-compact; the classification of connected one-dimensional manifolds with boundary identifies C with $[0, \infty)$. Properness implies that β tends to infinity along its sole non-compact end. On this component ξ is analytic and non-constant. Its critical points are isolated by the analytic identity theorem and form a definable set, so there are only finitely many. Choose a compact initial arc containing them all, and choose b greater than the largest sensitivity on that arc. On the remaining end ξ has nonzero derivative, is strictly monotone, and tends to infinity. Thus $C \cap (K \times (b, \infty))$ is a single graph $\{(\varphi(\beta), \beta) : \beta > b\}$, where φ is definable and real analytic. Each bounded coordinate of φ has a limit by definable monotonicity. The resulting vector is the unique high-sensitivity limit of C and belongs to $\mathcal{V}(\infty) \subseteq \mathcal{VE}$.

Generic hyperbolicity: We now strengthen regularity to hyperbolicity for every finite n. On the universal equilibrium manifold $\mathcal{M} = \Phi^{-1}(0)$ above, which has dimension $d + 1$, set

$$J(v, \beta, \theta) := D_v g(v; \beta, \theta) - I_n, \qquad H(v, \beta, \theta) := \det(J \otimes I_n + I_n \otimes J)\,, \tag{B.6}$$

where $\otimes$ denotes the Kronecker product. The eigenvalues of the matrix inside the determinant are the sums $\lambda_i + \lambda_j$ of eigenvalues of the real matrix J, counted with multiplicity. If J has a purely imaginary eigenvalue $i\omega$, including zero, then $-i\omega$ is also an eigenvalue, so $H = 0$. Only this implication is needed: H can also vanish at a hyperbolic matrix with opposite nonzero eigenvalues.

Fix any $(v, \beta, p, \pi) \in \mathcal{M}$ and, keeping v fixed as a center, vary the reduced payoffs according to

$$\pi_s^t(\omega) = v_s + t\big(\pi_s(\omega) - v_s\big). \tag{B.7}$$

This perturbation is admissible in the stated parameter space and leaves p unchanged. Since the conditional weights depend on valuations, sensitivity, and p, but not on payoffs, for every test valuation u we have

$$g(u; \beta, p, \pi^{\,t}) = v + t\big(g(u; \beta, p, \pi) - v\big).$$

Thus v remains an equilibrium for every t, and along this curve $J_t = tD_v g(v;\beta,p,\pi) - I_n$. The corresponding $H(t)$ is a polynomial in t with $H(0) = (-2)^{n^2} \neq 0$. Consequently, arbitrarily close to $t = 1$ there are points of $\mathcal{M}$ at which $H \neq 0$. The definable set $Z := \{(v,\beta,\theta) \in \mathcal{M} : H = 0\}$ therefore has empty relative interior in the $(d+1)$-dimensional manifold $\mathcal{M}$, and hence $\dim Z \leq d$.

For $\theta \in \Theta$, write $Z_\theta := \{(v,\beta) : (v,\beta,\theta) \in Z\}$. The definable fiber-dimension theorem implies that

$$E := \{\theta \in \Theta : \dim Z_\theta \geq 1\} \quad \text{is definable and satisfies} \quad \dim E \leq d-1.$$

Indeed, on the definable set of parameters with fiber dimension k, the total dimension is the parameter-set dimension plus k; for $k \geq 1$ this bounds the parameter-set dimension by $d-1$. Thus E has measure zero and is nowhere dense, and every fiber Z_θ outside E is finite (or empty), because a zero-dimensional definable set is finite. These are standard consequences of definable cell decomposition (van den Dries, 1998). Intersecting $\Theta \setminus E$ with the comeager full-measure parameter set already obtained above preserves all preceding conclusions. For each parameter in this intersection, the projection of Z_θ onto the sensitivity axis is finite and contains every positive sensitivity admitting a nonhyperbolic SVE. At $\beta = 0$, $D_v g = 0$ and $J = -I_n$, so hyperbolicity is automatic. Finally, Proposition A.7 shows that the joint linearization adds only $n-1$ eigenvalues equal to -1. The same conclusion therefore holds for both dynamics, with the finite exceptional set depending on (p,π). □

B.2 Two supporting claims in the equilibrium discussion

Lemma B.2 (A robust continuum of VE valuations)**.** *Consider three equally likely menus $\{A,B\}$, $\{B,C\}$, and $\{C,A\}$, with payoffs $(1,0)$ to the classes listed in that order. Every (t,t,t) with $0 < t < 1$ is a VE valuation. A local continuum of fully mixed VE valuations persists for every tree in an open neighborhood of these primitives, holding the three-menu support fixed. At the symmetric primitives, however, the unique SVE is $(1/2,1/2,1/2)$ at every finite sensitivity, so this is the unique LSVE.*

Proof. Let q_1, q_2, q_3 be the probabilities of choosing the first-listed class in the three menus. With all valuations equal to t, VE consistency at the symmetric primitives is

$$(1-t)q_1 + tq_3 = t, \qquad tq_1 + (1-t)q_2 = t, \qquad tq_2 + (1-t)q_3 = t.$$

The choice $q_1 = q_2 = q_3 = t$ solves these equations. Their derivative with respect to q has determinant $(1-t)^3 + t^3 > 0$. For each fixed $t_0 \in (0,1)$, the implicit-function theorem therefore gives a smooth interior solution $q = q(t,p,\pi)$ for every t in a neighborhood of t_0 and every (p,π) sufficiently near the symmetric primitives. The unnormalized consistency equations are smooth in these parameters; their derivative differs at the symmetric point from the displayed matrix only by a nonzero common scalar. Since all classes are tied, every such interior profile is greedy-optimal. Thus these are continua of distinct VE valuation vectors, not only continua of different mixing profiles at one valuation.

For finite $\beta > 0$, the assessment map at the symmetric primitives satisfies

$$\begin{aligned}
\operatorname{sgn}(g_A(v) - 1/2) &= \operatorname{sgn}(v_C - v_B),\\
\operatorname{sgn}(g_B(v) - 1/2) &= \operatorname{sgn}(v_A - v_C),\\
\operatorname{sgn}(g_C(v) - 1/2) &= \operatorname{sgn}(v_B - v_A).
\end{aligned}$$

These identities follow by comparing, for each class, its logit selection probabilities in its payoff-one and payoff-zero menus. At a fixed point choose a highest-valued class and, by cyclic relabeling, call it A. The second and third identities give $v_B \geq 1/2$ and $v_C \leq 1/2$. Hence $v_C \leq v_B$, and the first identity gives $v_A \leq 1/2$. Since $v_A \geq v_B$, both equal $1/2$. The first identity then forces $v_C = v_B$. Uniform choice at $\beta = 0$ yields the same unique fixed point directly. □

Lemma B.3 (The stated odds restrictions characterize logit)**.** *On a rich domain admitting every real valuation profile and menus with at least three distinct labels, suppose choice probabilities are positive and continuous. Suppose the odds of any two available classes with valuations a, b equal a common function $r(a, b)$, independent of the other available classes and of their labels. If choice is invariant to common valuation shifts and is weakly increasing in a class's own valuation, then, for some $\beta \geq 0$,*

$$\sigma^s_\omega(v) = \frac{\exp(\beta v_s)}{\sum_{j\in\omega} \exp(\beta v_j)}.$$

Conversely every such logit rule has these properties.

Proof. Consistency of odds in a three-class menu gives $r(a, c) = r(a, b)r(b, c)$ for every $a, b, c \in \mathbb{R}$. With $w(a) := r(a, 0) > 0$, this yields $r(a, b) = w(a)/w(b)$ and $w(0) = 1$. Common-shift invariance implies $r(a + c, c) = r(a, 0)$, or $w(a + c) = w(a)w(c)$. The continuous function $\log w$ is additive on $\mathbb{R}$, so $w(a) = \exp(\beta a)$ for a real constant β. Own-valuation monotonicity implies $\beta \geq 0$, and strict monotonicity rules out zero. Normalizing the pairwise odds gives the displayed formula. Its converse properties follow immediately. This characterization concerns the specified rich domain. □

Lemma B.4 (Closest-loss concentration)**.** *Suppose $\beta_j \to \infty$ and $v^j \to v^*$, without requiring that v^j be equilibria. Define*

$$\Delta_s(\omega; v) := \max_{r\in\omega} v_r - v_s, \qquad \Omega^\star_s(v^*) := \arg\min_{\omega\in\Omega^+_s} \Delta_s(\omega; v^*).$$

Then $\sum_{\omega\notin\Omega^\star_s(v^*)} w_s(\omega; v^j, \beta_j) \to 0$ for every class s. Consequently every limiting SVE satisfies

$$v^*_s \in \mathrm{co}\{\pi_s(\omega) : \omega \in \Omega^\star_s(v^*)\}.$$

Proof. The assertion is immediate if every supported menu containing s is a closest-loss menu. Otherwise let $\delta > 0$ be the smallest excess loss of a non-minimizing menu at v^*. By continuity and finiteness, for all large j each such menu has loss at least $\delta/2$ larger than the current minimum $d_j := \min_{\omega\in\Omega^+_s} \Delta_s(\omega; v^j)$. For every menu,

$$\frac{1}{n} e^{-\beta_j \Delta_s(\omega;v^j)} \leq \sigma^s_\omega(v^j; \beta_j) \leq e^{-\beta_j \Delta_s(\omega;v^j)}.$$

Choose a current minimizing menu and set $p_{\min,s} := \min_{\omega\in\Omega^+_s} p(\omega) > 0$. The denominator Λ_s is at least $(p_{\min,s}/n)e^{-\beta_j d_j}$, whereas the total numerator over nonminimizing menus is at most $e^{-\beta_j(d_j+\delta/2)}$. Their ratio is at most $(n/p_{\min,s})e^{-\beta_j\delta/2}$ and tends to zero. Passing to a subsequential limit of the finitely many weights and using $v^j = g(v^j; \beta_j)$ gives the convex-hull conclusion for SVE. □

B.3 Strict equilibria and the closest-loss characterization

A *strict VE* has a unique winner $a(\omega)$ in each supported menu and pairwise distinct valuations among its winner classes $\mathcal{W} := \{a(\omega) : \omega \in \Omega^+\}$. For such a vector, define

$$\bar{g}_s(v^*) := \frac{\sum_{\omega\in\Omega^\star_s(v^*)} p(\omega)\pi_s(\omega)}{\sum_{\omega\in\Omega^\star_s(v^*)} p(\omega)}. \tag{B.8}$$

Lemma B.5 (Uniform conditional-weight and derivative estimates)**.** *Let v^* be menu-wise strict. Suppose, for every s, there is a unique class*

$$t_s = \arg\min\{v^*_t : t = a(\omega),\ \omega \in \Omega^+_s\}.$$

This condition holds for every strict VE. There exist $r, \eta, C > 0$ such that, on $B := \{v : \|v - v^*\|_\infty \le r\}$ and for $\beta \ge 1$,

$$\sup_{v \in B} \sum_{\omega \in \Omega_s^+} \left| w_s(\omega; v, \beta) - \frac{p(\omega)\mathbf{1}_{\{\omega \in \Omega_s^\star(v^*)\}}}{\sum_{\omega' \in \Omega_s^\star(v^*)} p(\omega')} \right| \le Ce^{-\beta\eta}, \tag{B.9}$$

$$\sup_{v \in B} \|g(v;\beta) - \bar{g}(v^*)\|_\infty \le Ce^{-\beta\eta}, \tag{B.10}$$

$$\sup_{v \in B} \|D_v g(v;\beta)\|_\infty \le C\beta e^{-\beta\eta}. \tag{B.11}$$

Here the matrix norm is the induced maximum row-sum norm.

Proof. Write $\Omega_s^\star := \Omega_s^\star(v^*)$. The unique closest winner condition implies $\Omega_s^\star = \{\omega \in \Omega_s^+ : a(\omega) = t_s\}$. Choose r sufficiently small that, throughout B, all menu winners remain $a(\omega)$ and all the positive differences

$$v_{a(\omega)} - v_j \quad (j \in \omega \setminus \{a(\omega)\}), \qquad v_{a(\omega)} - v_{t_s} \quad (\omega \in \Omega_s^+ \setminus \Omega_s^\star)$$

are bounded below by one common $\eta > 0$. Empty collections of comparisons impose no restriction. Define

$$Z_\omega := 1 + \sum_{j \in \omega \setminus \{a(\omega)\}} e^{-\beta(v_{a(\omega)} - v_j)}.$$

Then $1 \le Z_\omega \le n$ and $|Z_\omega - 1| \le (n-1)e^{-\beta\eta}$. The exact identity

$$\sigma_\omega^s = e^{-\beta(v_{a(\omega)} - v_s)}/Z_\omega$$

allows the common factor $e^{-\beta(v_{t_s} - v_s)}$ to be canceled from the conditional weights. Put

$$a_{s\omega} := \frac{e^{-\beta(v_{a(\omega)} - v_{t_s})}}{Z_\omega}, \quad A_s := \sum_{\omega \in \Omega_s^+} p(\omega) a_{s\omega}, \quad P_s^\star := \sum_{\omega \in \Omega_s^\star} p(\omega) > 0.$$

Thus $w_s(\omega) = p(\omega)a_{s\omega}/A_s$. On $\Omega_s^\star$, $a_{s\omega} = 1/Z_\omega$, whereas outside it $a_{s\omega} \le e^{-\beta\eta}$. Consequently

$$A_s \ge P_s^\star/n, \qquad E_s := \sum_{\omega \in \Omega_s^+} p(\omega)|a_{s\omega} - \mathbf{1}_{\{\omega \in \Omega_s^\star\}}| \le ne^{-\beta\eta}, \qquad |A_s - P_s^\star| \le E_s.$$

Normalizing the weights changes their total variation by at most $2E_s/A_s \le (2n^2/P_s^\star)e^{-\beta\eta}$. This proves (B.9), and bounded payoffs give (B.10). The lower bound is for the *factored* denominator A_s; Λ_s itself need not be uniformly positive as sensitivity increases.

For the derivative estimate, differentiating the logit quotient and using exact payoff centering,

$$\begin{aligned} \partial_{v_k} g_s &= -\beta \sum_{\omega \in \Omega_s^+} w_s(\omega)\sigma_\omega^k(\pi_s(\omega) - g_s) \\ &= -\beta \sum_{\omega \in \Omega_s^+} w_s(\omega)(\sigma_\omega^k - \mathbf{1}_{\{k=t_s\}})(\pi_s(\omega) - g_s). \end{aligned} \tag{B.12}$$

Indeed, $\sum_\omega w_s(\omega)(\pi_s(\omega) - g_s) = 0$, so the terms constant across menus cancel. With $R := \max_s(M_s - m_s)$, the absolute row sum is bounded by

$$\beta R \sum_{\omega \in \Omega_s^+} w_s(\omega)\|\sigma_\omega - e_{t_s}\|_1.$$

On $\Omega_s^\star$, the last norm is at most $2(n-1)e^{-\beta\eta}$. On its complement it is at most two, and the total conditional mass there is $O(e^{-\beta\eta})$ by (B.9). This proves (B.11), uniformly on B. □

Theorem B.6 (Strict VE: necessity, sufficiency, and local stability)**.** *Let v^* be a strict VE. Then $v^* \in \mathcal{V}(\infty)$ if and only if $v^* = \bar{g}(v^*)$. Whenever this condition holds, a fixed neighborhood of v^* contains exactly one SVE v^β for every sufficiently large β, and*

$$\|v^\beta - v^*\|_\infty \le Ce^{-\beta\eta}$$

for some constants $C, \eta > 0$. This equilibrium is locally exponentially stable under the normalized system $\dot{v} = g(v;\beta) - v$, and $(v^\beta, \Lambda(v^\beta;\beta))$ is locally exponentially stable under the joint system

$$\dot{v}_s = \frac{\Lambda_s(v;\beta)}{x_s}(g_s(v;\beta) - v_s), \qquad \dot{x}_s = \Lambda_s(v;\beta) - x_s, \qquad x \in \operatorname{ri}\Delta(\mathcal{S}).$$

Proof. The strict ordering of winner classes makes each t_s in Lemma B.5 unique. Suppose first that equilibria $v^j \to v^*$ occur along $\beta_j \to \infty$. Eventually they belong to the fixed ball B of that lemma. The fixed-point identity and (B.10) imply

$$\|v^j - \bar{g}(v^*)\|_\infty \le Ce^{-\beta_j\eta} \longrightarrow 0.$$

Thus $v^* = \bar{g}(v^*)$, proving necessity even for classes that are never chosen at the limiting VE.

Conversely, suppose $v^* = \bar{g}(v^*)$. For all sufficiently large β, the lemma gives $g(B;\beta) \subseteq B_{r/2}(v^*)$ and $\sup_B \|D_v g\|_\infty \le q_\beta < 1/2$. Since B is convex, g is a contraction on B, so Banach's theorem gives exactly one fixed point there. It belongs to K because every value of g belongs to K. Evaluating (B.10) at the fixed point proves the exponential approximation bound, and hence the LSVE property. In particular the interior of B is the fixed neighborhood required in the paper. On each boundary face of B the normalized field points strictly inward. Thus B is positively invariant. If $v(t)$ starts in B, variation of constants for $z(t) := v(t) - v^\beta$ and the contraction bound give

$$e^t\|z(t)\|_\infty \le \|z(0)\|_\infty + q_\beta \int_0^t e^u\|z(u)\|_\infty\, du.$$

Gronwall's inequality yields $\|z(t)\|_\infty \le e^{-(1-q_\beta)t}\|z(0)\|_\infty$. This proves local exponential stability for the normalized system. Equivalently, every eigenvalue of $Dg(v^\beta) - I$ has real part at most $-1 + q_\beta < -1/2$.

At a joint equilibrium $x^\beta = \Lambda(v^\beta;\beta)$, the factors Λ_s/x_s equal one and all prediction errors vanish. In local coordinates on the simplex tangent space, the joint derivative is the block-triangular matrix

$$\begin{pmatrix} Dg(v^\beta) - I & 0 \\ D\Lambda(v^\beta) & -I \end{pmatrix}. \tag{B.13}$$

Its extra $n-1$ eigenvalues are -1. It is therefore Hurwitz, and the standard linearization theorem for a smooth ODE gives local exponential stability of the joint equilibrium. □

Corollary B.7 (Coverage, including full singleton support)**.** *If a menu-wise strict VE selects every class in at least one supported menu, it is an LSVE and all the nearby-equilibrium and local-stability conclusions of Theorem B.6 hold. In particular this applies when every class has a supported singleton menu. Distinct winners need not have distinct valuations in this corollary.*

Proof. For each s, some menu is won by s. In every other menu containing s, the winner either is s or has valuation strictly above v_s^*. Thus the unique closest winner is $t_s = s$, and the closest-loss menus are exactly those won by s. VE consistency gives $v_s^* = \bar{g}_s(v^*)$. The uniform estimates and the sufficiency proof above apply. A class always wins its own singleton menu, proving the assertion. □

Lemma B.8 (Generic ordering of winning classes)**.** *For generic reduced-tree parameters, every menu-wise strict VE orders its winning classes strictly.*

Proof. There are finitely many pure menu-choice profiles. Fix one and write A_s for the nonempty collection of menus it assigns to winner s. VE consistency determines

$$v_s = \frac{\sum_{\omega \in A_s} p(\omega)\pi_s(\omega)}{\sum_{\omega \in A_s} p(\omega)}.$$

For two distinct winners s, t, their equality is a smooth hypersurface in parameter space: its derivative with respect to any $\pi_s(\omega)$ with $\omega \in A_s$ is nonzero. It is relatively closed, nowhere dense, and Lebesgue null. The union over the finitely many profiles and winner pairs has the same properties. Outside that union, no menu-wise strict VE has tied winner valuations. □

C Constant competitive-menu payoffs and two-class convergence

Suppose $\pi_s(\omega) = b_s$ whenever $s \in \omega$ and $|\omega| \geq 2$. Put $a_s = p(\{s\})$, interpreting this probability as zero if the singleton is unsupported, and put $d_s = a_s(\pi_s(\{s\}) - b_s)$, with $d_s = 0$ when $a_s = 0$. If a class appears only alone, choose any b_s and define d_s by the same identity. Then

$$R_s(v) = b_s\Lambda_s(v) + d_s, \qquad g_s(v) = b_s + \frac{d_s}{\Lambda_s(v)}. \tag{C.1}$$

Theorem C.1 (Constant competitive-menu payoffs)**.** *For each fixed finite β, every normalized trajectory approaches the SVE set, and every joint trajectory approaches the corresponding joint equilibrium set. If the SVE set is finite, each trajectory converges to one equilibrium. For each fixed β, finiteness holds generically within the family satisfying the stated payoff restriction. If $d_s \geq 0$ for every class, the SVE is unique.*

Proof. The case $\beta = 0$ is the constant-target benchmark. Assume $\beta > 0$. Let $\rho = (\rho^s_\omega)$ be an interior menu-wise choice profile and set $X_s(\rho) = \sum_{\omega \ni s} p(\omega)\rho^s_\omega$. Consider

$$\Phi(\rho) = \sum_s \{b_s X_s(\rho) + d_s \log X_s(\rho)\} - \frac{1}{\beta}\sum_\omega p(\omega) \sum_{s \in \omega} \rho^s_\omega \log \rho^s_\omega. \tag{C.2}$$

All expressions below are evaluated on compact interior sets, so their derivatives are bounded. In a simplex tangent direction its gradient is represented by

$$\frac{1}{p(\omega)}\frac{\partial \Phi}{\partial \rho^s_\omega} = b_s + \frac{d_s}{X_s(\rho)} - \frac{1}{\beta}(1 + \log \rho^s_\omega).$$

Normalized dynamics. Set $\rho = \sigma(v)$ and $F = g(v) - v$. Since $X(\rho) = \Lambda(v)$ and $\log \sigma^s_\omega = \beta v_s - \log \sum_{j \in \omega} e^{\beta v_j}$, the simplex gradient is represented by F_s plus a menu constant. Differentiating the logits gives

$$\dot\rho^s_\omega = \beta\rho^s_\omega \left(F_s - \sum_{j \in \omega} \rho^j_\omega F_j \right), \qquad \frac{d}{dt}\Phi(\sigma(v)) = \beta \sum_\omega p(\omega) \operatorname{Var}_{\sigma_\omega(v)}(F) \geq 0.$$

Valuations stay in a compact box. LaSalle's invariance argument places their omega-limit set in the zero-derivative set. There F_s is constant on each connected component of the class co-occurrence

graph. Along a trajectory remaining in this set, all within-component valuation gaps are constant, so g is constant and the remaining component-wise common shift solves $\dot{c} = -c$. The only such complete trajectories in a compact set are stationary. Thus the largest compact invariant subset of the zero-derivative set consists of SVE, proving approach to the SVE set.

Joint dynamics. In sufficient statistics, $e_s = y_s - b_s x_s - d_s$ satisfies $\dot{e}_s = -e_s$ by (C.1). Introduce an auxiliary profile, with any interior initial condition, through

$$\dot{\rho}^s_\omega = \sigma^s_\omega(v) - \rho^s_\omega.$$

It remains a menu-wise probability profile. Moreover $X(\rho) - x$ solves the same equation $\dot{h} = -h$. The compact valuation box bounds all logits away from zero, so eventually ρ, $X(\rho)$, and x stay in compact interior sets. Consequently

$$b_s + d_s / X_s(\rho) - v_s = O(e^{-t}).$$

Writing $\sigma = \sigma(v)$ and using the logit odds identity gives

$$\begin{aligned} \frac{d}{dt}\Phi(\rho(t)) &= \frac{1}{\beta}\sum_\omega p(\omega) \sum_{s\in\omega} (\sigma^s_\omega - \rho^s_\omega) \log \frac{\sigma^s_\omega}{\rho^s_\omega} + O(e^{-t}) \\ &=: D(t) + O(e^{-t}), \qquad D(t) \geq 0. \end{aligned} \tag{C.3}$$

The error is absolutely integrable and Φ is bounded along the trajectory. Thus $\int_0^\infty D(t)\,dt < \infty$. Smoothness on the eventual compact set makes D uniformly continuous; the elementary integrability argument for a nonnegative uniformly continuous function gives $D(t) \to 0$. Each summand is nonnegative and vanishes precisely when its two arguments agree. Hence $\rho - \sigma(v) \to 0$ and $x - \Lambda(v) \to 0$. The decaying e_s and (C.1) now imply $v - g(v) \to 0$. Compactness proves approach to the joint equilibrium set.

If that set is finite, the connectedness of an omega-limit set gives convergence to one point. For generic finiteness at a fixed sensitivity, vary every payoff of class s by an independent constant t_s. This operation preserves constant competitive-menu payoffs and changes g_s by exactly t_s. The equilibrium equation is therefore a submersion in these class shifts. Parametric regular-value theory gives regular, hence isolated, equilibria for a comeager full-measure set of parameters; compactness makes their number finite. The parameter topology is the relative Euclidean topology of this restricted payoff family.

Finally, when all $d_s \geq 0$, the first sum in (C.2) is concave and the entropy term is strictly concave on the product of menu simplices. At any SVE its profile is an interior stationary point of Φ; strict concavity gives at most one such profile. That profile determines g and therefore the valuations uniquely. Existence was proved above. This also covers classes with $d_s = 0$. □

C.1 Robustness to payoff heterogeneity

The robustness argument concerns the entire recurrent set, not merely the continuation of individual equilibria. We use the following perturbation fact on a fixed compact trapping region. All derivatives and hyperbolicity statements are understood in the affine state space; in particular, frequency and menu-probability vectors are restricted to their respective simplex hyperplanes.

Lemma C.2 (Persistence of a finite hyperbolic recurrent set)**.** *Let F_θ be C^1 vector fields on a neighborhood of a compact set K, with $F_\theta \to F_0$ in C^1 there as $\theta \to 0$. Suppose K is positively invariant for every sufficiently small θ. If the chain-recurrent set of the benchmark semiflow on K*

consists of finitely many hyperbolic equilibria in the interior of K, then, for sufficiently small θ, its counterpart consists precisely of their hyperbolic continuations. Their stable and unstable dimensions are unchanged, and every trajectory in K converges to one of them.

Proof. Write φ_t^θ for the semiflow. A point is chain recurrent if, for every $\epsilon, T > 0$, it can return to itself by a finite sequence of orbit segments of lengths at least T, with jumps smaller than ϵ between successive segments. The segment lengths may always be taken in $[T, 2T]$ by subdivision.

The chain-recurrent set is upper semicontinuous in θ. To see this, suppose $\theta_j \to 0$ and chain-recurrent points $z_j \in K$ converge to z. Fix $\epsilon, T > 0$. Continuous dependence of solutions gives uniform convergence of $\varphi_t^{\theta_j}$ to φ_t^0 on $[0, 2T] \times K$. A sufficiently accurate returning chain for z_j therefore becomes an (ϵ, T)-chain for the benchmark after replacing its initial and final points by z. Uniform continuity of φ_t^0 on the same compact set controls the change in the first segment. Thus z is benchmark chain recurrent. Compactness proves that every open neighborhood of the benchmark recurrent set contains the perturbed recurrent set for all sufficiently small θ.

Let the benchmark equilibria be $z_1, \dots, z_N$. Hyperbolicity and the implicit function theorem give continuations $z_i(\theta)$ with the same stable and unstable dimensions. Choose pairwise disjoint compact neighborhoods U_i in the interior of K sufficiently small that the only complete orbit remaining in U_i is $z_i(\theta)$, uniformly for small θ. This familiar local consequence of hyperbolicity can also be seen from an indefinite quadratic form: in stable and unstable coordinates, take the difference of positive quadratic forms on the unstable and stable subspaces. Its derivative is positive definite for the linearization, and remains so near the continued equilibrium for C^1-close fields. Along a complete orbit confined to this neighborhood, boundedness of the quadratic form makes the squared distance to the equilibrium integrable over $\mathbb{R}$. Uniform continuity implies that this distance tends to zero in both time directions. The quadratic form consequently has the same limit zero at both ends, forcing the orbit to be stationary.

By upper semicontinuity, the perturbed recurrent set lies in $\bigcup_i \operatorname{int} U_i$. It is invariant, and continuity of orbits prevents a complete orbit in this set from moving between the disjoint neighborhoods. The local property just proved therefore makes it exactly $\{z_i(\theta) : 1 \le i \le N\}$. Every trajectory in K has a nonempty connected omega-limit set contained in the chain-recurrent set. That limit set is a singleton, proving convergence. □

Corollary C.3 (Robustness to payoff heterogeneity)**.** *Fix a finite sensitivity $\beta \ge 0$ and the menu distribution p. Suppose the benchmark satisfies constant competitive-menu payoffs and every SVE is regular:*

$$\det\big(I - Dg(v^*)\big) \neq 0 \quad \text{for every } v^* \in \mathcal{V}(\beta). \tag{C.4}$$

There exists $\varepsilon > 0$ such that every conditional-mean payoff specification $\widetilde{\pi}$ with

$$\|\widetilde{\pi} - \pi\|_\infty := \max_{\substack{\omega \in \Omega^+ \\ s \in \omega}} |\widetilde{\pi}_s(\omega) - \pi_s(\omega)| < \varepsilon$$

has the same finite number of SVE as the benchmark, with unchanged stable and unstable dimensions under each mean field. Every normalized trajectory from finite valuations, and every joint trajectory from finite valuations and strictly positive frequencies, converges to an equilibrium. Condition (C.4) holds on a comeager, full-measure subset of the restricted benchmark payoff family at each fixed β and p.

Proof. If $\beta = 0$, all choice probabilities and sampled-payoff targets are constant, for every payoff specification. The normalized equation is a stable linear filter. In joint sufficient statistics both x

and y are stable linear filters with constant targets. There is one equilibrium, all its eigenvalues are -1, and every trajectory converges. Henceforth assume $\beta > 0$.

Step 1: common compact trapping regions and continuous dependence. Restrict provisionally to $\|\widetilde{\pi} - \pi\|_\infty < 1$, and put

$$m_s = \min_{\omega \ni s} \pi_s(\omega), \qquad M_s = \max_{\omega \ni s} \pi_s(\omega), \qquad B = \prod_s [m_s - 2, M_s + 2].$$

Every perturbed target $\widetilde{g}_s(v)$ lies in $[m_s - 1, M_s + 1]$. Thus B is strictly inward pointing for the normalized valuation field and for the valuation coordinates of the joint field. It is also eventually entered from every finite initial valuation vector. For the joint system, first use the invariant box spanned by the initial values and $[m_s - 1, M_s + 1]$. On this box Λ_s has a positive lower bound, so the rates Λ_s / x_s are bounded below, since $x_s \leq 1$. Coordinate comparison then gives finite-time entry into the enlarged box B. This argument uses the enlargement, not finite-time entry into the payoff box itself.

Let

$$\lambda = \min_{s,\, v \in B} \Lambda_s(v) > 0, \qquad 0 < \delta_x < \lambda.$$

The set $K_J = B \times \{x \in \Delta(\mathcal{S}) : x_s \geq \delta_x\}$ is compact and positively invariant, and every joint trajectory eventually enters it. Indeed, after entry into B, the equation $\dot{x}_s = \Lambda_s(v) - x_s$ forces x_s above δ_x in finite time. The normalized system uses $K_N = B$. Neither region depends on $\widetilde{\pi}$. On these regions the corresponding vector fields converge in C^1 to their benchmark fields as $\widetilde{\pi} \to \pi$: the weights and denominators in g are smooth, the denominators are bounded away from zero, and the reward numerators depend linearly on the conditional means.

Step 2: regularity implies hyperbolicity in the benchmark. Define

$$L(v) = \frac{1}{\beta} \sum_\omega p(\omega) \log \sum_{s \in \omega} e^{\beta v_s}.$$

Then $\Lambda = \nabla L$ and $H(v) := D\Lambda(v) = D^2 L(v)$ is symmetric positive semidefinite. With $D(v) = \operatorname{diag}(d_s / \Lambda_s(v)^2)$, identity (C.1) gives

$$Dg(v) = -D(v)H(v), \qquad J_N(v) = -I - D(v)H(v). \tag{C.5}$$

The matrices DH and $H^{1/2} D H^{1/2}$ have the same characteristic polynomial, by $\det(tI - AB) = \det(tI - BA)$. The latter is symmetric. Consequently every eigenvalue of J_N is real, even if some d_s are negative or H is singular. Regularity excludes zero eigenvalues and therefore makes every normalized equilibrium hyperbolic. Proposition A.7 adds only $|\mathcal{S}| - 1$ stable eigenvalues equal to -1 for the joint equilibrium. Compactness of the equilibrium set and regularity make the number of equilibria finite.

Step 3: the normalized benchmark has no other chain recurrence. For a compact positively invariant region K, its maximal invariant set $\mathcal{A} = \bigcap_{t \geq 0} \varphi_t(K)$ consists of the complete orbits that remain in K. It attracts K, contains its chain-recurrent set, and the latter is internally chain recurrent. We use the finite-critical-values Lyapunov criterion: a strict Lyapunov function on $\mathcal{A}$, with finitely many equilibrium values, excludes chain recurrence outside equilibria (Benaïm, 1999, Theorem 5.5 and Proposition 6.4). One can see the exclusion directly by choosing, between the function's value at a nonequilibrium point and its value later on that orbit, a level distinct from all equilibrium values. Within the compact recurrent set, the corresponding upper level set traps all sufficiently accurate chains once they enter it, preventing a return to the original point. The finiteness qualification matters; monotonicity of a scalar function alone would not suffice.

On K_N, the preceding potential calculation gives

$$\frac{d}{dt}\Phi(\sigma(v)) = \beta \sum_{\omega} p(\omega)\,\mathrm{Var}_{\sigma_\omega(v)}(g(v) - v) \geq 0.$$

Suppose its derivative vanishes at v. Write $c = g(v) - v$. The entries of c coincide within each connected component of the class co-occurrence graph. Hence every menu's logits are unchanged by adding c, and $u := v + c = g(v)$ satisfies $g(u) = u$. The entire orbit through v is

$$v(t) = u + e^{-t}(v - u), \qquad t \in \mathbb{R},$$

because the logits and g stay constant along these component-wise translations. If $v \neq u$, this orbit is unbounded backwards in time. It cannot lie in the maximal invariant subset $\mathcal{A}_N$ of K_N. Therefore, on $\mathcal{A}_N$, the potential increases strictly outside the finite equilibrium set. The Lyapunov criterion makes that set the entire chain-recurrent set on K_N.

Step 4: the joint benchmark has no other chain recurrence. Adjoin the auxiliary menu-probability profile

$$\dot{\rho}^s_\omega = \sigma^s_\omega(v) - \rho^s_\omega. \tag{C.6}$$

This equation does not affect (v, x) and is adjoined to both the benchmark and perturbed systems. Each equilibrium has the unique lift $\rho = \sigma(v)$, and the augmented Jacobian adds only $\sum_\omega(|\omega| - 1)$ eigenvalues equal to -1. Choose $\delta_\rho > 0$ smaller than $\min_{v\in B,\,\omega,\,s\in\omega} \sigma^s_\omega(v)$. Taking the product of K_J with the menu simplices restricted to $\rho^s_\omega \geq \delta_\rho$ gives a common compact trapping region K_L for the augmented systems. The linear filters in (C.6) give eventual entry, and the augmented fields converge in C^1 on a neighborhood of K_L.

In the benchmark, let $y_s = x_s v_s$ and

$$e_s = y_s - b_s x_s - d_s, \qquad h_s = X_s(\rho) - x_s.$$

As in the potential proof, $\dot{e} = -e$ and $\dot{h} = -h$. Every complete orbit in K_L therefore has $e = h = 0$. On its maximal invariant subset $\mathcal{A}_L$ we consequently have $x = X(\rho)$ and $v_s = b_s + d_s/X_s(\rho)$. The potential identity is now exact, with no error term:

$$\frac{d}{dt}\Phi(\rho) = \frac{1}{\beta}\sum_{\omega} p(\omega) \sum_{s\in\omega} (\sigma^s_\omega(v) - \rho^s_\omega) \log \frac{\sigma^s_\omega(v)}{\rho^s_\omega} \geq 0. \tag{C.7}$$

Equality holds exactly when $\rho = \sigma(v)$. Together with $e = h = 0$, this implies $x = \Lambda(v)$ and $v = g(v)$, so the augmented state is an equilibrium. The potential is thus strict off the finite augmented equilibrium set on $\mathcal{A}_L$. The same Lyapunov criterion shows that this is the entire chain-recurrent set on K_L.

Step 5: perturbation and genericity. Steps 1–4 verify all hypotheses of Lemma C.2, separately for the normalized field on K_N and the augmented joint field on K_L. All benchmark equilibria are in the interiors of these regions; shrink the payoff neighborhood to one that works for both applications. Every perturbed trajectory in either region converges to a continued equilibrium. Eventual entry extends the conclusion to arbitrary finite valuations and positive initial frequencies. Projecting an augmented trajectory onto (v, x) proves the claimed joint convergence. The implicit-function continuations of $g(v) - v = 0$ give the same valuation equilibria in both systems. Their numbers and stable and unstable dimensions are unchanged; the auxiliary profile only adds stable directions.

Finally, the class-wise payoff shifts used in the proof of Theorem C.1 preserve the benchmark restriction and shift g by arbitrary constant vectors. Parametric regular-value theory therefore gives (C.4) on a comeager, full-measure subset of the benchmark family, with p and β fixed. This proves the generic form of the main-text corollary. □

The neighborhood in Corollary C.3 is unrestricted in the conditional means: competitive-menu means need not remain equal, and singleton means may also change.

C.2 Convergence with at most two classes

Corollary C.4 (At most two classes)**.** *For every decision tree with at most two classes and every finite $\beta \geq 0$, every trajectory of either mean field converges to an equilibrium. If the equilibrium is unique, it is globally asymptotically stable under both systems.*

Proof. Only one competitive class menu is possible, so Theorem C.1 applies. If there are two classes, write $u = v_1 - v_2$. The normalized gap field is $h(u) = g_1(u) - g_2(u) - u$. It is real analytic on $\mathbb{R}$ and is not identically zero because g is bounded. All its zeros lie in a compact interval, so it has finitely many zeros. Each zero determines exactly one valuation equilibrium. The one-class case is immediate.

For completeness, suppose the equilibrium is unique. If the binary menu is absent, both targets are constant and Lyapunov stability is immediate. Otherwise write $p > 0$ for its probability and define

$$Q := \frac{x_1 - a_1}{p}, \qquad e_s := x_s(v_s - b_s) - d_s.$$

Thus Q is a historical-frequency coordinate. Since $a_1 + a_2 + p = 1$, we have

$$x_1 = a_1 + pQ, \qquad x_2 = a_2 + p(1 - Q).$$

The joint equations give the exact identities

$$\dot{e}_s = -e_s, \qquad \dot{Q} = \ell_\beta\left(b_1 + \frac{d_1 + e_1}{a_1 + pQ} - b_2 - \frac{d_2 + e_2}{a_2 + p(1 - Q)}\right) - Q,$$

where $\ell_\beta(w) = (1 + e^{-\beta w})^{-1}$.

Let $H(Q)$ denote the right-hand side with $e_1 = e_2 = 0$. Its zeros in $(0, 1)$ correspond exactly to joint equilibria, so it has a unique zero $Q^* \in (0, 1)$. Moreover, $H(0) > 0$ and $H(1) < 0$, using the continuous extension of d_s/x_s when $a_s = d_s = 0$. Hence H is positive below Q^* and negative above it. Choose any sufficiently small closed interval about Q^*. Its endpoint fields point strictly inward, and both denominators are bounded away from zero on that interval. By continuity, these inequalities persist for sufficiently small discrepancies (e_1, e_2). Since $e_s(t) = e_s(0)e^{-t}$, sufficiently small initial discrepancies remain small for all time, and the interval is forward invariant. Finally,

$$v_s = b_s + \frac{d_s + e_s}{x_s}$$

shows that (Q, e_1, e_2) are smooth local coordinates near the equilibrium. Taking the interval and discrepancies sufficiently small therefore proves joint Lyapunov stability.

For the normalized system the scalar gap field h has the same inward sign property about its unique zero; the remaining valuation follows a stable linear filter. This proves normalized Lyapunov stability as well. The already proved global attraction completes the claim. □

C.3 The Bordeaux–Chianti illustrations

For the three equally likely menus $\{B, C\}$, $\{B\}$, and $\{C\}$, competitive payoffs $(2, 1)$, and common singleton payoff z, put $q = \ell_\beta(v_B - v_C)$. Then

$$g_B = \frac{2q + z}{1 + q}, \qquad g_C = \frac{1 - q + z}{2 - q}.$$

At any SVE the normalized eigenvalues are

$$-1, \qquad -1+\beta q(1-q)\left\{\frac{2-z}{(1+q)^2}+\frac{1-z}{(2-q)^2}\right\}. \tag{C.8}$$

Indeed $Dg = g'(q)\beta q(1-q)(1,-1)$ has rank at most one. The joint system has one further eigenvalue -1.

When $z=0$, the strict equilibria associated with the two rankings are $(1,0)$ and $(0,1/2)$. They produce two locally exponentially stable branches at high sensitivity by the strict-equilibrium theorem. Every interior limiting probability must solve $g_B(q)=g_C(q)$, or $q^2-4q+1=0$. The only root in $(0,1)$ is $q_0=2-\sqrt{3}$, with common valuation $1-1/\sqrt{3}$. The derivative of g_B-g_C is strictly positive on $[0,1]$. The implicit function theorem applied to

$$g_B(q)-g_C(q)-\beta^{-1}\log\frac{q}{1-q}=0$$

therefore gives exactly one branch near q_0; its second eigenvalue in (C.8) is positive for sufficiently large β. Compactness of $[0,1]$ and the three possible limiting probabilities show that these three branches exhaust the SVE for all sufficiently large sensitivities. The stable branches have both normalized eigenvalues tending to -1, since $\beta q(1-q)\to 0$ there.

When $z=11/3$, both d_s are positive. The SVE is unique, globally attracting, and locally exponentially stable at every finite sensitivity; for $\beta>0$ its second normalized eigenvalue is strictly below -1. The limiting probability cannot be zero or one, because the corresponding limiting valuation gap would have the opposite sign. Its interior limit solves $g_B(q)=g_C(q)$, with unique solution $q=2/3$ and valuations $(3,3)$. This proves the indifference and mixing values used.

C.4 A useful lower bound on the size of a cycling example

Lemma C.5 (At most one class with heterogeneous menu means)**.** *If g_s is constant for all but at most one class, every trajectory of either mean field converges to an equilibrium at every finite sensitivity. Consequently cycling requires at least three classes and at least five objective alternatives. Deleting any supported menu from the three-menu construction below eliminates cycling.*

Proof. For each constant-target class, normalized valuations converge exponentially to their target; under the joint system the identity $x_s(v_s-c_s)=e^{-t}x_s(0)(v_s(0)-c_s)$ gives the same conclusion because x_s stays away from zero. For the remaining class j, after rescaling time by its bounded, strictly positive joint coefficient Λ_j/x_j (or without rescaling in the normalized system), its equation is an asymptotically autonomous scalar equation with limiting field $g_j(v_j,c_{-j})-v_j$. This field is analytic, bounded minus v_j, and has only finitely many zeros in the trapping interval. A scalar asymptotic pseudotrajectory with compact range has its limit set among these zeros: away from a zero the field has a fixed nonzero sign, which excludes chain recurrence there. Connectedness gives convergence to one zero. The frequency filters then converge as well.

With at most two classes the preceding corollary applies. With three or more classes but at most four alternatives, at most one class can contain more than one alternative. Every other class has a constant conditional mean, so the lemma applies. The five-alternative construction attains this lower bound. Removing either of its binary menus leaves constant competitive-menu payoffs; removing its singleton leaves only B with heterogeneous menu means. These observations prove the last assertion. The alternative-count claim concerns fixed objective reward laws and a fixed within-class implementation rule, as maintained here. □

D Corrective feedback and singleton-payoff shifts

Throughout this section, $\Omega^+ = \{\omega : p(\omega) > 0\}$, $\Omega_s^+ = \{\omega \in \Omega^+ : s \in \omega\}$, and

$$\Lambda_s(v) = \sum_{\omega \in \Omega_s^+} p(\omega)\sigma_\omega^s(v), \qquad g_s(v) = \frac{\sum_{\omega \in \Omega_s^+} p(\omega)\sigma_\omega^s(v)\pi_s(\omega)}{\Lambda_s(v)}, \qquad F(v) = g(v) - v.$$

The sensitivity $\beta < \infty$ is fixed unless otherwise indicated. Every class occurs in a supported menu, so $\Lambda_s(v) > 0$ at every finite v. The normalized dynamics are $\dot{v} = F(v)$. The coefficient-one harmonic visit-count dynamics are

$$\dot{v}_s = \frac{\Lambda_s(v)}{x_s} F_s(v), \qquad \dot{x}_s = \Lambda_s(v) - x_s, \qquad x \in \operatorname{int} \Delta(\mathcal{S}). \tag{D.1}$$

Write $m_s = \min_{\Omega_s^+} \pi_s$, $M_s = \max_{\Omega_s^+} \pi_s$, and $K = \prod_s [m_s, M_s]$; some coordinate intervals may be singletons.

D.1 Negative payoff–crowding covariance

Lemma D.1 (Derivative and row-sum identities)**.** *For distinct s, r,*

$$\partial_{v_r} g_s(v) = \frac{\beta}{\Lambda_s(v)} \sum_{\omega \in \Omega_s^+} p(\omega)\sigma_\omega^s(v)\sigma_\omega^r(v)[g_s(v) - \pi_s(\omega)] \tag{D.2}$$

$$= -\beta \operatorname{Cov}_{w_s^v}(\pi_s(\omega), \sigma_\omega^r(v)), \qquad w_s^v(\omega) = \frac{p(\omega)\sigma_\omega^s(v)}{\Lambda_s(v)}. \tag{D.3}$$

Here $\sigma_\omega^r = 0$ if $r \notin \omega$. Moreover, $Dg(v)\mathbf{1} = 0$ and $DF(v)\mathbf{1} = -\mathbf{1}$.

Proof. For $s \neq r$, $\partial_{v_r} \sigma_\omega^s = -\beta \sigma_\omega^s \sigma_\omega^r$. Differentiating the quotient defining g_s gives (D.2). Since $\sum_\omega w_s^v(\omega)\pi_s(\omega) = g_s(v)$, the same expression is the covariance in (D.3). Logit probabilities are unchanged by $v \mapsto v + c\mathbf{1}$, hence so is g. Differentiation in c proves both row-sum identities. These formulas remain valid at $\beta = 0$, when g is constant. □

Theorem D.2 (Corrective feedback: both dynamics and arbitrary initial values)**.** *Suppose the covariances in (D.3) are nonpositive for every $v \in K$ and every $s \neq r$. There is a unique SVE v^*. The normalized equilibrium v^* and the joint equilibrium $(v^*, \Lambda(v^*))$ are globally asymptotically stable, for arbitrary finite initial valuations and strictly positive initial frequencies. For two normalized solutions initialized in K,*

$$\|v(t) - u(t)\|_\infty \leq e^{-t}\|v(0) - u(0)\|_\infty. \tag{D.4}$$

Under the maintained stochastic-approximation assumptions, the harmonic visit-count recursion converges almost surely to $(v^*, \Lambda(v^*))$.

Proof. Existence and contraction on K. The continuous map g takes all of $\mathbb{R}^{\mathcal{S}}$ into K, so Brouwer's theorem on the possibly lower-dimensional box K gives a fixed point. For $v, u \in K$, put

$$A(v, u) = \int_0^1 DF(u + \theta(v - u))\, d\theta.$$

Lemma D.1 makes A Metzler with row sums -1. If $z = v - u$ and $|z_i| = \|z\|_\infty$, then

$$\operatorname{sgn}(z_i)(Az)_i \leq \left(A_{ii} + \sum_{j \neq i} A_{ij} \right) \|z\|_\infty = -\|z\|_\infty.$$

The upper right Dini derivative therefore satisfies $D^+\|z(t)\|_\infty \le -\|z(t)\|_\infty$ for normalized solutions in K. The box is positively invariant because $g_s \in [m_s, M_s]$. Gronwall's inequality proves (D.4), and applying this estimate to two fixed points proves uniqueness.

Bounds along trajectories. Let

$$H = \prod_s [\min\{m_s, v_s(0)\}, \max\{M_s, v_s(0)\}].$$

For either dynamics, H is positively invariant. On this compact box, set $d = \min_{s, v\in H} \Lambda_s(v) > 0$. The frequency equation gives

$$x_s(t) = e^{-t}x_s(0) + \int_0^t e^{-(t-u)}\Lambda_s(v(u))\,du \ge \min\{x_s(0), d\} > 0.$$

Consequently the joint adjustment rates $a_s(t) = \Lambda_s(v(t))/x_s(t)$ are bounded above and below by strictly positive constants $\overline{a}$ and $\underline{a}$. For normalized dynamics take both constants equal to one. These bounds also imply global existence. Coordinate comparison at the endpoints of K yields

$$\delta(t) := \operatorname{dist}_\infty(v(t), K) \le e^{-\underline{a}t}\delta(0). \tag{D.5}$$

Indeed each positive part $(v_s - M_s)_+$ and $(m_s - v_s)_+$ has Dini derivative at most $-a_s(t)$ times itself.

Convergence even when the covariance condition is imposed only on K. Let $\widehat{v}$ be the coordinatewise projection of v onto K and $e = v - v^*$. On the compact set H, g is Lipschitz. Put

$$B(\widehat{v}) = \int_0^1 Dg(v^* + \theta(\widehat{v} - v^*))\,d\theta.$$

Its off-diagonal entries are nonnegative and its row sums vanish. There is a finite constant C_H such that

$$F(v) = (B(\widehat{v}) - I)e + r(v), \qquad \|r(v)\|_\infty \le C_H\|v - \widehat{v}\|_\infty.$$

This identity follows by expanding $g(\widehat{v}) - g(v^*)$ and absorbing $g(v) - g(\widehat{v}) - B(\widehat{v})(v - \widehat{v})$ into $r(v)$. The previous active-coordinate calculation, now multiplied by the positive rate $a_i(t)$, gives

$$D^+\|e(t)\|_\infty \le -\underline{a}\|e(t)\|_\infty + \overline{a}C_H\delta(t).$$

Together with (D.5), this implies

$$\|v(t) - v^*\|_\infty \le e^{-\underline{a}t}\big(\|v(0) - v^*\|_\infty + \overline{a}C_H\delta(0)t\big) \longrightarrow 0. \tag{D.6}$$

Continuity of Λ and variation of constants in the frequency equation then give $x(t) \to \Lambda(v^*)$.

The estimates also prove stability. For initial states in a fixed small neighborhood of the equilibrium, H can be chosen as one fixed compact box and the rate bounds chosen uniformly. Since $\delta(0) \le \|v(0) - v^*\|_\infty$, the right side of (D.6) is bounded by a fixed multiple of the initial valuation error, uniformly in time. The frequency variation-of-constants estimate adds a fixed multiple of that error to the initial frequency error. This proves Lyapunov stability and, together with global attraction, global asymptotic stability. Finally the stochastic-approximation proposition established for the harmonic recursion identifies its almost-sure limit set with an internally chain transitive set of the joint flow. A globally asymptotically stable equilibrium is its only such limit set in the compact region visited by a bounded trajectory, proving the almost-sure assertion. □

Remark D.3. The box K is globally attracting and positively invariant. It need not be entered in finite time. For example, when g_s is constant, a coordinate initialized above that constant approaches it from above at every finite time. No finite-time absorption claim is needed in the preceding proof.

D.2 High singleton payoffs

Assume every singleton is supported, there are at least two classes, and the co-occurrence graph is connected. Write $q_s = p(\{s\})$, $a_s = \pi_s(\{s\})$, and $\pi_s^z(\omega) = \pi_s(\omega) + z\mathbf{1}\{\omega = \{s\}\}$. To give explicit sufficient cutoffs, set

$$P_s = \sum_{\omega \in \Omega_s^+} p(\omega), \quad \alpha_s = q_s/P_s, \quad \ell_s = \min_{\substack{\omega \in \Omega_s^+ \\ |\omega| \geq 2}} \pi_s(\omega), \quad u_s = \max_{\substack{\omega \in \Omega_s^+ \\ |\omega| \geq 2}} \pi_s(\omega),$$

and

$$z_{\mathrm{cov}} = \max_s \left\{ \frac{u_s - (1-\alpha_s)\ell_s}{\alpha_s} - a_s \right\}, \tag{D.7}$$

$$M = \max_{s, \omega \in \Omega_s^+} \pi_s(\omega), \quad a_{\min} = \min_s a_s, \quad p_C = \min_{\substack{\omega \in \Omega^+ \\ |\omega| \geq 2}} p(\omega), \quad q_{\max} = \max_s q_s,$$

$$\lambda_* = \frac{q_{\max}}{q_{\max} + p_C} < 1, \qquad z_{\mathrm{tie}} = \frac{M - a_{\min}}{1 - \lambda_*}. \tag{D.8}$$

Connectedness ensures that all competitive-payoff extrema and p_C are well defined.

Theorem D.4 (High singleton shifts)**.** *Let $\bar{z}_0 = \max\{0, z_{\mathrm{cov}}, z_{\mathrm{tie}}\}$. For each fixed $z > \bar{z}_0$, every finite sensitivity has a unique SVE and both dynamics converge globally to the corresponding equilibrium. The harmonic stochastic recursion converges almost surely under the maintained assumptions. As $\beta \to \infty$, the unique SVE and its entire menu-choice profile converge to a VE $(\sigma^\infty, v^\infty)$ with at least one supported competitive menu containing multiple valuation maximizers.*

Moreover, for every fixed $z > \bar{z}_0$, the limiting choice profile σ^∞ assigns positive probability to at least two classes in some supported competitive menu.

Proof. Uniform corrective feedback. The conditional singleton weight is $q_s/\Lambda_s(v) \geq \alpha_s$. For $z > z_{\mathrm{cov}}$, $a_s + z > \ell_s$ and therefore, uniformly over all finite v and all finite sensitivities,

$$g_s^z(v) \geq \alpha_s(a_s + z) + (1 - \alpha_s)\ell_s > u_s.$$

Singletons contribute zero to every cross derivative in (D.2); all competitive contributions are nonnegative. Theorem D.2 proves all finite-sensitivity dynamic conclusions.

No menu-wise strict VE at high shifts. Suppose a VE has a unique maximizer in every supported menu and let i have minimum valuation. It cannot be selected in a competitive menu: if it were that menu's unique maximizer, every other available valuation would be strictly smaller, contradicting minimality. Hence $v_i = a_i + z \geq a_{\min} + z$. Choose a supported competitive menu containing i, and denote its unique winner by j. If D_j is the total probability of menus won by j, then

$$v_j = b_j + \lambda_j z, \qquad b_j \leq M, \qquad \lambda_j = q_j/D_j \leq \frac{q_j}{q_j + p_C} \leq \lambda_*.$$

For $z > z_{\mathrm{tie}}$, this gives $v_i - v_j \geq (1 - \lambda_*)z + a_{\min} - M > 0$, contradicting the victory of j.

A limit of the unique SVE path. For fixed z, the graph of $v^\beta = g^z(v^\beta; \beta)$ is definable in $\mathbb{R}_{\exp}$. Uniqueness makes it the graph of a single bounded function of β, since all equilibria lie in $K(z)$. The o-minimal monotonicity theorem implies convergence of each coordinate at infinity. Each associated logit probability is also bounded and definable, so the complete choice profile converges. If s is not a maximizer in ω at v^∞, its logit probability is eventually at most $e^{-\beta\eta}$ for some $\eta > 0$. The limiting choices are thus supported on menu maximizers. Since $\Lambda_s(v^\beta) \geq q_s > 0$, passage to the limit in

the consistency equations is legitimate and proves VE consistency. The preceding step excludes menu-wise strictness of this limiting VE.

Genuine mixing of the selected limit. Suppose, to the contrary, that σ^∞ is pure in every supported menu, and write $\varphi(\omega)$ for its chosen class. Since the menu support is finite, for all sufficiently large finite $\beta > 0$,

$$\sigma_\omega^{\varphi(\omega)}(v^\beta) > \frac{1}{2} \qquad \text{for every } \omega \in \Omega^+.$$

The logit formula therefore implies $v^\beta_{\varphi(\omega)} > v^\beta_r$ for every $r \in \omega \setminus \{\varphi(\omega)\}$. Fix one such β and break any remaining valuation ties to obtain a strict total ranking implementing φ in every supported menu.

Let i be the lowest-ranked class. It is selected only in its singleton menu, so limiting VE consistency gives

$$v_i^\infty = a_i + z \geq a_{\min} + z.$$

Connectedness supplies a supported competitive menu containing i; let $j \neq i$ be its selected class. Define

$$D_j = \sum_{\omega:\varphi(\omega)=j} p(\omega), \qquad b_j = \frac{\sum_{\omega:\varphi(\omega)=j} p(\omega)\pi_j(\omega)}{D_j}.$$

Since j is selected in its singleton and in this competitive menu,

$$D_j \geq q_j + p_C, \qquad b_j \leq M, \qquad v_j^\infty = b_j + \frac{q_j}{D_j} z \leq M + \lambda_* z.$$

Consequently, $z > z_{\text{tie}}$ implies

$$v_i^\infty - v_j^\infty \geq a_{\min} - M + (1 - \lambda_*)z > 0.$$

This contradicts VE maximization in the competitive menu containing i and selected class j. Thus σ^∞ mixes in at least one supported competitive menu. □

D.2.1 Monotone menu probabilities and total indifference

In this subsection suppose every nonempty menu is supported, $q_s = q > 0$ for every class, and $p(\omega) \leq p(\omega')$ whenever $\omega \subseteq \omega'$. Let $M_0 = \max_{s,\omega\in\Omega_s^+} |\pi_s(\omega)|$.

Proposition D.5 (Total indifference at every VE)**.** *For $z > 2M_0/q^2$, every VE of the shifted tree has equal valuations across all classes. For $z > \max\{\bar{z}_0, 2M_0/q^2\}$, the unique SVE path selects one such VE whose limiting choice profile satisfies $\sigma_\omega^\infty(s) > 0$ for every supported menu ω and every $s \in \omega$. No genericity restriction is needed.*

Proof. Fix any VE (σ, v) and write

$$D_s = \sum_{\omega \ni s} p(\omega)\sigma_\omega(s), \qquad v_s = b_s + (q/D_s)z, \qquad |b_s| \leq M_0.$$

Suppose $v_i < v_j$, and let $\mathcal{M}_i = \{\omega : \sigma_\omega(i) > 0\}$. For every $\omega \in \mathcal{M}_i$, all classes in ω have value at most v_i, so $j \notin \omega$ and j is the unique maximizer in $\omega \cup \{j\}$. The injection $\omega \mapsto \omega \cup \{j\}$ has image disjoint from $\{j\}$. Consequently

$$D_j \geq q + \sum_{\omega\in\mathcal{M}_i} p(\omega \cup \{j\}) \geq q + \sum_{\omega\in\mathcal{M}_i} p(\omega) \geq q + D_i.$$

Since $D_i, D_j \le 1$, $q/D_i - q/D_j \ge q^2$. Consistency now gives $v_i - v_j \ge q^2 z - 2M_0 > 0$, a contradiction. Thus all valuations tie.

For $z > \max\{\bar{z}_0, 2M_0/q^2\}$, write σ^β for the logit profile at the unique SVE v^β. By Theorem D.4, these profiles and valuations converge to a VE $(\sigma^\infty, v^\infty)$. Suppose some supported menu ν contains a class i with $\sigma^\infty_\nu(i) = 0$. Choose $j \in \nu$ with $\sigma^\infty_\nu(j) > 0$. Their logit odds satisfy

$$R_\beta := \exp\{\beta(v^\beta_j - v^\beta_i)\} = \frac{\sigma^\beta_\nu(j)}{\sigma^\beta_\nu(i)} \longrightarrow \infty.$$

Consider any menu ω with $\sigma^\infty_\omega(i) > 0$. It excludes j, since otherwise $\sigma^\beta_\omega(j) = R_\beta \sigma^\beta_\omega(i)$ would exceed one. Adding j to this menu gives

$$\sigma^\beta_{\omega\cup\{j\}}(j) = \frac{R_\beta \sigma^\beta_\omega(i)}{1 + R_\beta \sigma^\beta_\omega(i)} \longrightarrow 1.$$

The same injection thus gives $D_j \ge q + D_i$ for the selection masses of σ^∞, and hence $q/D_i - q/D_j \ge q^2$. Since all limiting valuations are equal, payoff consistency gives $0 = v^\infty_i - v^\infty_j \ge -2M_0 + zq^2 > 0$, a contradiction. Every available class thus has positive probability in the selected limiting profile. □

D.3 Additional primitive sources of corrective feedback

High singleton payoffs are one way to ensure corrective feedback in the sense of Theorem D.2. More generally, the relevant comparison is between a class's assessment and the payoffs of the observations displaced by each rival. In this subsection, all menu sums and extrema here range over the supported set Ω^+. For distinct co-occurring classes s, r, define

$$Z_\omega(v) := \sum_{j\in\omega} \exp(\beta v_j), \qquad h_{sr}(v) := \frac{\sum_{\omega \ni s,r} p(\omega)\pi_s(\omega)/Z_\omega(v)^2}{\sum_{\omega\ni s,r} p(\omega)/Z_\omega(v)^2}.$$

Its weights are proportional to $p(\omega)\sigma^s_\omega(v)\sigma^r_\omega(v)$. Consequently,

$$\mathrm{Cov}_{w^v_s}\big(\pi_s(\omega), \sigma^r_\omega(v)\big) = \mathbb{E}_{w^v_s}[\sigma^r_\omega(v)]\big[h_{sr}(v) - g_s(v)\big]. \tag{D.9}$$

Nonpositive covariance is equivalent to $h_{sr}(v) \le g_s(v)$. Pairs that never co-occur have zero covariance.

Proposition D.6 (A primitive sufficient condition)**.** *Suppose every singleton menu is supported and every co-occurring pair s, r has a supported binary menu. Write*

$$q_s = p(\{s\}), \qquad a_s = \pi_s(\{s\}), \qquad P_s = \sum_{\omega\ni s} p(\omega),$$

and, for each co-occurring pair, let

$$p_{sr} = p(\{s,r\}), \qquad H_{sr} = \sum_{\omega \supsetneq \{s,r\}} p(\omega), \qquad U_{sr} = \max_{\substack{\omega\in\Omega^+ \\ s,r\in\omega}} \pi_s(\omega).$$

If

$$m_s + \frac{q_s}{P_s}(a_s - m_s) \ge \frac{p_{sr}\pi_s(\{s,r\}) + H_{sr}U_{sr}}{p_{sr} + H_{sr}} \quad \text{for every co-occurring ordered pair } s, r, \tag{D.10}$$

then all payoff–crowding covariances are nonpositive for every $v \in \mathbb{R}^{\mathcal{S}}$ and every finite $\beta \ge 0$.

Proof. Since $\Lambda_s(v) \le P_s$ and every payoff is at least m_s,

$$g_s(v) \ge m_s + \frac{q_s}{\Lambda_s(v)}(a_s - m_s) \ge m_s + \frac{q_s}{P_s}(a_s - m_s).$$

For every $\omega \supsetneq \{s, r\}$, $Z_\omega(v) \ge Z_{\{s,r\}}(v)$. The binary menu therefore receives at least the fraction $p_{sr}/(p_{sr} + H_{sr})$ of the weights defining h_{sr}. Since $\pi_s(\{s, r\}) \le U_{sr}$,

$$h_{sr}(v) \le \frac{p_{sr}\pi_s(\{s, r\}) + H_{sr}U_{sr}}{p_{sr} + H_{sr}}.$$

The conclusion follows from (D.9). □

This condition permits favorable experiences in larger menus, provided each rival predominantly displaces poorer observations. The following example also shows that intermediate singleton payoffs can sustain robust limiting indifference and mixing.

Corollary D.7 (Intermediate singleton payoffs)**.** *Consider three classes and all seven nonempty menus. Each singleton menu and each binary menu has probability* $10/61$*, and the ternary menu has probability* $1/61$*. Every represented class receives expected payoff* 1 *in singleton menus,* 0 *in binary menus, and* 2 *in the ternary menu. At every finite sensitivity, all ordered-pair payoff–crowding covariances are strictly negative, and the unique SVE is*

$$v^\beta = \frac{32}{61}\mathbf{1}.$$

Choices are uniform within every menu. In the relative parameter space of reduced primitives (p, π) *with this seven-menu support, there is an open neighborhood of the stated specification in which the covariance condition holds at every finite sensitivity, every VE has equal class valuations, and the unique SVE converges as* $\beta \to \infty$ *to total indifference with strictly positive choice probabilities in every menu.*

Proof. For every class and co-occurring rival,

$$g_s(v) \ge \frac{10}{31}, \qquad h_{sr}(v) \le \frac{2}{11}, \qquad g_s(v) - h_{sr}(v) \ge \frac{48}{341} > 0.$$

Proposition D.6 and Theorem D.2 give the covariance signs and uniqueness. Symmetry forces equal equilibrium valuations. At equal valuations the conditional payoff is

$$\frac{10 + 2/3}{10 + 10 + 1/3} = \frac{32}{61},$$

independently of sensitivity.

We establish robustness explicitly. For each nonempty proper subset $H \subset \{A, B, C\}$, consider the compact set $\mathcal{P}_H$ of choice profiles that choose only from H whenever the menu intersects H. Write $G_s(\sigma)$ for the conditional payoff induced by a profile σ. At the displayed primitives, every $\sigma \in \mathcal{P}_H$ satisfies, for $s \in H$ and $r \notin H$,

$$\begin{aligned} |H| = 1: &\quad G_s(\sigma) = 12/31 < 1/2 \le G_r(\sigma), \\ |H| = 2: &\quad G_s(\sigma) \le 12/21 < 1 = G_r(\sigma). \end{aligned}$$

For the first comparison, the sole member of H wins both its binary menus and the ternary menu, while each other class receives at most one binary-menu type alongside its singleton observations.

For the second, each member of H must win its binary menu against the remaining class; that class receives only singleton observations.

Singleton probabilities bound all sampling denominators away from zero. Continuity, compactness of the sets $\mathcal{P}_H$, and the finitely many possible H therefore make these strict comparisons persist throughout one open neighborhood of primitives. Shrink this neighborhood so that the strict inequalities in (D.10) also persist. A VE with unequal valuations would have a proper highest-valued group H, whose choices belong to $\mathcal{P}_H$. Its payoff consistency would contradict the displayed comparisons. Every nearby VE is therefore totally indifferent.

Fix any specification in this neighborhood. The unique SVE and its logit choice profile are bounded functions of β definable in $\mathbb{R}_{\exp}$. The monotonicity argument used in the proof of Theorem B.1 therefore gives limits as $\beta \to \infty$, without a genericity restriction. The same argument applies to

$$\lambda_s^\beta := \frac{\exp(\beta v_s^\beta)}{\sum_j \exp(\beta v_j^\beta)}.$$

For every choice profile, singleton support gives $D_s(\sigma) := \sum_{\omega \ni s} p(\omega)\sigma_\omega^s \geq q_s > 0$. Continuity of G_s consequently implies

$$v_s^\infty = \lim_{\beta \to \infty} G_s(\sigma^\beta) = G_s(\sigma^\infty).$$

Let $H = \{s : \lambda_s^\infty > 0\}$. If H were proper, the limiting choice profile would belong to $\mathcal{P}_H$, implying $v_s^\infty < v_r^\infty$ for $s \in H$, $r \notin H$. But $\lambda_s^\beta / \lambda_r^\beta \to \infty$ implies $v_s^\beta > v_r^\beta$ eventually, and hence $v_s^\infty \geq v_r^\infty$, a contradiction. Thus every $\lambda_s^\infty > 0$, and

$$\sigma_\omega^{\infty,s} = \frac{\lambda_s^\infty}{\sum_{j \in \omega} \lambda_j^\infty} > 0 \qquad (s \in \omega).$$

This proves full limiting mixing. □

A different construction dispenses with singleton menus altogether. Here, broader assortments generate poorer focal-class experiences, while balanced availability ensures that each rival disproportionately displaces those experiences.

Proposition D.8 (Corrective feedback without singleton menus)**.** *Suppose there is one focal class s and $m \geq 2$ benchmark classes, each yielding the same constant expected payoff c. There are no singleton menus. Every menu containing s has the form $\{s\} \cup B$, where B is a nonempty subset of benchmark classes. Conditional on $k = |B|$, these subsets are equally likely, and*

$$\pi_s(\{s\} \cup B) = h(k),$$

where h is nonincreasing on the supported sizes. Then all payoff–crowding covariances are nonpositive on K at every finite sensitivity. No ordering between c and the values of h is required.

Proof. On K, every benchmark valuation equals c. Put $a = \exp(\beta(v_s - c))$. In a menu with k benchmark classes, $\sigma^s = a/(a+k)$, so conditioning on selection of s preserves uniformity among subsets of a given size. For any benchmark class r,

$$\mathbb{E}_{w_s^v}[\sigma_\omega^r(v) \mid k] = \frac{k}{m(a+k)}.$$

Writing ν_s^v for the distribution of k conditional on choosing s, we obtain

$$\mathrm{Cov}_{w_s^v}(\pi_s, \sigma^r) = \mathrm{Cov}_{\nu_s^v}\left(h(k), \frac{k}{m(a+k)}\right) \leq 0,$$

because the first argument is nonincreasing and the second strictly increasing. Benchmark rows have zero covariance because their payoffs are constant. □

For example, support AB, AC, ABC, with $p(AB) = p(AC) > 0$ and $p(ABC) > 0$. Let A yield payoff 2 in the binary menus and 0 in the ternary menu, while B, C always yield 1. The condition holds on $K = [0, 2] \times \{1\} \times \{1\}$ at every finite sensitivity. Theorem D.2 therefore applies even from initial valuations outside K.

D.4 Low singleton payoffs and strict equilibria

For a pure selector φ on the supported menus, define

$$D_s^\varphi = \sum_{\omega:\varphi(\omega)=s} p(\omega), \qquad b_s^\varphi = \frac{\sum_{\omega:\varphi(\omega)=s} p(\omega)\pi_s(\omega)}{D_s^\varphi}, \qquad \lambda_s^\varphi = q_s/D_s^\varphi.$$

All denominators are positive because a selector must choose s in $\{s\}$. Set

$$v_s^\varphi(z) = b_s^\varphi + \lambda_s^\varphi z. \tag{D.11}$$

These are the unique valuations consistent with that selector. If they strictly order the classes in the order used by φ, the resulting pair is a strict VE.

Lemma D.9 (Priority construction)**.** *There is a generic set of fixed-support decision trees with singleton support for which some strict VE exists at every sufficiently negative shift. If every binary menu is supported, then, without any genericity restriction, for every designated class b there is a strict VE at every sufficiently negative shift with b uniquely lowest.*

Proof. The construction. Initially every supported menu is unassigned. At a priority stage, let R be the remaining candidate classes, and for $i \in R$ let

$$T_i(R) = \sum_{\substack{\omega \text{ unassigned} \\ i \in \omega}} p(\omega), \qquad \ell_i(R) = q_i/T_i(R).$$

Choose a class minimizing $\ell_i(R)$, assign to it every still unassigned menu containing it, and remove that class from R. The assigned menus are exactly those in which it is highest in the resulting priority order. Its final singleton share is its current $\ell_i(R)$. For a later class, removing menus can only increase its eventual share.

Generic existence. Run the algorithm with all classes as candidates. Impose the following generic conditions. For each possible remaining set R, whenever distinct $i, j \in R$ both have positive residual competitive mass $T_i(R) - q_i$ and $T_j(R) - q_j$, their residual shares differ; also all singleton baseline payoffs a_i differ. There are finitely many remaining sets. Each excluded equality of shares is the nonzero polynomial

$$q_i[T_j(R) - q_j] - q_j[T_i(R) - q_i] = 0.$$

It is nonzero because the competitive masses do not involve any singleton probabilities and are both nonzero linear forms in competitive-menu probabilities. Varying one singleton weight on the positive cone makes the equality fail, and homogeneity shows that its restriction to the probability simplex is not identically zero. The excluded payoff equalities are likewise proper hyperplanes. Thus these conditions hold on an open dense set of full measure in the fixed-support parameter space.

While any competitive menu remains, the minimizing share is less than one and is strictly smaller than every other remaining share. Hence it is strictly smaller than the final share of every later class.

Once no competitive menu remains, all remaining classes are selected only in their singleton menus. Order these last classes by decreasing a_i. Their consistent valuations are $a_i + z$ and remain strictly ordered. For an earlier class and any later class the strict singleton-share gap makes the earlier valuation larger for all sufficiently negative z. There are only finitely many comparisons. Taking the minimum of their finite cutoffs gives one strict ordering for every sufficiently negative shift, and (D.11) supplies its VE.

A designated bottom class, without genericity. Fix b, run the same algorithm with candidates $R = \mathcal{S} \setminus \{b\}$, and leave b in every original menu. In particular $\{b, i\}$ remains an unassigned competitive menu until i is chosen, so every candidate's final share is strictly below one. At a stage when i is chosen and j is a later candidate, the unassigned binary menu $\{i, j\}$ is assigned to i. This removes strictly positive mass from the denominator for j, while its singleton numerator remains $q_j > 0$. Therefore

$$\lambda_j^{\varphi} > \ell_j(R) \geq \ell_i(R) = \lambda_i^{\varphi}.$$

The first inequality holds even when the current minimizers tie. It proves strict increase of final singleton shares along the entire selected order. Finally assign the only remaining menu, $\{b\}$, to b, whose final share is one. We have constructed

$$\lambda_{i_1}^{\varphi} < \cdots < \lambda_{i_{n-1}}^{\varphi} < \lambda_b^{\varphi} = 1.$$

For every sufficiently negative z, the affine valuations in (D.11) strictly decrease in this priority order. Thus they implement φ, and b is uniquely lowest. For $n = 2$ the argument consists simply of the binary menu $\{b, i_1\}$ and the two singleton menus. □

Lemma D.10 (Refinement and stability of covered strict equilibria)**.** *Suppose every singleton is supported. Every menu-wise strict VE v^* is approached, for every sufficiently large β, by a locally unique SVE v^β. This SVE is locally exponentially stable under the normalized dynamics, as is the corresponding joint equilibrium under* (D.1)*.*

Proof. Choose a closed ball B about v^* in the ambient valuation space, small enough that each menu retains its unique winner and all winning margins are at least $\eta > 0$. It is unnecessary, and can be impossible, to choose this ball inside K. Uniformly on B, losing logit probabilities are $O(e^{-\beta\eta})$ and every logit derivative is $O(\beta e^{-\beta\eta})$. The denominators of g_s are at least q_s. Thus

$$\sup_{v \in B} \|g(v;\beta) - v^*\|_\infty = O(e^{-\beta\eta}), \qquad \sup_{v \in B} \|Dg(v;\beta)\|_\infty = O(\beta e^{-\beta\eta}).$$

Here the first limit is exactly VE consistency for the fixed pure selector, with its menu weights proportional to p. For large β, g maps B into itself and is a strict contraction. Its unique fixed point v^β converges to v^*. The normalized Jacobian is $Dg(v^\beta;\beta) - I$, whose eigenvalues approach -1. At a joint equilibrium the derivative of the valuation equation with respect to x vanishes, and its derivative with respect to v is the same normalized Jacobian because $x_s = \Lambda_s(v^\beta)$. The joint Jacobian is block triangular:

$$\begin{pmatrix} Dg(v^\beta;\beta) - I & 0 \\ D\Lambda(v^\beta;\beta) & -I \end{pmatrix}.$$

On the frequency-simplex tangent space it adds $n - 1$ eigenvalues -1. All eigenvalues have strictly negative real parts for large β, proving local exponential stability of both systems. □

Theorem D.11 (Low shifts and multiplicity of strict VE)**.** *Fix a tree with singleton support and connected co-occurrence graph. For a generic such tree, a strict VE exists for every sufficiently negative shift. If every binary menu is supported, then for every sufficiently negative shift there are*

at least $n = |\mathcal{S}|$ distinct strict VE, one with each class uniquely lowest, with no genericity restriction. For each fixed such shift, every strict VE has a unique nearby SVE for every sufficiently large sensitivity, and all these nearby equilibria are locally exponentially stable under both dynamics.

Proof. The first two conclusions are Lemma D.9, taking the minimum of its finitely many cutoffs over designated bottom classes. The constructed valuations are distinct because their uniquely lowest classes differ. Lemma D.10 gives refinement and stability. There are only finitely many strict VE at a fixed shift: each is induced by one of the finitely many pure menu selectors and has the unique consistent valuation (D.11). Choose disjoint sufficiently small neighborhoods around these finitely many distinct vectors, and take the maximum of their finite sensitivity thresholds. This supplies one $\widehat{\beta}(z)$ for all strict VE, as stated in the main text. □

Proposition D.12 (Every ranking under monotone menu probabilities)**.** *Under the menu-probability restrictions of Proposition D.5, every strict total order is represented by a strict VE whenever*

$$z < \underline{z}_1 := -\frac{2M_0 + 1}{q^2}.$$

For each fixed such z, all sufficiently large sensitivities admit at least $n!$ distinct SVE locally exponentially stable under both dynamics.

Proof. Fix an order $s_1 \prec \cdots \prec s_n$, with s_1 lowest, and let φ choose the highest available class. Let $\mathcal{M}_r$ be the menus won by s_r and $D_r = \sum_{\omega \in \mathcal{M}_r} p(\omega)$. For $r < n$, the injection $\omega \mapsto \omega \cup \{s_{r+1}\}$ takes $\mathcal{M}_r$ into $\mathcal{M}_{r+1}$, preserves or raises each menu's probability, and misses the singleton $\{s_{r+1}\}$. Thus $D_{r+1} \geq D_r + q$ and

$$\frac{q}{D_r} - \frac{q}{D_{r+1}} \geq q^2.$$

The affine consistent valuations obey, uniformly over all orders,

$$v^{\varphi}_{s_{r+1}}(z) - v^{\varphi}_{s_r}(z) \geq q^2|z| - 2M_0 > 1.$$

They therefore reproduce the proposed strict order and define a strict VE. Distinct orders give distinct valuation vectors. Lemma D.10 and the maximum of the finitely many sensitivity thresholds prove the claim about $n!$ stable SVE. □

D.5 Robustness to nearly equal singleton probabilities

The equal-singleton benchmark in Propositions D.5 and D.12 can be relaxed while retaining both conclusions. Maintain $n = |\mathcal{S}| \geq 2$, positive probability for every nonempty menu, and inclusion monotonicity: $p(\omega) \leq p(\omega')$ whenever $\omega \subseteq \omega'$. Allow unequal singleton probabilities and define

$$q_s = p(\{s\}), \quad q_- = \min_s q_s, \quad q_+ = \max_s q_s, \quad c = q_-^2 - (q_+ - q_-), \quad M_0 = \max_{s, \omega \ni s} |\pi_s(\omega)|. \tag{D.12}$$

The payoffs entering M_0 are the unshifted payoffs; as before, $\pi^z_s(\omega) = \pi_s(\omega) + z\mathbf{1}\{\omega = \{s\}\}$.

Proposition D.13 (Robustness of total indifference and every ranking)**.** *Suppose $c > 0$.*

(i) If $z > 2M_0/c$, every VE of the shifted tree has equal valuations across all classes. If also $z > \bar{z}_0$, where $\bar{z}_0$ is the threshold in Theorem D.4 for the given tree, all conclusions of that theorem hold and its selected limiting VE has equal valuations and assigns positive probability to every available class in every supported menu.

(ii) If $z < -2M_0/c$, every strict total order on the classes is represented by a strict VE. For each fixed such z, there is a finite $\widehat{\beta}(z)$ such that every $\beta \geq \widehat{\beta}(z)$ admits at least $n!$ distinct SVE, one near each of these strict VE, locally exponentially stable under both the normalized and joint harmonic visit-count dynamics.

Neither conclusion requires genericity.

Proof. A uniform gap in singleton shares. For any menu-choice profile σ, write

$$D_s = \sum_{\omega \ni s} p(\omega)\sigma_\omega(s), \qquad b_s = \frac{\sum_{\omega \ni s} p(\omega)\sigma_\omega(s)\pi_s(\omega)}{D_s}, \qquad \lambda_s = \frac{q_s}{D_s}.$$

Singleton support gives $D_s \geq q_s > 0$, and $|b_s| \leq M_0$. The unique valuations consistent with this profile are $v_s = b_s + \lambda_s z$.

Suppose that, whenever i is chosen with positive probability in ω, the menu excludes j and j is chosen with probability one in $\omega \cup \{j\}$. Let $\mathcal{M}_i = \{\omega : \sigma_\omega(i) > 0\}$. The map $\omega \mapsto \omega \cup \{j\}$ is injective on $\mathcal{M}_i$, and its image excludes $\{j\}$ because every image menu contains i. Inclusion monotonicity gives

$$D_j \geq q_j + \sum_{\omega \in \mathcal{M}_i} p(\omega \cup \{j\}) \geq q_j + \sum_{\omega \in \mathcal{M}_i} p(\omega) \geq q_j + D_i.$$

Since $D_i \leq 1$,

$$q_i D_j - q_j D_i \geq q_i q_j + (q_i - q_j)D_i \geq q_-^2 - (q_+ - q_-) = c.$$

Dividing by $0 < D_i D_j \leq 1$ yields the uniform bound

$$\lambda_i - \lambda_j = \frac{q_i D_j - q_j D_i}{D_i D_j} \geq c > 0. \tag{D.13}$$

Positive shifts. Fix an arbitrary VE (σ, v) and suppose $v_i < v_j$. Every menu in which i is chosen with positive probability contains only classes with valuations at most v_i. It therefore excludes j, which is the unique maximizer after being added. The preceding argument applies irrespective of how ties within the original menu are resolved. Consistency and (D.13) imply

$$v_i - v_j = (b_i - b_j) + z(\lambda_i - \lambda_j) \geq -2M_0 + zc > 0,$$

contradicting $v_i < v_j$. Thus every VE has equal valuations. Full menu support implies the singleton support and connectedness required by Theorem D.4. For $z > \max\{\bar{z}_0, 2M_0/c\}$, that theorem supplies uniqueness and global convergence at every finite sensitivity, almost-sure convergence of harmonic learning, and convergence of the SVE and choice profiles as $\beta \to \infty$. The limiting VE has equal valuations by the preceding argument. If an available class i had zero limiting probability in some menu, the logit-odds argument in the proof of Proposition D.5 would give a class j that wins with probability one when added to any menu in which i has positive limiting probability. Each such menu excludes j. Thus (D.13) applies to the limiting profile and gives $\lambda_i - \lambda_j \geq c$. Payoff consistency would then imply $0 = v_i^\infty - v_j^\infty \geq -2M_0 + zc > 0$. This proves full support of the selected limiting profile.

Negative shifts. Fix any proposed order $s_1 \prec \cdots \prec s_n$, with s_1 lowest, and let φ choose the highest available class in each menu. Compute D_s^φ, b_s^φ, and λ_s^φ from this fixed selector, before assigning valuations. For each adjacent pair $i = s_r$, $j = s_{r+1}$, every menu won by i excludes j and is won by j after adjoining it. Hence $\lambda_{s_r}^\varphi - \lambda_{s_{r+1}}^\varphi \geq c$. The consistent valuations $v_s^\varphi(z) = b_s^\varphi + \lambda_s^\varphi z$ satisfy

$$v_{s_{r+1}}^\varphi(z) - v_{s_r}^\varphi(z) \geq -2M_0 + |z|c > 0.$$

They reproduce the proposed order, so φ is optimal in every menu and defines a strict VE. Distinct orders yield distinct valuation vectors. For each fixed shifted tree, choose disjoint sufficiently small neighborhoods of these $n!$ vectors. Lemma D.10 gives a locally unique, locally exponentially stable SVE in each neighborhood at every sufficiently large sensitivity. Taking the maximum of the finitely many sensitivity thresholds proves (ii). □

The condition $c > 0$ is sufficient, not necessary, and is satisfied near every equal-singleton benchmark. Explicitly, fix such a benchmark with common singleton probability $q > 0$ and fixed unshifted payoffs. For any inclusion-monotone probability distribution with

$$\varepsilon := \max_s |q_s - q| < q^2/8,$$

we have

$$c \geq (q - \varepsilon)^2 - 2\varepsilon > q^2/2,$$

where $q \leq 1$. Thus $z > 4M_0/q^2$ suffices uniformly for total indifference at every VE, and $z < -4M_0/q^2$ suffices uniformly for every ranking to be a strict VE. The same bounds apply to payoff perturbations sharing the bound M_0. These uniform statements concern the VE conclusions: the additional global-convergence threshold $\bar{z}_0$ and the sensitivity threshold $\widehat{\beta}(z)$ are taken for each fixed perturbed environment.

E The three-menu cycling construction

E.1 Primitives and statements

There are three similarity classes A, B, C and three menus, sampled independently across periods with probabilities 1/3. Their class-menu expected payoffs are

menu	A	B	C
AB	0	1	–
BC	–	-1	$3/2$
C	–	–	$-3/4$

These primitives are implemented by five objective alternatives: one in class A, two in B, and two in C. Each alternative occurs in its indicated menu and has a stationary reward law with the displayed mean. Throughout this section, rewards satisfy the maintained independence and uniform square-integrability assumptions. Thus menu-dependent class feedback does not require menu-dependent objective reward laws. The class co-occurrence graph is the path $A - B - C$.

Let $\ell_\beta(z) = (1 + e^{-\beta z})^{-1}$, where $0 \leq \beta < \infty$. In a binary menu, classes are chosen by logit; the singleton is chosen with probability one. With positive prior counts κ_s and arbitrary finite initial valuations, the learning rule is

$$v_{n+1,s} = v_{n,s} + \frac{\mathbf{1}\{S_{n+1} = s\}}{\kappa_s + N_{n,s} + 1}\big(r_{n+1} - v_{n,s}\big). \tag{E.1}$$

Here $N_{n,s}$ counts visits through period n. The leading coefficient in the harmonic gain is exactly one, and the benchmark update uses only the observed reward.

Theorem E.1 (Exact construction)**.** *For the primitives in (E.1):*

(i) At every finite $\beta \geq 0$, there is exactly one SVE. Its valuations are $v^* = (0, 0, 0)$ and its class frequencies are $x^* = (1/6, 1/3, 1/2)$.

(ii) The joint visit-count mean field has an invariant, exponentially attracting affine plane. On this plane the equilibrium is a hyperbolic source for every $\beta > 8$. In the full five-dimensional state space it has two unstable eigenvalues and three stable eigenvalues equal to -1.

(iii) For every $\beta > 8$, every nonstationary trajectory on the plane has a periodic orbit as its omega-limit set. Every full-system trajectory outside the equilibrium's global stable manifold also has one periodic orbit as its omega-limit set.

(iv) A supercritical Hopf bifurcation occurs at $\beta = 8$. For every $\beta > 8$ sufficiently close to 8, there is a hyperbolic attracting periodic orbit, also attracting in the full joint system.

(v) For every fixed $\beta > 8$, the stochastic process (E.1), from every finite initial valuation vector and positive prior counts, almost surely has one nonconstant mean-field periodic orbit as its limit set in joint sufficient statistics. Valuations and menu-choice probabilities do not converge. For the attracting Hopf cycle in (iv), the probability that it is the limit set is positive.

The periodic orbit in (iii) or (v) may depend on the initial condition or sample path. Neither uniqueness nor hyperbolicity of periodic orbits is asserted for every $\beta > 8$. Set convergence to a periodic orbit is different from convergence to a fixed phase, or periodicity in calendar time.

E.2 Exact sample-mean coordinates and reduction

Set $K = \sum_s \kappa_s$, $T_n = K + n$, and

$$x_{n,s} = \frac{\kappa_s + N_{n,s}}{T_n}, \qquad y_{n,s} = x_{n,s} v_{n,s}.$$

Then $\sum_s x_{n,s} = 1$. With $H_{n+1} = (e_{S_{n+1}}, r_{n+1} e_{S_{n+1}})$ and $z_n = (x_n, y_n)$, the exact recursion is

$$z_{n+1} = z_n + \frac{1}{T_n + 1}\big[H_{n+1} - z_n\big]. \tag{E.2}$$

Its mean field is

$$\dot{x} = \Lambda(v) - x, \qquad \dot{y} = R(v) - y, \qquad v_s = y_s / x_s. \tag{E.3}$$

Write $a(v) = \ell_\beta(v_A - v_B)$ and $c(v) = \ell_\beta(v_C - v_B)$. For (E.1),

$$\Lambda(v) = \tfrac{1}{3}(a, 2 - a - c, 1 + c), \qquad R(v) = \tfrac{1}{3}(0, c - a, \tfrac{3}{2}c - \tfrac{3}{4}).$$

The three affine discrepancies

$$e_A = y_A, \quad e_B = y_B - x_C + x_A + \tfrac{1}{3}, \quad e_C = y_C - \tfrac{3}{2}x_C + \tfrac{3}{4} \tag{E.4}$$

satisfy $\dot{e}_s = -e_s$. Their common zero set is an invariant attracting plane. With $q = 3x_A$ and $r = 3x_C - 1$, this plane is

$$\begin{aligned} x &= \tfrac{1}{3}(q, 2 - q - r, 1 + r), \\ y &= \big(0, (r - q)/3, (r - 1/2)/2\big). \end{aligned} \tag{E.5}$$

On its relevant domain $D = (0, 1)^2$, valuations are

$$v_A = 0, \qquad B(q, r) := v_B = \frac{r - q}{2 - q - r}, \qquad C(r) := v_C = \frac{3(r - 1/2)}{2(1 + r)}, \tag{E.6}$$

and the planar system is

$$\dot{q} = \ell_\beta(-B(q,r)) - q, \qquad \dot{r} = \ell_\beta(C(r) - B(q,r)) - r. \tag{E.7}$$

The variables q, r are filtered menu-choice frequencies, not the current logit probabilities along a nonequilibrium trajectory. At equilibrium the two coincide.

For arbitrary initial conditions of the ODE, each v_s stays in the interval spanned by its initial value and its class's expected payoffs: indeed

$$\dot{v}_s = \frac{\Lambda_s(v)}{x_s}\big(g_s(v) - v_s\big).$$

Hence valuations stay bounded. At finite β, binary-menu choice probabilities are then bounded away from zero and one. The linear filter equations for x, q, r imply that x_s is eventually bounded away from zero and (q, r) eventually enters a compact subset of D. On that subset the full projected system differs from (E.7) by $O(e^{-t})$.

On $D = (0,1)^2$, $-1 \le B \le 1$ and $-3/4 \le C \le 3/8$. Thus both logit targets are bounded strictly between zero and one uniformly on D. A sufficiently small $\delta > 0$ makes $[\delta, 1-\delta]^2$ strictly inward pointing and absorbing. The denominator that vanishes at $(1,1)$ never affects recurrent dynamics.

E.3 Uniqueness at every finite sensitivity

We prove a slightly stronger fact: replace the competitive C payoff by h and its singleton payoff by $-h/2$. The equilibrium remains unique for $0 < h \le 2$.

At $\beta = 0$, both probabilities are $1/2$. At $\beta > 0$, equilibrium requires

$$\operatorname{logit} q = \beta \frac{q - r}{2 - q - r}, \tag{E.8}$$

$$\operatorname{logit} r - \operatorname{logit} q = \beta h \frac{r - 1/2}{1 + r}. \tag{E.9}$$

If either probability is $1/2$, both are. Otherwise monotonicity of logit implies that every noncentral equilibrium has either $q > 1/2 > r$ or $r > 1/2 > q$.

Let $L(t) = \log[(1/2 + t)/(1/2 - t)]$, $0 < t < 1/2$. This function is positive, increasing and strictly convex, with $L(0) = 0$. If $q = 1/2 + a$, $r = 1/2 - b$, eliminating β gives

$$h = \left(1 + \frac{L(b)}{L(a)}\right) \frac{a+b}{b} \frac{3/2 - b}{1 - a + b}. \tag{E.10}$$

For $a \ge b$, the second factor is at least two and the last exceeds one, so $h > 2$. For $a \le b$, convexity gives $L(b)/L(a) \ge b/a$; the first two factors have product at least $(a+b)^2/(ab) \ge 4$, while the last exceeds $2/3$. Hence $h > 8/3$.

If $q = 1/2 - a$, $r = 1/2 + b$, the corresponding identity is

$$h = \left(1 + \frac{L(b)}{L(a)}\right) \frac{a+b}{b} \frac{3/2 + b}{1 + a - b}. \tag{E.11}$$

For $a \ge b$,

$$(a+b)(3/2 + b) - 3b(1 + a - b) = (a - b)(3/2 - 2b) + 2b^2 > 0,$$

so $h > 3$. For $a \le b$, the first two factors again have product at least four and the last exceeds $3/2$, giving $h > 6$. All cases contradict $h \le 2$. Therefore $q = r = 1/2$ is the unique equilibrium. Every joint equilibrium lies on (E.5), proving Theorem E.1(i).

E.4 Instability and the Hopf calculation

For $h = 3/2$, the planar Jacobian at $(1/2, 1/2)$ is

$$J_\beta = \begin{pmatrix} \beta/4 - 1 & -\beta/4 \\ \beta/4 & -1 \end{pmatrix}, \qquad \lambda_\pm = -1 + \beta/8 \ \pm i\sqrt{3}\,\beta/8.$$

The three discrepancy coordinates add three eigenvalues -1. The equilibrium is a planar source for $\beta > 8$, but a saddle in the full five-dimensional system, with a three-dimensional stable manifold.

At $\beta = 8$, use

$$X = q - 1/2, \qquad Y = \frac{2(r - 1/2) - (q - 1/2)}{\sqrt{3}}.$$

The linear part is $\dot{X} = -\sqrt{3}Y$, $\dot{Y} = \sqrt{3}X$. Let f, g denote the nonlinear terms. All derivatives below are evaluated at the origin. Direct differentiation yields

$$\begin{aligned} (f_{XX}, f_{XY}, f_{YY}) &= (3, -\sqrt{3}, -3), \\ (g_{XX}, g_{XY}, g_{YY}) &= (5\sqrt{3}/9, -7/3, -7\sqrt{3}/3), \\ (f_{XXX}, f_{XYY}, g_{XXY}, g_{YYY}) &= (11/2, -63/2, -49/6, -49/2). \end{aligned}$$

In this coordinate and time normalization the cubic radial coefficient is

$$\begin{aligned} a_1 &= \frac{f_{XXX} + f_{XYY} + g_{XXY} + g_{YYY}}{16} \\ &\quad + \frac{f_{XY}(f_{XX} + f_{YY}) - g_{XY}(g_{XX} + g_{YY}) - f_{XX}g_{XX} + f_{YY}g_{YY}}{16\sqrt{3}} \\ &= -\frac{11}{3} + \frac{2}{27} = -\frac{97}{27} < 0. \end{aligned}$$

Since the real part crosses zero with derivative $1/8$, Hopf's theorem (Kuznetsov, 2013) gives a hyperbolic attracting cycle for $\beta > 8$ sufficiently close to 8. In discrepancy coordinates the variational system is block triangular; its three transverse Floquet multipliers are e^{-T}, where T is the cycle period. Thus the cycle attracts in the full joint system as well.

E.5 Planar recurrence and the full mean field

For $\beta > 8$, a small Lyapunov neighborhood of the unique planar equilibrium is strictly repelling. No nonstationary planar trajectory has that equilibrium in its omega-limit set. Compact trapping and Poincaré–Bendixson therefore give a periodic omega-limit orbit for every nonstationary planar trajectory.

For the stochastic conclusion, we also need to rule out a continuum of cycles. The following argument supplies this extra step.

Lemma E.6 below proves that every periodic orbit of this analytic source field is isolated; its proof includes the compactness and return-map argument needed to exclude a period annulus.

A bounded asymptotic pseudotrajectory has an internally chain transitive limit set. For the present planar field, Lemma E.6 leaves only the equilibrium singleton or a single periodic orbit. To see why no additional equilibrium-containing case remains, the source is isolated in its chain component by its strictly repelling neighborhood, and no nonconstant orbit enters it forward in time. There can be no cyclic chain of connecting orbits involving it.

The projection of every full-system trajectory is an asymptotic pseudotrajectory of (E.7), because its discrepancy forcing is $O(e^{-t})$ on an eventual compact interior domain. Thus its omega-limit

is either the equilibrium or one periodic orbit. The former occurs precisely on the global stable manifold of the full hyperbolic equilibrium. This proves (iii) without treating attraction to a plane as, by itself, a Poincaré–Bendixson theorem for a five-dimensional system.

E.6 The stochastic process

Boundedness, persistent sampling and limit sets

Let $\mathcal{F}_n$ be the history through period n and let $\mathcal{G}_{n+1}$ additionally reveal the next menu and implemented alternative a_{n+1}, before its reward. With this shift of the indices in (A.5), the maintained assumptions give

$$\mathbb{E}[r_{n+1} \mid \mathcal{G}_{n+1}] = \mu_{a_{n+1}}, \qquad \mathbb{E}[r_{n+1}^2 \mid \mathcal{G}_{n+1}] \leq M_2. \tag{E.12}$$

Proposition A.4 implies that valuations are almost surely bounded and $\liminf_n \min_s x_{n,s} > 0$. The valuation bound may depend on the sample path. At fixed finite β, it bounds every feasible menu-choice probability away from zero along that path. Applying the martingale strong law to the five menu-choice indicators shows that each has positive lower limiting frequency, while each menu's empirical frequency tends to $1/3$. Consequently $(3x_{n,A}, 3x_{n,C} - 1)$ eventually stays in a compact subset of D.

Write $U_{n+1} = H_{n+1} - \mathbb{E}[H_{n+1} \mid \mathcal{F}_n]$. This is martingale-difference noise with

$$\mathbb{E}[\|U_{n+1}\|^2 \mid \mathcal{F}_n] \leq 1 + M_2.$$

Recursion (E.2) has harmonic step size $\gamma_{n+1} = 1/(K + n + 1)$, so $\sum \gamma_n = \infty$ and $\sum \gamma_n^2 < \infty$. Proposition A.5 therefore applies, without bounded rewards or a moment of order greater than two.

For completeness, write the affine error vector in (E.4) as $e(z) = Lz + d$. The mean-field identity $\dot{e} = -e$ is equivalent to $L\mathbb{E}[H_{n+1} \mid \mathcal{F}_n] + d = 0$. Hence, with $\zeta_{n+1} = LH_{n+1} + d$,

$$e(z_n) = \frac{Ke(z_0) + \sum_{j=1}^{n} \zeta_j}{K + n}, \qquad \mathbb{E}[\zeta_{n+1} \mid \mathcal{F}_n] = 0, \qquad \mathbb{E}[\|\zeta_{n+1}\|^2 \mid \mathcal{F}_n] \leq C. \tag{E.13}$$

The martingale strong law gives $e(z_n) \to 0$ almost surely. Thus random rewards need not keep sample paths on the plane, but their distance from it vanishes. On the eventual compact interior domain, the projected drift differs from (E.7) by $O(\|e(z_n)\|)$, and the accumulated martingale noise vanishes on bounded learning-time windows. The projected interpolation is consequently an asymptotic pseudotrajectory of the planar field. Its limit set is internally chain transitive and, by the preceding section, is either the equilibrium or one periodic orbit.

Avoiding the unstable equilibrium

We verify the needed unstable-direction noise condition explicitly. Put $\bar{H}_{n+1} = \mathbb{E}[H_{n+1} \mid \mathcal{G}_{n+1}]$ and $\bar{U}_{n+1} = \mathbb{E}[U_{n+1} \mid \mathcal{G}_{n+1}]$. In coordinates (x_A, x_B, y_B, y_C), the five conditional mean outcome vectors $\bar{H}_{n+1}$, suppressing their zero y_A coordinate, are

$$(1,0,0,0),\ (0,1,1,0),\ (0,1,-1,0),\ (0,0,0,3/2),\ (0,0,0,-3/4).$$

Their affine differences span $\mathbb{R}^4$. In the five-dimensional tangent state space $\sum_s dx_s = 0$, this span is the invariant hyperplane $dy_A = 0$. The quotient normal to this hyperplane has eigenvalue -1, so the unstable generalized eigenspace lies in the hyperplane. Let P_+ project onto that two-dimensional space parallel to the stable generalized eigenspace, at the equilibrium. The projected conditional

mean noise $P_+\bar{U}_{n+1}$ spans the unstable space. All five menu-choice events have positive probability at equilibrium; by continuity there are a neighborhood of the equilibrium and $\delta > 0$ on which $\mathbb{E}[\|P_+\bar{U}_{n+1}\| \mid \mathcal{F}_n] \geq \delta$. Conditional Jensen's inequality gives, on this neighborhood,

$$\mathbb{E}[\|P_+U_{n+1}\| \mid \mathcal{F}_n] \geq \mathbb{E}[\|P_+\bar{U}_{n+1}\| \mid \mathcal{F}_n] \geq \delta > 0. \tag{E.14}$$

Thus centered reward noise cannot eliminate the unstable-direction excitation already supplied by random menus and choices.

The field is analytic locally, the noise has uniformly bounded conditional second moments, and the step size is harmonic. Together with the first-moment excitation condition (E.14), these properties verify the hypotheses of Brandière and Duflo (1996, Theorem 1), with noise coefficient equal to the step size and zero remainder. No $2+\delta$ reward moment or positive reward-noise variance is required. Thus

$$\mathbb{P}(z_n \longrightarrow z^*) = 0 \qquad (\beta > 8).$$

This argument covers every finite initial valuation vector; it does not require $v_{0,A} = 0$. Full ambient covariance may be singular, as in the deterministic special case; only excitation in the unstable directions is used.

Combining avoidance with the limit-set classification proves that almost surely the limit set is one nonconstant periodic orbit. The map from the plane to valuations is injective: $C(r)$ is strictly increasing, and at fixed $r < 1$,

$$\partial_q B(q,r) = -\frac{2(1-r)}{(2-q-r)^2} < 0.$$

Hence the valuation image of a cycle is nonconstant. The pair of logit menu-choice probabilities determines B, C at finite positive β, so those probabilities also fail to converge.

Capture by the attracting Hopf cycle

The Hopf cycle is an attractor of the full joint flow. For any point $w = (x, y)$ on the cycle, mix the five menu-choice reward laws with respective weights

$$q/3, \quad (1-q)/3, \quad (1-r)/3, \quad r/3, \quad 1/3.$$

The resulting integrable outcome vector $H = (e_S, re_S)$ has expectation w, so w lies in the closed convex hull of the supported outcome vectors. The finite-word construction in the proof of Proposition A.10 therefore makes every such point attainable at arbitrarily late dates. Specifically, approximate w by a finite convex combination of supported outcomes, approximate its weights by empirical frequencies, and require rewards to lie in small bounded neighborhoods of the selected support points. Each prescribed finite menu-choice word and these reward events jointly have positive probability under adaptive logit. Exact averaging in (E.2) then brings the process arbitrarily close to w once prior contributions are negligible. This argument does not require any reward to equal its mean.

The conditional martingale maximal inequality and trapping argument in that proof apply equally to a compact neighborhood of the hyperbolic attracting cycle contained in its basin. At a sufficiently late date, entry into a sufficiently small neighborhood therefore has positive conditional probability of subsequent capture, using only (E.12). The asymptotic-pseudotrajectory limit set then lies in the cycle and, as a nonempty invariant subset of a periodic orbit, equals the whole cycle. Thus the attracting Hopf cycle has positive capture probability from every finite prior. The natural mean-field clock is $\tau_n = \sum_{j \leq n} \gamma_j \sim \log n$.

E.7 An open set with uniform eventual uniqueness

For the same three-menu support, write

$$a = p(AB), \qquad b = p(BC), \qquad c = p(C),$$

and parameterize the five class-menu expected payoffs by

$$\pi_A(AB) = u, \quad \pi_B(AB) = v, \quad \pi_B(BC) = w, \quad \pi_C(BC) = z, \quad \pi_C(C) = t.$$

Thus the parameter tuple and its space are

$$\theta = (a, b, c, u, v, w, z, t) \in \Theta := \{(a, b, c) \in (0, 1)^3 : a + b + c = 1\} \times \mathbb{R}^5, \tag{E.15}$$

with the relative Euclidean topology. Reward laws may otherwise vary subject to the maintained assumptions. The centered construction has $\theta_0 = (1/3, 1/3, 1/3, 0, 1, -1, 3/2, -3/4)$.

Let $Q_A = \ell_\beta(v_A - v_B)$ and $Q_C = \ell_\beta(v_C - v_B)$ denote current binary choice probabilities. The mean-field targets are

$$\begin{aligned} \Lambda &= (aQ_A,\ a(1 - Q_A) + b(1 - Q_C),\ c + bQ_C), \\ R &= (auQ_A,\ av(1 - Q_A) + bw(1 - Q_C),\ ct + bzQ_C). \end{aligned}$$

Define the filtered-frequency coordinates

$$q = \frac{x_A}{a}, \qquad r = \frac{x_C - c}{b}, \qquad \dot{q} = Q_A - q, \qquad \dot{r} = Q_C - r. \tag{E.16}$$

They need not equal the current choice probabilities away from equilibrium. The affine discrepancies

$$\begin{aligned} e_A &= y_A - ux_A, \\ e_B &= y_B - v(a - x_A) - w(1 - a - x_C), \\ e_C &= y_C - tc - z(x_C - c) \end{aligned} \tag{E.17}$$

satisfy $\dot{e}_s = -e_s$ by direct substitution in (E.3). Their common zero set is therefore an invariant, exponentially attracting affine plane, parameterized by

$$\begin{aligned} x &= \big(aq,\ a(1 - q) + b(1 - r),\ c + br\big), \\ y &= \big(auq,\ av(1 - q) + bw(1 - r),\ ct + bzr\big). \end{aligned} \tag{E.18}$$

On this plane, $v_A = u$ and

$$B(q, r) = \frac{av(1 - q) + bw(1 - r)}{a(1 - q) + b(1 - r)}, \qquad C(r) = \frac{ct + bzr}{c + br}. \tag{E.19}$$

The dependence of B, C on θ is suppressed. The reduced dynamics are

$$\dot{q} = \ell_\beta(u - B(q, r)) - q, \qquad \dot{r} = \ell_\beta(C(r) - B(q, r)) - r. \tag{E.20}$$

At each fixed finite sensitivity, bounded valuations keep Q_A, Q_C bounded away from zero and one. The filters in (E.16) therefore place (q, r) eventually in a compact subset of $D = (0, 1)^2$, even if it starts outside D. The centered reduction (E.4)–(E.7) is recovered at $\theta = \theta_0$.

Theorem E.2 (Robust high-sensitivity cycling)**.** *Let θ_+ have $a = b = c = 1/3$ and $(u, v, w, z, t) = (1/100, 1, -1, 3/2, -3/4)$. There exist an open neighborhood $\mathcal{U}$ of θ_+ in Θ and a finite $\widehat{\beta}$ such that, for every $\theta \in \mathcal{U}$ and every $\beta \geq \widehat{\beta}$:*

(i) there is exactly one SVE, which is a source in the reduced plane and linearly unstable in the full visit-count mean field;

(ii) every reduced nonstationary trajectory, and every full trajectory outside the equilibrium's stable manifold, has one periodic omega-limit orbit;

(iii) under the maintained reward assumptions, including uniform square-integrability, and harmonic visit-count learning, the stochastic joint process almost surely has one nonconstant periodic orbit as its limit set, from every finite initial valuation vector and positive prior counts. Neither valuations nor menu-choice probabilities converge.

Proof. Step 1: strict inequalities. In a neighborhood of θ_+ we have

$$w < t < u < C(1) < v, \qquad k := B(0,0) < u, \tag{E.21}$$

$$B(0,r) > C(r) \quad \text{for every } r \in [0,1]. \tag{E.22}$$

Also $C(r^*) = u$ and $B(q^*, r^*) = u$ have the unique interior solution

$$r^* = \frac{c(u-t)}{b(z-u)}, \qquad q^* = 1 - \frac{b(u-w)}{a(v-u)}(1-r^*). \tag{E.23}$$

At θ_+,

$$r^* = \frac{76}{149}, \qquad q^* = \frac{7378}{14751}.$$

All inequalities are strict. In particular, at the specified probabilities and B, C payoffs,

$$B(0,r) - C(r) = \frac{(5/2)r^2 - (11/4)r + 3/2}{(2-r)(1+r)} > 0.$$

The numerator has negative discriminant and positive leading coefficient. These properties persist on the closure of a sufficiently small neighborhood.

Step 2: exclude every boundary limit. Consider any sequence of equilibria with $\beta_n \to \infty$ and primitives in that compact neighborhood, converging to $\bar{\theta}$. Pass to convergent probability subsequences $(q_n, r_n) \to (q, r)$. Away from $(1,1)$, the functions B, C extend continuously. An interior limiting probability requires its limiting valuation gap to vanish; a zero probability requires a nonpositive gap; a probability one requires a nonnegative gap.

If $q = 0$, (E.22) forces $r = 0$; then the q equation requires $u \le k$, contradicting (E.21). If $r = 0$ and $q \in (0,1)$, then $B(q,0) = u$, whereas $B(q,0) \le k < u$. At $(1,0)$, $B = w < t = C(0)$ contradicts $r = 0$. If $q = 1$ and $r < 1$, then $B = w < C(r)$, forcing $r = 1$, a contradiction. If $r = 1$ and $q < 1$, then $B = v > C(1)$, again a contradiction.

It remains to exclude the singular corner $(1,1)$. Along such a sequence the exact logits imply

$$\frac{1-r_n}{1-q_n} = \frac{1 + e^{\beta_n(u_n - B_n)}}{1 + e^{\beta_n(C_n - B_n)}} \longrightarrow 0.$$

Indeed $q_n \to 1$ gives $\beta_n(u_n - B_n) \to \infty$, while $C_n - u_n$ is bounded below by a strictly positive constant. Formula (E.19) then gives $B_n \to v$, which contradicts $u < v$ and $q_n \to 1$.

Every equilibrium limit is consequently interior and satisfies $B = u = C$. It must equal (E.23). By the sequential argument, this localization is uniform over the compact primitive neighborhood.

Step 3: a fixed-neighborhood implicit-function argument. Put $\varepsilon = 1/\beta$ and

$$H(q, r, \varepsilon; \theta) = \begin{pmatrix} u - B(q,r) - \varepsilon \operatorname{logit} q \\ C(r) - B(q,r) - \varepsilon \operatorname{logit} r \end{pmatrix}.$$

At $(q^*, r^*, 0)$,

$$\det D_{q,r} H = (-B_q) C' > 0,$$

because $v > w$, $z > t$, and q^*, r^* are interior. The parameter-dependent implicit function theorem gives one equilibrium branch in a fixed neighborhood, uniformly after shrinking the primitive neighborhood. Step 2 puts every equilibrium in that neighborhood for all sufficiently large β, with a common finite threshold. This proves uniform eventual uniqueness.

Step 4: instability. At equilibrium the planar Jacobian is $J_\beta = \beta M - I$, where

$$M = \begin{pmatrix} q(1-q)(-B_q) & q(1-q)(-B_r) \\ r(1-r)(-B_q) & r(1-r)(C' - B_r) \end{pmatrix}.$$

Its determinant is $q(1-q)r(1-r)(-B_q)C' > 0$. At the limiting point for θ_+,

$$\operatorname{tr} M = \frac{3240251}{13410000} > 0, \qquad (\operatorname{tr} M)^2 - 4 \det M = -\frac{33852981109031}{179828100000000} < 0.$$

Continuity and the uniform localization preserve strict positive real parts for the eigenvalues of M. Increasing the common threshold makes J_β an unstable focus uniformly. The full system adds three eigenvalues -1.

Step 5: periodic and stochastic limits. The finite expected payoffs again provide compact interior trapping for the ODE, and Proposition A.4 supplies almost-sure boundedness and persistent sampling for the stochastic process. The planar field is analytic on D. The preceding Poincaré–Bendixson, no-period-annulus, and chain-transitivity arguments apply unchanged. The affine discrepancies vanish almost surely by the martingale argument (E.13), applied to the general discrepancies (E.17) defining the plane (E.18). The five conditional mean outcome vectors have affine span of dimension four because $v \neq w$ and $z \neq t$. Their centered span is the invariant hyperplane $dy_A = u\, dx_A$, whose quotient has eigenvalue -1; its unstable projection therefore has rank two. Conditional Jensen's inequality again yields (E.14), so unstable-equilibrium avoidance and the stochastic limit-set argument hold under the maintained second-moment bound. Finally $C' > 0$ and $B_q < 0$ make the valuation map injective, proving nonconvergence as well. □

Why the positive perturbation is necessary for this robustness claim

Keep all base primitives but set $u = -\eta$ with $\eta > 0$ arbitrarily small. The corner $(q, r) = (0, 0)$ has $B = 0 > u$ and $B = 0 > t = -3/4$. On a sufficiently small rectangle about that corner, both logit targets are exponentially small in β. For all sufficiently large β, the target map sends that rectangle to itself and has derivative norm less than one. It therefore has a locally attracting fixed point tending to the corner. The interior equilibrium branch near $(1/2, 1/2)$ also persists for small η by the implicit-function argument, and remains unstable at sufficiently large β. Thus arbitrary neighborhoods of the centered example contain models with an additional stable equilibrium. One cannot infer the open-set theorem by perturbing the centered example without this qualification.

E.8 Corrective feedback fails, and extra classes do not remove the example

For the functions in (E.19), $B_q < 0$ and $C' > 0$ throughout the cycling neighborhood. Using the current choice probabilities $Q_A = \ell_\beta(v_A - v_B)$ and $Q_C = \ell_\beta(v_C - v_B)$, the payoff targets are $g_B(\mathbf{v}) = B(Q_A, Q_C)$ and $g_C(\mathbf{v}) = C(Q_C)$. Thus, for $\beta > 0$,

$$\partial_{v_A} g_B = B_q(Q_A, Q_C)\, \beta Q_A(1 - Q_A) < 0, \qquad \partial_{v_B} g_C = -C'(Q_C)\, \beta Q_C(1 - Q_C) < 0.$$

Thus the nonnegative cross-derivative requirement equivalent to negative payoff–crowding covariance fails throughout the cycling neighborhood. This makes precise the claim that competition removes favorable observations in the two indicated interactions.

Corollary E.3 (Any number of classes at least three)**.** *For each $n \geq 3$ there is a fixed menu support on n classes and a nonempty relatively open family of deterministic decision trees on that support for which the conclusions of the persistent-cycling theorem hold at all sufficiently large finite sensitivities.*

Proof. Add $n-3$ classes, each occurring only in its own singleton menu, with arbitrary deterministic payoffs. Assign positive probabilities to these menus and total probability $h > 0$ to the three original menus; conditional on an original menu, choose its probabilities and payoffs in the open family of Theorem E.2. This describes an open family relative to the enlarged fixed support.

Normalized dynamics for the original classes are unchanged and the additional valuations solve independent stable linear equations. For the joint dynamics let $z = \sum_{s\in\{A,B,C\}} x_s$. Then $\dot{z} = h - z$. In the normalized core sufficient statistics $\widetilde{x}_s = x_s/z$, $\widetilde{y}_s = y_s/z$, direct differentiation gives

$$\dot{\widetilde{x}}_s = \frac{h}{z}\big(\Lambda_s^{\mathrm{core}}(\widetilde{y}/\widetilde{x}) - \widetilde{x}_s\big), \qquad \dot{\widetilde{y}}_s = \frac{h}{z}\big(R_s^{\mathrm{core}}(\widetilde{y}/\widetilde{x}) - \widetilde{y}_s\big).$$

This is exactly the original joint system with a positive time change whose derivative tends to one. All extra frequencies converge to their menu probabilities and all extra valuations to their payoffs. Hence uniqueness, instability, and the periodic limit-set conclusion lift from the three-class subsystem; the additional directions are stable.

For the stochastic process, restrict attention to the successive dates at which an original menu occurs. Their conditional menus are i.i.d. with the original distribution, and the class visit-count updates are exactly the original three-class recursion. The number of these dates divided by calendar time tends almost surely to h. The core periodic limit set therefore lifts by multiplying its frequencies and sufficient rewards by h; the extra coordinates converge by the ordinary strong law. This proves the stochastic assertion as well. □

E.9 An exact three-class stability criterion

The following result applies to any decision tree with three classes. Theorem B.1(iii) shows that its hyperbolicity assumption holds for generic (p, π) at all but finitely many sensitivities.

Proposition E.4 (Three-class stability and the equilibrium index)**.** *Fix a finite sensitivity $\beta \geq 0$ and suppose $|\mathcal{S}| = 3$. Assume every SVE is hyperbolic. For each SVE v^*, define*

$$T(v^*) := -2 + \beta \sum_{s\neq r} \mathrm{Cov}_{w_s^{v^*}}\big(\pi_s(\omega), \sigma_\omega^r(v^*)\big), \qquad d(v^*) := \det\big(I_3 - Dg(v^*)\big). \tag{E.24}$$

Under either the normalized or joint asynchronous dynamics, an SVE is locally asymptotically stable if and only if $d(v^*) > 0$ and $T(v^*) < 0$. If $d(v^*) < 0$, it has exactly one unstable direction; if $d(v^*) > 0$ and $T(v^*) > 0$, it has exactly two unstable directions. The equilibrium set $\mathcal{V}(\beta)$ is finite and satisfies

$$\sum_{v^*\in\mathcal{V}(\beta)} \operatorname{sgn} d(v^*) = 1. \tag{E.25}$$

Consequently, there are no asymptotically stable SVE if and only if every positive-determinant SVE has $T(v^*) > 0$. In that case, at least one SVE has two unstable relative-valuation directions.

Proof. Fix an SVE v^* and write $L = Dg(v^*)$ and $J = L - I_3$. Lemma D.1 gives

$$L\mathbf{1} = 0, \qquad L_{ss} = \beta \sum_{r \neq s} \mathrm{Cov}_{w_s^{v^*}}\big(\pi_s(\omega), \sigma_\omega^r(v^*)\big), \qquad \operatorname{tr} L = T(v^*) + 2.$$

The common-shift direction therefore has eigenvalue -1 under J. The induced linear map $\bar{J}$ on the two-dimensional quotient $\mathbb{R}^3/\operatorname{span}\{\mathbf{1}\}$ has trace $T(v^*)$. Moreover, $\det \bar{J} = -\det J = \det(-J) = d(v^*)$. Hence

$$\det(\lambda I_3 - J) = (\lambda + 1)\big[\lambda^2 - T(v^*)\lambda + d(v^*)\big]. \tag{E.26}$$

If $d(v^*) < 0$, the two quotient eigenvalues are real and have opposite signs. If $d(v^*) > 0$, their real parts are both negative when $T(v^*) < 0$ and both positive when $T(v^*) > 0$. Hyperbolicity excludes $d(v^*) = 0$ and the case $d(v^*) > 0$, $T(v^*) = 0$. The linearization theorem proves the normalized classification, including local exponential stability in the stable case. Proposition A.7 adds two eigenvalues equal to -1 for the joint dynamics, so their classification is identical.

For the global index calculation, write $m_s = \min_{\omega \in \Omega_s^+} \pi_s(\omega)$ and $M_s = \max_{\omega \in \Omega_s^+} \pi_s(\omega)$. Every SVE lies in the compact payoff box $K = \prod_s [m_s, M_s]$, and the equilibrium set is closed because g is continuous. Hyperbolicity makes $I_3 - Dg(v^*)$ nonsingular, so the inverse function theorem makes every SVE isolated. Compactness then makes the equilibrium set finite. To compute its total index, take the open box

$$Q = \prod_s (m_s - 1, M_s + 1)$$

and any $c \in K$. Since $g(v) \in K$ for every $v \in \mathbb{R}^3$, the homotopy

$$H_t(v) = v - \big[(1-t)c + t g(v)\big], \qquad 0 \leq t \leq 1,$$

has no zero on ∂Q: its bracketed term belongs to $K \subset Q$. Homotopy invariance of Brouwer degree gives $\deg(H_1, Q, 0) = \deg(H_0, Q, 0) = 1$. Every zero is regular, so the sum of its local degrees is exactly (E.25). This argument also covers classes with constant payoffs, for which a coordinate interval of K is a singleton.

The degree identity guarantees a positive-determinant equilibrium. The local classification shows that none is asymptotically stable exactly when every positive-determinant equilibrium has $T(v^*) > 0$. At least one equilibrium must then have two unstable directions in the quotient by common shifts. These are relative-valuation directions; the common-shift direction and the additional joint frequency directions remain stable. □

E.10 The normalized cycling dynamics and the planar limit-set argument

This subsection supplies the normalized-dynamics part of the persistent cycling theorem and states the additional planar fact needed when a projection of the joint process is only an asymptotic pseudo-trajectory. The latter distinction matters: ordinary Poincaré–Bendixson alone does not classify arbitrary asymptotic pseudo-trajectories.

E.10.1 An exact planar quotient for normalized learning

Use the three-menu support $\{AB, BC, C\}$. Write the positive menu probabilities as a, b, c, with $a + b + c = 1$, and the five payoffs as

$$\pi_A(AB) = a_0, \quad \pi_B(AB) = b_+, \quad \pi_B(BC) = b_-, \quad \pi_C(BC) = c_+, \quad \pi_C(C) = c_0.$$

Thus $(a_0, b_+, b_-, c_+, c_0) = (u, v, w, z, t)$ in (E.15), and the functions below are precisely those in (E.19). These alternative payoff labels are used only within this section. Set $\ell_\beta(z) = (1 + e^{-\beta z})^{-1}$ and

$$B(q, r) = \frac{ab_+(1-q) + bb_-(1-r)}{a(1-q) + b(1-r)}, \qquad C(r) = \frac{cc_0 + bc_+ r}{c + br}.$$

The strict inequalities $b_+ > b_-$ and $c_+ > c_0$ hold throughout the open parameter neighborhood of the cycling construction. In particular, $B_q < 0$ and $C' > 0$ on $(0, 1)^2$.

Let $\xi = v_B - v_A$ and $\eta = v_C - v_A$. The current binary choice probabilities are

$$q = \ell_\beta(-\xi), \qquad r = \ell_\beta(\eta - \xi).$$

Since $g_A \equiv a_0$, the normalized system is exactly equivalent to

$$\begin{aligned} \dot{\xi} &= B\big(\ell_\beta(-\xi), \ell_\beta(\eta - \xi)\big) - a_0 - \xi, \\ \dot{\eta} &= C\big(\ell_\beta(\eta - \xi)\big) - a_0 - \eta, \\ \dot{v}_A &= a_0 - v_A. \end{aligned} \tag{E.27}$$

The first two equations are an autonomous analytic planar system on $\mathbb{R}^2$. They are not the filtered-frequency equations governing the joint invariant plane.

Proposition E.5 (Periodic limits under normalized learning)**.** *Suppose the SVE is unique and is a hyperbolic source on the invariant plane of the joint harmonic-visit dynamics. Then it is also a hyperbolic source in the normalized quotient* (E.27)*. Every normalized trajectory outside the equilibrium's stable manifold has a non-constant periodic orbit as its omega-limit set. Its global stable manifold is the affine line*

$$\{v^* + d(1, 1, 1) : d \in \mathbb{R}\}.$$

Thus these conclusions apply to the same open neighborhood and the same sufficiently large finite sensitivities as the joint cycling theorem. They also apply to the centered example $a_0 = 0$ for every $\beta > 8$.

Proof. Equilibria of the quotient are precisely SVE valuations expressed in relative coordinates: an equilibrium (ξ, η) lifts uniquely to $v = (a_0, a_0 + \xi, a_0 + \eta)$. Thus uniqueness transfers.

The joint Jacobian consists of its two planar eigenvalues and three eigenvalues -1. Proposition A.7 implies that the normalized Jacobian has these same two planar eigenvalues and one eigenvalue -1. In (E.27), the last eigenvalue belongs to the v_A equation, so the quotient equilibrium is a hyperbolic source. One can also verify this directly. Put $K(q, r) = (B(q, r) - a_0, C(r) - a_0)$ and $L = \left(\begin{smallmatrix} -1 & 0 \\ -1 & 1 \end{smallmatrix}\right)$. At an equilibrium, the planar joint and quotient Jacobians are

$$J_{\text{joint}} = \beta D_\ell L D K - I_2, \qquad J_{\text{norm}} = \beta D K D_\ell L - I_2, \quad D_\ell = \operatorname{diag}(q(1-q), r(1-r)).$$

Because $\det DK = B_q C' \neq 0$, these matrices are similar. In the centered example their eigenvalues are $-1 + \beta/8 \pm i\sqrt{3}\beta/8$.

The target $B(q, r)$ lies between b_- and b_+, and $C(r)$ lies in the finite interval with endpoints $C(0)$ and $C(1)$. Consequently a sufficiently large rectangle in (ξ, η) has strictly inward vector field and contains the forward orbit of any prescribed initial point. For example, enlarge each of the corresponding target intervals, shifted by $-a_0$, by one unit at either endpoint; that rectangle is absorbing. The quotient trajectory is therefore precompact.

A hyperbolic source admits a sufficiently small neighborhood with strictly outward vector field. More explicitly, if J is its linearization, choose a positive definite P satisfying $J^\top P + PJ = I_2$; this

is possible because $-J$ is Hurwitz. The function $V(z) = (z - z^*)^\mathsf{T} P(z - z^*)$ has positive derivative away from z^* in a small neighborhood. A nonstationary trajectory starting there exits a fixed small sublevel neighborhood in finite time and cannot re-enter it. Thus its omega-limit set does not contain the unique equilibrium. The Poincaré–Bendixson theorem now implies that the omega-limit set is a single periodic orbit.

Finally $v_A(t) = a_0 + (v_A(0) - a_0)e^{-t}$. Hence a quotient periodic orbit lifts, at the limit, to the periodic orbit $(a_0, a_0 + \xi(t), a_0 + \eta(t))$ of the normalized system. A quotient trajectory converges to its source only if it is identically that equilibrium. The corresponding full normalized trajectories are exactly $v^* + de^{-t}(1, 1, 1)$, proving the stable-manifold assertion. □

E.10.2 Why a stochastic planar limit is one orbit, not an annulus

Lemma E.6 (Analytic source fields have no period annulus)**.** *Let F be an analytic planar vector field on a simply connected open domain D, with exactly one equilibrium e, which is a hyperbolic source. Suppose the closed Jordan disk bounded by any periodic orbit under consideration is a compact subset of D. Then no such periodic orbit belongs to an open annulus foliated by periodic orbits. Every such periodic orbit is isolated. A compact internally chain transitive set in D is either $\{e\}$ or a single periodic orbit.*

Proof. Every periodic orbit encloses an equilibrium by the planar index theorem, and therefore encloses e. Distinct periodic orbits are disjoint and nested. Choose a closed disk-like source neighborhood U with strictly outward flow. A periodic orbit cannot intersect U: one starting in $U \setminus \{e\}$ exits in finite time and cannot return. Hence U lies in the interior of every periodic orbit.

Suppose a period annulus exists, fix a cycle in it, and continue the annulus maximally inward. Choose nested cycles Γ_n cofinal at its inner end. Their closed Jordan disks D_n decrease, contain U, and are contained in one compact Jordan disk in D. Put $D_\infty = \bigcap_n D_n$ and $K = \partial D_\infty$. The cycles converge in Hausdorff distance to K. Indeed, an accumulation point of the cycles lies in every D_n, but cannot lie in the interior of D_∞; conversely a boundary point of D_∞ at distance bounded away from some tail of the cycles would have a neighborhood inside every D_n, a contradiction. Compactness makes these assertions uniform. Thus K is compact, connected, invariant, and has empty interior. It is disjoint from the interior of U. It cannot meet ∂U either, because its backward orbit would immediately enter that interior, contradicting invariance. Thus it is disjoint from all equilibria.

The Hausdorff limit of these periodic orbits is internally chain transitive. To verify this directly, fix an error tolerance and a lower time bound T. Join nearby approximations of any two points of K along a sufficiently close Γ_n. Subdivide its orbit arc into segments with lengths in $[T, 2T]$, traversing additional revolutions if necessary. Approximate the segment endpoints by points of K. Uniform continuity of the flow over $[T, 2T]$ and Hausdorff convergence make every resulting jump smaller than the chosen tolerance.

For any compact equilibrium-free internally chain transitive set $L \subset D$, choose a compact neighborhood $W \subset D$ of L without equilibria and a smooth cutoff $\chi \in C_c^\infty(D)$ equal to one near W. Extend χF by zero to $\mathbb{R}^2$. Its flow is complete and leaves invariant a closed disk B containing its support in its interior. Restriction to the compact, locally path connected space B preserves the internal dynamics on L. By Benaïm (1999, Theorem 5.7(ii)), there is an asymptotic pseudotrajectory X in B with limit set L. Thus $\operatorname{dist}(X(t), L) \to 0$, and X eventually remains in W. Corollary 6.14 of the same reference makes L a periodic orbit or a cylinder of periodic orbits, also for the original field. Applying this conclusion to K, whose interior is empty, shows that K is a periodic orbit. Take a local transverse section to K. If its analytic return-map displacement is not identically zero, its

sign is fixed on each sufficiently small punctured side of the fixed point. Points on each side then converge to K in forward or backward time. Thus a sufficiently small tubular neighborhood of K cannot meet any other periodic orbit. This contradicts the accumulating cycles Γ_n. The analytic identity theorem therefore makes the return map equal to the identity on a full interval about its fixed point. Thus periodic orbits fill a neighborhood on both sides of K, extending the original annulus inward, a contradiction. Similarly, nonisolation of any periodic orbit would make its analytic return map identically the identity and produce a period annulus. This proves isolation.

For the last assertion, let $L \subset D$ be compact and internally chain transitive. If $e \in L$, time reversal preserves internal chain transitivity and makes e an attractor for the flow on L. An internally chain transitive set admits no proper attractor (Benaïm, 1999, Proposition 5.3), so $L = \{e\}$. Otherwise the localized classification above makes L a periodic orbit or a cylinder of periodic orbits. Isolation of every periodic orbit excludes the latter possibility. □

For the normalized quotient $D = \mathbb{R}^2$. For the joint invariant plane in filtered-frequency coordinates, $D = (0,1)^2$; its convexity ensures that every periodic orbit has its closed Jordan disk compactly inside D. The compact interior trapping already established for that plane supplies the precompactness required above. Hence any of the projected asymptotic pseudo-trajectories used in the joint and stochastic proofs has limit set either the equilibrium or a single periodic orbit. Unstable-equilibrium avoidance is a separate stochastic step and is not implied by this lemma.

E.11 Additional instability examples with a cyclic comparison graph

The following examples justify the auxiliary instability claim in the main-text footnote. They establish linearly unstable SVE under both mean fields; they do not require a conclusion about periodic orbits or the absence of other equilibria.

There are three classes A, B, C. In both examples the binary menus have the following deterministic class payoffs:

ω	A	B	C
$\{A,B\}$	-1	1	–
$\{B,C\}$	–	-1	1
$\{A,C\}$	1	–	-1

(E.28)

Thus each class has a payoff advantage over one class and a disadvantage against the other. Each supported class-menu payoff can be implemented by a distinct objective alternative in that class, with a deterministic reward law. The examples are consequently stationary decision problems of the maintained kind.

Proposition E.7 (Intermediate singleton payoffs)**.** *Support the three binary menus in* (E.28) *and the three singleton menus, each with probability* $1/6$*. Give every singleton the payoff* $z \in (-1, 0)$*. For every finite* β*, there is an SVE*

$$v^* = (z/2, z/2, z/2), \qquad x^* = (1/3, 1/3, 1/3).$$

Its normalized eigenvalues are

$$-1, \qquad -1 - \frac{3z\beta}{16} \pm i\frac{\sqrt{3}\beta}{8}. \tag{E.29}$$

The joint harmonic visit-count system has these eigenvalues together with two additional eigenvalues -1*. The equilibrium is therefore linearly unstable for* $\beta > -16/(3z)$*.*

Proposition E.8 (A ternary menu without singleton menus)**.** *Support the three binary menus in* (E.28) *and the ternary menu* $\{A, B, C\}$*, each with probability* $1/4$*. In the ternary menu give all three classes the same payoff* $h > 0$*. There are no singleton menus. For every finite* β*, there is an SVE*

$$v^* = (h/4, h/4, h/4), \qquad x^* = (1/3, 1/3, 1/3).$$

Its normalized eigenvalues are

$$-1, \qquad -1 + \frac{3h\beta}{64} \pm i\frac{3\sqrt{3}\beta}{16}. \tag{E.30}$$

The joint system adds two eigenvalues -1*, so this equilibrium is linearly unstable for* $\beta > 64/(3h)$*. For example,* $h = 1/2$ *gives the threshold* $128/3$ *and uses a ternary payoff strictly between the binary payoffs* -1 *and* 1*.*

Proof of Propositions E.7 and E.8. A common calculation covers both cases. Give each binary menu mass a, each singleton mass b, and the ternary menu mass c, omitting a menu type when its mass is zero. Thus $3a + 3b + c = 1$. At equal valuations each binary choice probability is $1/2$ and each ternary choice probability is $1/3$. For every class,

$$\Lambda_s = a + b + c/3 = 1/3, \qquad R_s = bz + ch/3, \qquad V := g_s = 3bz + ch.$$

Hence (V, V, V) is an SVE. Use the quotient-rule derivative identity to define

$$\kappa = 3\beta\left(\frac{aV}{4} + \frac{c(V-h)}{9}\right), \qquad \lambda = \frac{3a\beta}{4}.$$

In the coordinate order (A, B, C) the Jacobian of g is

$$Dg = \begin{pmatrix} -2\kappa & \kappa+\lambda & \kappa-\lambda \\ \kappa-\lambda & -2\kappa & \kappa+\lambda \\ \kappa+\lambda & \kappa-\lambda & -2\kappa \end{pmatrix}.$$

Indeed each shared binary menu contributes $3\beta(a/4)(V - \pi_s)$ to its cross derivative and the ternary menu contributes $3\beta(c/9)(V - h)$. Translation invariance gives the displayed diagonal entries. The eigenvalues are 0 and $-3\kappa \pm i\sqrt{3}\lambda$, so the normalized Jacobian $Dg - I$ has eigenvalues -1 and $-1 - 3\kappa \pm i\sqrt{3}\lambda$.

For $(a, b, c) = (1/6, 1/6, 0)$, $V = z/2$, $\kappa = z\beta/16$, and $\lambda = \beta/8$. For $(a, b, c) = (1/4, 0, 1/4)$, $V = h/4$, $\kappa = -h\beta/64$, and $\lambda = 3\beta/16$. These give the two claimed spectra. Finally the harmonic joint Jacobian at any equilibrium has blocks

$$\begin{pmatrix} Dg - I & 0 \\ D\Lambda & -I \end{pmatrix}.$$

Restriction to the frequency-simplex tangent space adds two copies of -1, proving both joint conclusions. □

Corollary E.9 (Robustness on each fixed menu support)**.** *Around either of the preceding specifications there is a nonempty open set of reduced payoff and menu-probability parameters, with that example's menu support fixed, and a common finite* β_0 *such that every tree in the open set has a linearly unstable SVE for every* $\beta \geq \beta_0$*. The assertion is existence of an unstable branch; it does not assert global uniqueness of SVE.*

Proof. Denote the parameters by θ and their symmetric value by θ_0. Use scaled relative scores $w = (\beta(v_A - v_C), \beta(v_B - v_C))$ and let $G_\theta(w)$ be the conditional payoff vector when the logit score vector is $(w_1, w_2, 0)$. It is smooth jointly in (w, θ) near the specified parameters, since all denominators are positive. Put

$$\Psi_\theta(w) = (G_{\theta,A}(w) - G_{\theta,C}(w), G_{\theta,B}(w) - G_{\theta,C}(w)).$$

With $\varepsilon = 1/\beta$, the relative equilibrium equation is $\Psi_\theta(w) - \varepsilon w = 0$. At $(\theta, w, \varepsilon) = (\theta_0, 0, 0)$ it holds, and the two eigenvalues of $D\Psi_{\theta_0}(0)$ have strictly positive real parts: $-3z/16 \pm i\sqrt{3}/8$ in the singleton example and $3h/64 \pm i3\sqrt{3}/16$ in the ternary example. In particular this derivative is invertible. The implicit function theorem gives a smooth local solution $w = W(\theta, \varepsilon)$ for θ in one open neighborhood and ε near zero. Define

$$v_C = G_{\theta,C}(W), \qquad v_A = v_C + \varepsilon W_1, \qquad v_B = v_C + \varepsilon W_2.$$

The relative equilibrium equation makes this a genuine SVE. By continuity, after shrinking the parameter neighborhood and choosing $\varepsilon_0 > 0$, the eigenvalues of $D\Psi_\theta(W(\theta, \varepsilon)) - \varepsilon I$ retain strictly positive real parts throughout $0 \le \varepsilon \le \varepsilon_0$. The normalized relative Jacobian is β times this matrix. Its remaining valuation eigenvalue is -1; the joint system adds two more copies of -1. Taking $\beta_0 = 1/\varepsilon_0$ proves the uniform assertion. □

F Coarse representation and the welfare cost of learning

This section establishes the claims in the main text's welfare discussion. The within-class implementation rule remains fixed. Its expected payoff is incorporated in $\pi_s(\omega)$, so changing that rule is outside the policy comparison below. Sums and extrema are over the supported set Ω^+.

F.1 The informed common-score benchmark

For $\theta \in \mathbb{R}^S$, write

$$q_s(\omega; \theta) = \frac{e^{\theta_s}}{\sum_{j \in \omega} e^{\theta_j}}, \qquad W(q) = \sum_\omega p(\omega) \sum_{s \in \omega} q_s(\omega) \pi_s(\omega).$$

The unrestricted score index θ absorbs the product βv at positive sensitivity. Let $\mathcal{L}$ be the family of these menu-wise logit policies and let $\overline{\mathcal{L}}$ be its closure in the finite product of menu simplexes. For a strict class ranking $\succ$, let $q^\succ$ choose $\max_\succ \omega$ from each menu and put $W_\succ = W(q^\succ)$.

Proposition F.1 (The informed coarse benchmark)**.** *Allowing deterministic limits of logit choice, an optimal oracle policy is attained by a strict class ranking, and its expected payoff is*

$$W^{\text{coarse}} := \max_{q \in \overline{\mathcal{L}}} W(q) = \max_\succ W_\succ = \max_\succ \sum_\omega p(\omega) \pi_{\max_\succ \omega}(\omega). \tag{F.1}$$

Every policy in $\overline{\mathcal{L}}$ is a mixture of strict-ranking policies, so allowing further randomization cannot raise this value. For generic reduced parameters (p, π) on the fixed menu support, the optimal policy on supported menus is unique and pure.

Suppose each supported menu has a unique payoff-maximizing class $s^*(\omega)$. Define the efficient-comparison graph G^{eff} on $\mathcal{S}$ by drawing an arrow $s \to r$ whenever $s = s^*(\omega)$ and $r \in \omega \setminus \{s\}$ for some supported menu. Then

$$W^{\text{coarse}} = W^{\text{menu}} := \sum_\omega p(\omega) \max_{s \in \omega} \pi_s(\omega) \quad \iff \quad G^{\text{eff}} \text{ is acyclic.} \tag{F.2}$$

Proof. Fix θ. Independently for every class draw an exponential random variable T_s of rate e^{θ_s}, and rank all classes by increasing T_s. Ties have probability zero. For every menu,

$$\Pr\{T_s = \min_{j\in\omega} T_j\} = \int_0^\infty e^{\theta_s} \exp\left(-t\sum_{j\in\omega} e^{\theta_j}\right) dt = q_s(\omega;\theta).$$

Thus a single random ranking, drawn independently of the menu, implements the entire logit policy. If $\mu_\theta(\succ)$ is its distribution, then $q(\theta) = \sum_\succ \mu_\theta(\succ) q^\succ$. There are finitely many rankings, so their convex hull is closed. It therefore contains $\overline{\mathcal{L}}$, and linearity gives $W(q) \le \max_\succ W_\succ$ throughout this closure. Conversely, fix distinct scores ordered by $\succ$ and multiply them by $t \to \infty$. Their logit policies tend to $q^\succ$, so every ranking policy belongs to $\overline{\mathcal{L}}$. This proves equality and attainment. Linearity also proves the assertion about further randomization.

For generic uniqueness, identify rankings that induce the same selector on Ω^+. Any two distinct induced selectors f, h differ on some supported menu ω_0. Their expected welfare difference is

$$\sum_{\omega\in\Omega^+} p(\omega)\big[\pi_{f(\omega)}(\omega) - \pi_{h(\omega)}(\omega)\big].$$

For each fixed p, the coefficient of $\pi_{f(\omega_0)}(\omega_0)$ is $p(\omega_0) > 0$, so equality defines a proper hyperplane in the payoff space. In the joint parameter space each tie set is relatively closed by continuity, has empty interior by varying that payoff alone, and has measure zero by Fubini's theorem. There are only finitely many pairs. On the open dense, full-measure complement, one induced selector uniquely maximizes welfare. An optimal mixture can then put weight only on rankings that induce this selector, so every optimal policy makes the same pure choice in each supported menu.

Every policy earns at most W^{menu}. If G^{eff} is acyclic, a topological ordering gives a strict ranking with $s \succ r$ for every arrow $s \to r$. It selects $s^*(\omega)$ in every supported menu and therefore attains W^{menu}. Conversely, suppose $W^{\text{coarse}} = W^{\text{menu}}$ and choose an optimal strict ranking, whose existence was just proved. Its menu-wise payoff shortfalls are nonnegative. Since every supported menu has positive probability and a unique maximizer, equality of total welfare forces the ranking to select $s^*(\omega)$ in every menu. It therefore respects every arrow. A directed cycle would require a strict ranking to rank a class above itself, a contradiction. □

For precision, the benchmark preserves both the prescribed within-class implementation and the use of a common score across menus. Allowing arbitrary class choices separately in each reduced menu instead gives $W^{\text{menu}} \ge W^{\text{coarse}}$, as defined in (F.2). In the underlying alternative model, write $f(\psi)$ for the physical menu distribution and μ_a for the expected reward of alternative a. Unrestricted informed alternative choice yields

$$W^{\text{full}} = \sum_\psi f(\psi) \max_{a\in\psi} \mu_a \;\ge\; W^{\text{menu}}.$$

Indeed, for any reduced-menu class rule its expected payoff in a physical menu is an average of the available alternatives' means, hence no larger than their maximum. The same conclusion holds with menu-dependent means after replacing μ_a by $\mu_a(\psi)$. Thus, for a policy $q \in \overline{\mathcal{L}}$, the identity

$$W^{\text{full}} - W(q) = \big(W^{\text{full}} - W^{\text{coarse}}\big) + \big(W^{\text{coarse}} - W(q)\big) \tag{F.3}$$

separates the loss imposed by the maintained representation from its additional policy shortfall. The second term at a finite sensitivity can include logit noise. Comparing high-sensitivity limits of long-run outcomes isolates losses that persist as logit noise vanishes.

F.2 Efficient scores within own-payoff intervals

Throughout this and the next subsection, each supported menu has a unique payoff-maximizing class, and G^{eff} is the graph in Proposition F.1. Define

$$\ell_s := \min_{\omega \ni s} \pi_s(\omega), \qquad u_s := \max_{\omega \ni s} \pi_s(\omega), \qquad I_s := [\ell_s, u_s], \qquad K := \prod_{s \in \mathcal{S}} I_s. \tag{F.4}$$

These intervals are well defined because every class has positive availability. Call a score vector *efficient* if $v_{s^*(\omega)} > v_r$ for every supported menu and every $r \in \omega \setminus \{s^*(\omega)\}$. Equivalently, holding these scores fixed and letting logit sensitivity increase selects the unique payoff-maximizing class in each menu with probability tending to one.

Proposition F.2 (Score feasibility)**.** *There exists an efficient score vector $v \in K$ if and only if G^{eff} is acyclic and*

$$u_s > \ell_t \qquad \text{for every nontrivial directed path } s \rightsquigarrow t. \tag{F.5}$$

Proof. An efficient vector satisfies $v_s > v_r$ along every arrow, which excludes directed cycles. Along any nontrivial directed path $s \rightsquigarrow t$, it also satisfies $u_s \geq v_s > v_t \geq \ell_t$. This proves necessity.

For sufficiency, assume the graph is acyclic and (F.5) holds. Let $D(s)$ contain s and all its descendants. For $t \in D(s)$, let $L(s,t)$ be the length of a longest directed path from s to t, with $L(s,s) = 0$. All these lengths are finite. Since there are finitely many reachable pairs and their endpoint inequalities are strict, choose $\varepsilon > 0$ such that

$$\ell_t + \varepsilon L(s,t) \leq u_s \quad \text{whenever } t \in D(s) \setminus \{s\}.$$

If there are no arrows, any positive ε suffices. Set

$$v_s := \max_{t \in D(s)} \{\ell_t + \varepsilon L(s,t)\}.$$

The term $t = s$ gives $v_s \geq \ell_s$, and every term is at most u_s, so $v \in K$. If $s \to r$, then $D(r) \subseteq D(s)$ and $L(s,t) \geq 1 + L(r,t)$ for every $t \in D(r)$. Hence $v_s \geq v_r + \varepsilon$. The vector therefore satisfies every required strict comparison. □

The criterion checks entire paths, including endpoints that never appear together. In the main-text example, the menus $\{A, C\}$ and $\{B, C\}$ have respective payoffs $(3/5, 1)$ and $(2/5, 0)$, so

$$B \longrightarrow C \longrightarrow A, \qquad I_B = \{2/5\}, \quad I_C = [0,1], \quad I_A = \{3/5\}.$$

Each individual arrow can be satisfied by scores in these intervals, but the path cannot: it requires $v_B > v_A$, while $u_B < \ell_A$. The proposition concerns feasibility of efficient score comparisons.

F.3 Welfare bounds from conflicting paths

A *common monotone binary rule* is a measurable function $Q : \mathbb{R}^2 \to [0,1]$ that is weakly increasing in its first argument, weakly decreasing in its second, and satisfies $Q(x,y) + Q(y,x) = 1$. In a binary menu $\{s, r\}$ it selects s with probability $Q(v_s, v_r)$. The same function is used across classes and menus. Logit is one such rule; continuity is not required.

Lemma F.3 (Errors along a path)**.** *For any common monotone binary rule, any integer $d \geq 2$, and any real scores $a_0, \ldots, a_d$ with $a_0 \leq a_d$,*

$$\sum_{i=1}^{d} Q(a_i, a_{i-1}) \geq 1. \tag{F.6}$$

Proof. First, for all real a, b, c,

$$Q(b, a) + Q(c, b) \geq Q(c, a).$$

If $b \geq c$, the first term alone is at least the right side. If $b \leq a$, the second term alone is at least the right side. In the remaining case $a < b < c$, both terms are at least $1/2$, because $Q(x, x) = 1/2$ and the rule is monotone. Thus their sum is at least one and again dominates the right side. Repeatedly applying this inequality to the first $d - 1$ terms gives

$$\sum_{i=1}^{d} Q(a_i, a_{i-1}) \geq Q(a_{d-1}, a_0) + Q(a_d, a_{d-1}) \geq Q(a_{d-1}, a_0) + Q(a_0, a_{d-1}) = 1,$$

where the last inequality uses $a_d \geq a_0$. □

Proposition F.4 (A path-based welfare bound)**.** *Suppose G^{eff} is acyclic and contains a directed path of distinct classes $s_0 \to s_1 \to \cdots \to s_d$, with $d \geq 2$. For each edge choose a supported menu ω_i whose unique payoff-maximizing class is s_{i-1} and which contains s_i. Put*

$$\Delta_i := \pi_{s_{i-1}}(\omega_i) - \max_{r \in \omega_i \setminus \{s_{i-1}\}} \pi_r(\omega_i) > 0, \qquad c := \min_{1 \leq i \leq d} p(\omega_i)\Delta_i > 0. \tag{F.7}$$

Every common-score logit policy with finite sensitivity $\beta \geq 0$ and $v_{s_0} \leq v_{s_d}$ satisfies

$$W(q) \leq W^{\text{coarse}} - c. \tag{F.8}$$

If all the witnessing menus ω_i are binary, the same bound holds for any common monotone binary rule on those menus, irrespective of choices in the other menus.

Proof. The menus ω_i are distinct: otherwise the same menu would have two different unique payoff maximizers. Let E_i be the probability of failing to select s_{i-1} in ω_i. Since the graph is acyclic, Proposition F.1 gives $W^{\text{coarse}} = W^{\text{menu}}$. Menu-wise welfare shortfalls are nonnegative, so

$$W^{\text{coarse}} - W(q) \geq \sum_{i=1}^{d} p(\omega_i)\Delta_i E_i \geq c \sum_{i=1}^{d} E_i.$$

Under logit, comparison with the two-class menu gives

$$E_i = 1 - \frac{e^{\beta v_{s_{i-1}}}}{\sum_{r \in \omega_i} e^{\beta v_r}} \geq \frac{e^{\beta v_{s_i}}}{e^{\beta v_{s_{i-1}}} + e^{\beta v_{s_i}}} = Q_\beta(v_{s_i}, v_{s_{i-1}}).$$

Lemma F.3 implies $\sum_i E_i \geq 1$, proving (F.8). When the witnessing menus are binary and use a common monotone rule, $E_i = Q(v_{s_i}, v_{s_{i-1}})$, and the same lemma proves the second assertion. □

For a binary witnessing menu, $p(\omega_i)\Delta_i$ is its probability-weighted payoff advantage. In the main-text example both advantages equal $(1/2)(2/5) = 1/5$, giving the ceiling $7/10 - 1/5 = 1/2$. A failed interval comparison cannot arise along a single arrow: if s is best in a menu containing t, then $u_s \geq \pi_s(\omega) > \pi_t(\omega) \geq \ell_t$. Thus any violation of (F.5) in an acyclic graph occurs along a path of length at least two, as required by Proposition F.4.

For the online version, let $\mathcal{H}_k$ contain the observed history and private randomization used before period k's menu is drawn. The current menu is independent of $\mathcal{H}_k$ and has distribution p. Scores and the choice rule (including any sensitivity) are $\mathcal{H}_k$-measurable, and the current menu, choice, and reward belong to $\mathcal{H}_{k+1}$. Assume

$$\mathbb{E}[r_k \mid \mathcal{H}_k, \omega_k, s_k] = \pi_{s_k}(\omega_k), \qquad \mathbb{E}[r_k^2 \mid \mathcal{H}_k] \leq M_2 < \infty.$$

These reward conditions hold under the maintained finite collection of stationary independent reward laws and fixed within-class rule.

Corollary F.5 (A persistent loss under endpoint calibration)**.** *Under the hypotheses of Proposition F.4, suppose*

$$u_{s_0} < \ell_{s_d}, \qquad \operatorname{dist}(v_k(s), I_s) \longrightarrow 0 \quad (s = s_0, s_d) \quad \textit{almost surely.} \tag{F.9}$$

In the preceding online setting, use either logit with any finite history-dependent sensitivities $\beta_k \geq 0$, or, when the witnessing menus are binary, any history-dependent common monotone binary rules Q_k on those menus. Then, almost surely,

$$\limsup_{k\to\infty} \mathbb{E}[r_k \mid \mathcal{H}_k] \leq W^{\text{coarse}} - c, \qquad \limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} r_k \leq W^{\text{coarse}} - c. \tag{F.10}$$

Proof. Strict separation of the endpoint intervals and their asymptotic calibration give $v_k(s_0) < v_k(s_d)$ eventually almost surely. Conditional on $\mathcal{H}_k$, the scores and choice rule are fixed before the independent menu is drawn. Proposition F.4 therefore bounds the conditional expected reward by $W^{\text{coarse}} - c$ eventually, proving the first assertion.

Let $D_k = r_k - \mathbb{E}[r_k \mid \mathcal{H}_k]$. These are martingale differences with $\mathbb{E}[D_k^2 \mid \mathcal{H}_k] \leq M_2$. The martingale $\sum_{k<K} D_k/(k+1)$ has uniformly bounded second moments because $\sum_k M_2/(k+1)^2 < \infty$, and hence converges almost surely. Kronecker's lemma gives $K^{-1}\sum_{k<K} D_k \to 0$ almost surely. The eventual conditional reward bound then gives the asserted bound on realized average rewards. □

If class-menu rewards are deterministic, endpoint calibration follows from any nonnegative own-payoff averaging rule with eventually positive denominators and vanishing initial-score shares for the two endpoints. Indeed, its endpoint score can be written $v_k(s) = \lambda_{s,k} v_0(s) + (1 - \lambda_{s,k}) z_{s,k}$ with $z_{s,k} \in I_s$ and $\lambda_{s,k} \to 0$; compactness of I_s gives the calibration in (F.9). This includes arbitrary history- and menu-dependent weights. For noisy rewards, unrestricted reward-dependent weights need not approach the interval of conditional means; endpoint calibration must instead be assumed or established. Neither intermediate-score convergence nor convergence of the choice rule is required.

For fixed menu support, strict endpoint separation and the unique menu-wise payoff maximizers persist under all sufficiently small perturbations of payoffs and positive menu probabilities. The constant c remains positive. Each perturbed problem therefore has the same type of welfare bound for learners satisfying its endpoint calibration.

Strict separation in (F.9) is material when sensitivity can grow. With equally likely menus $\{A, C\}$ and $\{B, C\}$ of payoffs $(0, 1)$ and $(0, -1)$, the graph is $B \to C \to A$, but $I_A = I_B = \{0\}$ and $I_C = [-1, 1]$. No efficient vector lies in K. Nevertheless, for $k \geq 1$ the scores $v_k(A) = -1/k$, $v_k(C) = 0$, $v_k(B) = 1/k$ approach K, while logit sensitivity $\beta_k = k^2$ selects the efficient class in each menu with probability tending to one. Thus equality of the endpoint bounds rules out exact score feasibility but does not by itself imply a persistent loss under asymptotic calibration.

F.4 Singleton shifts and the two-class illustrations

Lemma F.6 (Singleton payoffs do not change the oracle's ranking)**.** *Replace each supported singleton payoff $\pi_s(\{s\})$ by $\pi_s(\{s\}) + \delta_s$, leaving all other primitives unchanged. The welfare of every fixed policy changes by the same amount $D = \sum_s p(\{s\})\delta_s$. Hence W^{coarse} changes by D, its maximizing rankings are unchanged, and each fixed policy's gap to the benchmark is unchanged.*

Proof. Every feasible policy selects s with probability one from $\{s\}$. Its payoff change is therefore exactly D, independently of all choices on competitive menus. Taking the maximum gives the assertions. □

This invariance concerns a fixed policy. Learning may produce a different policy after the shift. In the Bordeaux–Chianti illustrations, the menus $\{B, C\}, \{B\}, \{C\}$ each have probability $1/3$. Competitive payoffs are $(2, 1)$ and both singleton payoffs equal the shift z. If q is the probability of Bordeaux in the competitive menu, then

$$W_z(q) = \frac{1 + q + 2z}{3}, \qquad W_z^{\text{coarse}} = \frac{2 + 2z}{3}, \qquad W_z^{\text{coarse}} - W_z(q) = \frac{1 - q}{3}. \tag{F.11}$$

The convergence and high-sensitivity calculations are proved in Section C. At $z = 0$, the Chianti-favoring strict branch has $q \to 0$, so its limiting loss is $1/3$. At $z = 11/3$, the unique branch has $q \to 2/3$ and both valuations tend to 3, so its limiting loss is $1/9$. To check the latter numbers directly, its limiting conditional means satisfy

$$\frac{2q + 11/3}{1 + q} = \frac{1 - q + 11/3}{2 - q},$$

whose only solution in $[0, 1]$ is $q = 2/3$; both sides then equal 3. These losses are measured after taking the long-run outcome at each fixed sensitivity and then letting sensitivity increase.

F.5 An obstruction to calibrated payoff aggregation

The example admits a useful open family. There are two menus, $\{A, C\}$ and $\{B, C\}$, of probabilities p and $1 - p$, with deterministic payoffs (a, h) and (b, ℓ), respectively, where

$$0 < p < 1, \qquad h > a > b > \ell. \tag{F.12}$$

The maintained within-class implementation has no choice to make in either menu.

Lemma F.7 (Realization by objective alternatives)**.** *The reduced tree in* (F.12) *is induced by four alternatives with menu-independent deterministic rewards. Moreover,*

$$W^{\text{full}} = W^{\text{coarse}} = ph + (1 - p)b,$$

and the ranking $B \succ C \succ A$ attains this value.

Proof. Take classes $A = \{a_0\}$, $B = \{b_0\}$, and $C = \{c_H, c_L\}$, with respective alternative rewards a, b, h, ℓ. Let the physical menus be $\{a_0, c_H\}$ and $\{b_0, c_L\}$, with probabilities p and $1 - p$. Each available class has exactly one available alternative. The stated ranking chooses c_H in the first menu and b_0 in the second, which are precisely the two menu-wise payoff maximizers. No policy can improve on their expected payoff. □

Proposition F.8 (Welfare bound under the calibrated ordering)**.** *In* (F.12)*, every common monotone binary rule Q and every common score vector satisfying $v_A \geq v_B$, regardless of v_C, satisfy*

$$W_Q(v) \leq \overline{W} := \max\{pa + (1 - p)b,\ ph + (1 - p)\ell\}, \tag{F.13}$$

$$W^{\text{coarse}} - \overline{W} = \min\{p(h - a), (1 - p)(b - \ell)\} > 0. \tag{F.14}$$

This bound is sharp within the logit family as sensitivity increases, even when each score lies in the convex hull of that class's own possible rewards.

Proof. Let u and w be the probabilities of choosing C in $\{A, C\}$ and $\{B, C\}$, respectively. Monotonicity and $v_A \geq v_B$ give

$$u = Q(v_C, v_A) \leq Q(v_C, v_B) = w.$$

Put $B_0 = pa + (1-p)b$, $X = p(h-a) > 0$, and $Y = (1-p)(b-\ell) > 0$. Welfare equals

$$W_Q(v) = B_0 + Xu - Yw \leq B_0 + (X-Y)u \leq B_0 + \max\{0, X-Y\}.$$

This last value equals $\overline{W}$, while $W^{\text{coarse}} = B_0 + X$, proving both formulas. For sharpness set $(v_A, v_B) = (a, b)$. Taking $v_C = \ell$ and $\beta \to \infty$ yields the first value in the maximum; taking $v_C = h$ yields the second. These scores belong to the respective own-payoff hulls. □

The ordering $v_A \geq v_B$ is precisely what constant-payoff calibration eventually implies. Use the online setting preceding Corollary F.5, with scores v_k and a common monotone binary rule Q_k determined from $\mathcal{H}_k$ before the current independent menu is drawn. This includes logit with arbitrary history-dependent finite sensitivities $\beta_k \geq 0$. Neither scores nor the choice rule may be selected separately for different current menus.

Theorem F.9 (The long-run aggregation bound)**.** *Suppose $v_k(A) \to a$ and $v_k(B) \to b$ almost surely in the environment (F.12). No convergence or calibration condition is imposed on $v_k(C)$, and Q_k need not converge. Then almost surely*

$$\limsup_{k\to\infty} \mathbb{E}[r_k \mid \mathcal{H}_k] \leq \overline{W}, \qquad \limsup_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} r_k \leq \overline{W}. \tag{F.15}$$

In particular, these conclusions hold for any nonnegative own-payoff aggregation rule

$$v_k(s) = \frac{a_{s,k} v_0(s) + \sum_{j<k:s_j=s} w_{s,k,j} r_j}{a_{s,k} + \sum_{j<k:s_j=s} w_{s,k,j}}, \qquad a_{s,k}, w_{s,k,j} \geq 0, \tag{F.16}$$

provided that for $s = A, B$ the denominators are eventually positive and

$$\lambda_{s,k} := \frac{a_{s,k}}{a_{s,k} + \sum_{j<k:s_j=s} w_{s,k,j}} \longrightarrow 0 \quad \textit{almost surely.} \tag{F.17}$$

The finite nonnegative weights can depend arbitrarily on the observed history, including past menus and rewards, and can be chosen with full knowledge of the primitives.

Proof. Since $a > b$, calibration implies $v_k(A) > v_k(B)$ eventually on an event of probability one. Applying Proposition F.8 conditional on $\mathcal{H}_k$ gives the eventual bound $\mathbb{E}[r_k \mid \mathcal{H}_k] \leq \overline{W}$. Independence of the current menu is what preserves its probabilities in this conditional calculation. Let $D_k = r_k - \mathbb{E}[r_k \mid \mathcal{H}_k]$. The variables D_k are bounded martingale differences. Hence the martingale series $\sum_{k\geq 0} D_k/(k+1)$ converges almost surely, since its increment variances have finite sum. Kronecker's lemma gives $K^{-1}\sum_{k<K} D_k \to 0$ almost surely. The Cesàro averages of the conditional rewards have limit superior at most $\overline{W}$; their finitely many initial terms vanish from the average. This proves (F.15).

Under (F.16), every observation of A equals a and every observation of B equals b. Consequently,

$$v_k(A) = a + \lambda_{A,k}(v_0(A) - a), \qquad v_k(B) = b + \lambda_{B,k}(v_0(B) - b)$$

whenever their denominators are positive. Condition (F.17) implies the required calibration, irrespective of how the observation weights were chosen. □

For the main-text example,

$$p = \frac{1}{2}, \quad h = 1, \quad a = \frac{3}{5}, \quad b = \frac{2}{5}, \quad \ell = 0,$$

so $W^{\text{coarse}} = 7/10$, $\overline{W} = 1/2$, and the persistent shortfall is at least $1/5$. The same obstruction applies on the open set (F.12); it requires neither singleton menus nor convergence of the full score vector. Vanishing prior influence is an assumption on the feedback retained by the aggregation rule, not an automatic consequence of positive logit choice probabilities. A rule that discards its feedback need not satisfy it. Subject to calibration, however, even unrestricted reweighting of C's observations cannot reverse the ordering of its two choice probabilities that drives the bound.

F.6 Convergence in the example

Lemma F.10 (Global convergence with a persistent welfare loss)**.** *Consider equally likely menus $\{A, C\}$ and $\{B, C\}$ with deterministic payoffs $(3/5, 1)$ and $(2/5, 0)$, respectively. For every fixed finite $\beta \geq 0$, there is a unique SVE $v^\beta = (3/5, 2/5, c_\beta)$. It is globally asymptotically stable and locally exponentially stable under the normalized dynamics, as is $(v^\beta, \Lambda(v^\beta))$ under the joint dynamics on $\mathbb{R}^3 \times \operatorname{ri}\Delta(\mathcal{S})$. From every finite initial valuation vector and positive prior counts, harmonic visit-count learning converges to this joint equilibrium almost surely. Moreover,*

$$v^\beta \longrightarrow (3/5, 2/5, 0) \qquad \text{as } \beta \to \infty.$$

This strict VE selects A and B in their respective menus. Equilibrium welfare therefore tends to $1/2$, below $W^{\text{coarse}} = 7/10$.

Proof. The conditional payoff targets satisfy $g_A \equiv 3/5$ and $g_B \equiv 2/5$, so every SVE has these two coordinates. Write $c = v_C$, $U = e^{3\beta/5}$, $V = e^{2\beta/5}$, and $t = e^{\beta c}$. At these fixed values of v_A, v_B, the probabilities of choosing C are $t/(t+U)$ and $t/(t+V)$. Its conditional payoff target is consequently

$$H_\beta(c) := g_C(3/5, 2/5, c) = \frac{t+V}{2t+U+V}, \qquad F_\beta(c) := H_\beta(c) - c.$$

At $\beta = 0$, $H_0 \equiv 1/2$, so $c_0 = 1/2$ and $F_0' = -1$. For $\beta > 0$, $U > V$ and $0 < H_\beta(c) < 1/2$ for every c. Thus every zero of F_β lies in $(0, 1/2)$, and $F_\beta(0) > 0 > F_\beta(1/2)$. Direct differentiation gives

$$H_\beta'(c) = \frac{\beta t(U-V)}{(2t+U+V)^2}, \qquad H_\beta''(c) = \frac{\beta^2 t(U-V)(U+V-2t)}{(2t+U+V)^3}.$$

For $0 \leq c \leq 1/2$, $2t \leq 2e^{\beta/2} = 2\sqrt{UV} \leq U+V$, so F_β is convex on this interval. Its endpoint signs give a zero; two distinct zeros would, by monotonicity of convex secant slopes, force $F_\beta(1/2) \geq 0$. The zero c_β is therefore unique, and

$$F_\beta'(c_\beta) \leq \frac{F_\beta(1/2) - F_\beta(c_\beta)}{1/2 - c_\beta} < 0.$$

The normalized Jacobian is lower triangular, with eigenvalues $-1, -1, F_\beta'(c_\beta)$. Proposition A.7 adds two eigenvalues equal to -1 for the joint dynamics on the frequency simplex. Both equilibria are thus locally exponentially stable. Because only C has a nonconstant conditional payoff target, Lemma C.5 gives convergence of every trajectory of either system to an equilibrium. Uniqueness and local stability therefore establish global asymptotic stability for both systems. The deterministic rewards satisfy the maintained moment assumption, so Corollary A.6 gives the almost-sure convergence of harmonic visit-count learning.

Finally, monotonicity of H_β gives, for $\beta > 0$,

$$0 < c_\beta = H_\beta(c_\beta) \le H_\beta(1/2) = \frac{1}{1 + e^{\beta/10}} \longrightarrow 0.$$

The two limiting valuation gaps against C are strictly positive, so its choice probabilities vanish in both menus. The limiting policy chooses A and B, whose constant payoffs equal their valuations; C is unchosen. Hence the limit is a strict VE, and equilibrium welfare tends to $\frac{1}{2}(3/5)+\frac{1}{2}(2/5) = 1/2$. □

G Algorithmic prominence and consumer learning

This section proves the two-state illustration, the propositions on optimal deterministic and randomized prominence, and the associated welfare and learning claims, including extensions to heterogeneous consumer partitions and prominence caps. The maintained specification has no commission on competitor sales, and uses a fixed stationary prominence rule. Prominence has no direct cost except in Subsection G.10. There is no cap except in Subsection G.8. A product state $i \in \{1, \dots, J\}$ is drawn independently and uniformly each period. The in-house product has expected payoff μ_i and margin $m_i \ge 0$; the competitor and outside option have respective expected payoffs $\mu_C > \mu_O$. Their reward laws satisfy the maintained moment assumptions. The platform chooses prominence with probability $\tau_i \in [0, 1]$ in state i, using fresh draws across periods that are independent of reward shocks and choice randomization. The resulting class menu is $\{H, O\}$ under prominence and $\{H, C, O\}$ otherwise. Consumers use their common class valuations and harmonic visit-count learning. Among policies yielding the same limiting profit, the platform prefers lower prominence frequency $T(\tau) := J^{-1} \sum_i \tau_i$. More precisely, among attained primary maximizers, lower prominence frequency is preferred. A minimum-frequency maximizer is selected whenever that secondary minimum exists. This convention selects among profit maximizers; it does not change their profit values.

Throughout this section,

$$\bar{\mu} := \frac{1}{J} \sum_i \mu_i < \mu_O < \mu_C, \qquad \max_i \mu_i > \mu_O. \tag{G.1}$$

Sensitivity is fixed during learning. All high-sensitivity conclusions first take the long-run limit at a fixed finite β, and only then let $\beta \to \infty$. Approximations to a profit supremum by different fixed rules are a further limit; no interchange of these limits is assumed. Classes that are never available are omitted, as in the standing convention.

G.1 Reduction, convergence, and long-run payoffs

For a fixed rule put

$$T = \frac{1}{J} \sum_i \tau_i.$$

If $0 < T < 1$, define the promoted and unpromoted means

$$a = \frac{\sum_i \tau_i \mu_i}{JT}, \qquad b = \frac{\sum_i (1 - \tau_i)\mu_i}{J(1 - T)}, \qquad Ta + (1 - T)b = \bar{\mu}. \tag{G.2}$$

The class-menu probabilities are T and $1 - T$, and the conditional means of H are a and b. Indeed, conditional on the class menu, the probability of choosing H does not depend further on the physical state. Thus choosing H does not change the conditional distribution of states within that menu. This is exactly the reduction in Lemma A.1; it gives Alice neither the state payoffs nor their frequencies.

At an SVE the valuations of C and O are their constant targets. Writing v for the valuation of H, define

$$A_\beta(v) = \frac{e^{\beta v}}{e^{\beta v} + e^{\beta \mu_O}}, \qquad B_\beta(v) = \frac{e^{\beta v}}{e^{\beta v} + e^{\beta \mu_O} + e^{\beta \mu_C}}, \qquad R_\beta(v) = \frac{B_\beta(v)}{A_\beta(v)}. \tag{G.3}$$

These are the in-house choice probabilities under prominence, without prominence, and their ratio. In particular $0 < R_\beta(v) < 1$. The scalar conditional payoff target is

$$\begin{aligned} g_\beta(v) &= \frac{Ta + (1-T)bR_\beta(v)}{T + (1-T)R_\beta(v)} \\ &= \bar{\mu} + \frac{T(a-\bar{\mu})e^{\beta \mu_C}}{e^{\beta v} + e^{\beta \mu_O} + Te^{\beta \mu_C}}. \end{aligned} \tag{G.4}$$

It lies between a and $\bar{\mu}$. Its derivative has the sign of $\bar{\mu} - a$ (and is zero at $\beta = 0$).

Proposition G.1 (Learning under a fixed prominence rule)**.** *Under* (G.1)*, every fixed stationary prominence rule admits a unique SVE at every finite sensitivity. Its joint equilibrium is globally asymptotically stable and locally exponentially stable on the active-class state space. Harmonic visit-count learning converges to it almost surely from every finite initial valuation vector and positive initial pseudo-counts.*

Proof. Suppose first that $0 < T < 1$. The equilibrium equation for H is $v = g_\beta(v)$. If $a \geq \bar{\mu}$, then g_β is non-increasing, so $g_\beta(v) - v$ is strictly decreasing and has exactly one zero; existence also follows from the payoff-box fixed-point argument. At that zero, $g'_\beta(v) - 1 \leq -1$.

If $a < \bar{\mu}$, every zero lies in $(a, \bar{\mu}) \subset (-\infty, \mu_O)$. For $\beta > 0$, differentiating (G.4) and using the fixed-point equation gives, at each such zero,

$$g'_\beta(v) = \frac{\beta(\bar{\mu} - v)e^{\beta v}}{e^{\beta v} + e^{\beta \mu_O} + Te^{\beta \mu_C}} < \frac{u}{1 + e^u} < 1, \qquad u = \beta(\mu_O - v) > 0.$$

At $\beta = 0$ the derivative is zero. Hence every zero of $g_\beta - v$ is a strict downward crossing. The function is analytic and is not identically zero, since g_β is bounded whereas v is not. Its zeros in the compact payoff interval are therefore finite; two downward crossings would require an intervening zero that is not a downward crossing. Thus the zero is unique in this case as well, and $g'_\beta(v) - 1 < 0$ there.

All other class targets are constant. Lemma C.5 therefore gives convergence of every joint trajectory to an equilibrium, which must be the unique one just identified. In coordinates ordered as H, C, O, the normalized valuation Jacobian is upper triangular, with eigenvalues $g'_\beta(v) - 1, -1, -1$. Proposition A.7 adds only negative frequency eigenvalues to the joint linearization. The joint equilibrium is thus locally exponentially stable and globally attracting, hence globally asymptotically stable. Almost-sure convergence follows from Corollary A.6.

If $T = 0$, every state generates the same full menu and $g_H \equiv \bar{\mu}$. If $T = 1$, every state generates $\{H, O\}$, the unavailable competitor is omitted, and again $g_H \equiv \bar{\mu}$. All active targets are constant in either case, so Proposition A.8 applies. □

Let v_β denote the resulting equilibrium valuation of H and write

$$\begin{aligned} d_i^\beta &= \tau_i A_\beta(v_\beta) + (1 - \tau_i)B_\beta(v_\beta), \\ c_\beta &= (1-T)\frac{e^{\beta \mu_C}}{e^{\beta v_\beta} + e^{\beta \mu_O} + e^{\beta \mu_C}}, \\ o_\beta &= T\big(1 - A_\beta(v_\beta)\big) + (1-T)\frac{e^{\beta \mu_O}}{e^{\beta v_\beta} + e^{\beta \mu_O} + e^{\beta \mu_C}}. \end{aligned} \tag{G.5}$$

Here d_i^β is conditional on state i, while c_β, o_β are unconditional market shares. These expressions remain valid at $T = 0, 1$ by ignoring the probability of choice in a menu that never occurs. The limiting calendar-time profit and consumer payoff at fixed sensitivity are, respectively,

$$\Pi_\beta(\tau) = \frac{1}{J}\sum_i m_i d_i^\beta, \qquad W_\beta(\tau) = \frac{1}{J}\sum_i \mu_i d_i^\beta + \mu_C c_\beta + \mu_O o_\beta. \tag{G.6}$$

These are also the almost-sure limits of realized time averages. To see this, subtract the conditional expectation from each state-choice indicator and from each period's reward or profit. Indicator and profit martingale differences are bounded, and reward differences have bounded conditional second moments. The martingale series obtained by dividing the period-k difference by $k+1$ converges, so Kronecker's lemma makes its Cesàro average vanish. Convergence of valuations and continuity of the finite-β logits make the corresponding conditional expectations converge to (G.5)–(G.6). In particular, no identification of a deterministic time average with a stochastic time average is being assumed without this argument.

G.2 The high-sensitivity limit of any fixed rule

Lemma G.2 (Fixed-rule purchase limits)**.** *Suppose $0 < T < 1$, and use (G.2). Equilibrium valuations and purchase probabilities have limits as $\beta \to \infty$. They are given by*

promoted mean	$\lim v_\beta$	$\lim A_\beta(v_\beta)$	$\lim B_\beta(v_\beta)$
$a < \mu_O$	a	0	0
$a = \mu_O$	μ_O	1/2	0
$\mu_O < a \le \mu_C$	a	1	0
$a > \mu_C$	μ_C	1	q_τ

(G.7)

In the last row,

$$q_\tau = \frac{T(a-\mu_C)}{(1-T)(\mu_C - b)} = \frac{\sum_i \tau_i(\mu_i - \mu_C)}{\sum_i (1-\tau_i)(\mu_C - \mu_i)} \in (0,1). \tag{G.8}$$

Thus the limiting sales vector is respectively 0, $\tau/2$, τ, or $(\tau_i + (1-\tau_i)q_\tau)_i$. At $T = 0$ and $T = 1$ it is zero, because $v_\beta = \bar\mu < \mu_O$. Consequently $\Pi^\infty(\tau) := \lim_{\beta\to\infty}\Pi_\beta(\tau)$ and $W^\infty(\tau) := \lim_{\beta\to\infty} W_\beta(\tau)$ are well defined for every fixed stationary rule.

Proof. Write $\rho_\beta = R_\beta(v_\beta)$. If $a < \mu_O$, every root is at most $u_0 := \max\{a, \bar\mu\} < \mu_O$. Therefore

$$0 < \rho_\beta = \frac{e^{\beta v_\beta} + e^{\beta\mu_O}}{e^{\beta v_\beta} + e^{\beta\mu_O} + e^{\beta\mu_C}} \le 2e^{-\beta(\mu_C - \mu_O)} \longrightarrow 0.$$

Equation (G.4) gives $v_\beta \to a$, and the two purchase probabilities tend to zero.

If $a \ge \mu_O$, then $a > \bar\mu$ and $b < \bar\mu < \mu_O$. For $a = \mu_O$ the equilibrium satisfies $v_\beta < a$, and

$$0 \le a - v_\beta = \frac{(1-T)(a-b)\rho_\beta}{T + (1-T)\rho_\beta} \le \frac{2(1-T)(a-b)}{T} e^{-\beta(\mu_C - \mu_O)}.$$

Hence $\beta(v_\beta - \mu_O) \to 0$, proving the limiting promoted purchase probability 1/2. This is a rate calculation, rather than an inference from equality of limiting valuations alone.

If $\mu_O < a < \mu_C$, then $v_\beta \le a < \mu_C$, which again forces $\rho_\beta \to 0$ and $v_\beta \to a$. The limits are now $A_\beta \to 1$ and $B_\beta \to 0$. If $a = \mu_C$, a subsequential valuation limit strictly below μ_C would force

$\rho_\beta \to 0$ and hence $v_\beta \to a$, a contradiction. Since $v_\beta \le a$, it follows that $v_\beta \to \mu_C$. Rearranging the fixed-point equation gives

$$\rho_\beta = \frac{T(a - v_\beta)}{(1-T)(v_\beta - b)}. \tag{G.9}$$

Its denominator stays positive in the limit and its numerator tends to zero. Thus $B_\beta \to 0$ even though $v_\beta \to \mu_C$.

Finally suppose $a > \mu_C$. A subsequential valuation limit below μ_C would force $\rho_\beta \to 0$ and $v_\beta \to a > \mu_C$. A limit above μ_C would force $\rho_\beta \to 1$ and $v_\beta \to \bar{\mu} < \mu_O$. Compactness therefore implies $v_\beta \to \mu_C$. Equation (G.9) gives $\rho_\beta \to q_\tau$; $A_\beta \to 1$ then gives $B_\beta \to q_\tau$. Its numerator is positive, and the denominator exceeds it by $\mu_C - \bar{\mu} > 0$, proving $0 < q_\tau < 1$. The remaining assertions follow from (G.5)–(G.6). □

G.3 The two-state illustration

The following lemma collects the claims in the main text's two-state illustration and its welfare comparison.

Lemma G.3 (Two-state prominence and welfare)**.** *Suppose $J = 2$,*

$$\mu_h > \mu_C > \mu_O > \mu_\ell, \qquad (\mu_h + \mu_\ell)/2 < \mu_O, \qquad m_h, m_\ell > 0.$$

High-only prominence uniquely maximizes limiting profit among deterministic rules. Its limiting valuation is μ_C, its sales are $(1, q_2)$, and

$$q_2 = \frac{\mu_h - \mu_C}{\mu_C - \mu_\ell} \in (0,1), \qquad \Pi^\infty = \frac{m_h + q_2 m_\ell}{2}, \qquad W^\infty = \mu_C. \tag{G.10}$$

For that same prominence rule, the informed coarse oracle earns consumer payoff $(\mu_h + \mu_C)/2$.

Proof. There are four deterministic rules. No prominence, low-only prominence, and universal prominence have zero limiting sales by Lemma G.2. For high-only prominence, $T = 1/2$, $a = \mu_h > \mu_C$ and $b = \mu_\ell$, giving (G.10) and strictly positive profit. Its welfare equals

$$\frac{\mu_h + q_2\mu_\ell + (1-q_2)\mu_C}{2} = \mu_C,$$

because $\mu_h - \mu_C = q_2(\mu_C - \mu_\ell)$. No prominence also gives welfare μ_C, since the competitor is then chosen in every state.

With high-only prominence, the highest available expected payoff is μ_h in state h and μ_C in state ℓ. The common ranking $C \succ H \succ O$ implements both choices, so it attains the menu-wise upper bound and is the coarse oracle optimum of Proposition F.1. The welfare gap is $(\mu_h - \mu_C)/2$. This efficient ranking cannot be a self-confirming CQL limit: its positive in-house purchases occur only in state h and therefore calibrate H to $\mu_h > \mu_C$, contrary to its ranking below C. Under complete information prominence does not change sales in either state, since the high product exceeds μ_C and the low product is below μ_O. The platform's tie-breaking convention therefore selects no prominence under complete information. □

G.4 Optimal deterministic prominence

This is the main text's deterministic-prominence proposition, with its maintained restriction on the mean payoff of the two promoted products.

Proposition G.4 (Optimal deterministic prominence). *Suppose* (G.1) *holds,* $J = 3$*, and*

$$\mu_h > \mu_C > \mu_m > \mu_O > \mu_\ell, \qquad \frac{\mu_h + \mu_m}{2} \leq \mu_C, \qquad m_h, m_m, m_\ell > 0. \tag{G.11}$$

Define

$$q_3 := \frac{\mu_h - \mu_C}{2\mu_C - \mu_m - \mu_\ell} \in (0, 1). \tag{G.12}$$

Only high-only and high-and-medium prominence can maximize limiting profit among deterministic rules. Their purchase probabilities in states (h, m, ℓ) *and their profits are*

$$\begin{aligned} d^h &= (1, q_3, q_3), & \Pi_h^\infty &= \frac{m_h + q_3(m_m + m_\ell)}{3}, \\ d^{hm} &= (1, 1, 0), & \Pi_{hm}^\infty &= \frac{m_h + m_m}{3}. \end{aligned} \tag{G.13}$$

High-only prominence uniquely maximizes profit if and only if

$$q_3 m_\ell > (1 - q_3) m_m. \tag{G.14}$$

The reverse strict inequality selects high-and-medium prominence. At equality, both rules maximize profit, and the tie-breaking convention selects high-only prominence.

Proof. No prominence has zero limiting sales. Every promoted subset containing ℓ also has mean below μ_O: for example,

$$\mu_h + \mu_\ell < 3\mu_O - \mu_m < 2\mu_O, \qquad \mu_m + \mu_\ell < 3\mu_O - \mu_h < 2\mu_O.$$

This covers both pairs; the singleton and the full set satisfy the same claim by assumption. These four rules have zero limiting sales by Lemma G.2.

Medium-only prominence has promoted mean $\mu_m \in (\mu_O, \mu_C)$ and therefore sells only in state m. High-only prominence has $a = \mu_h$ and $b = (\mu_m + \mu_\ell)/2$, giving d^h in (G.13) and limiting brand valuation μ_C. High-and-medium prominence has $a = (\mu_h + \mu_m)/2 \in (\mu_O, \mu_C]$ and $b = \mu_\ell$. Lemma G.2 gives limiting valuation a and sales $(1, 1, 0)$. This includes $a = \mu_C$: the lemma establishes that the unpromoted purchase probability tends to zero even at that boundary. Since $m_h > 0$, high-and-medium prominence strictly dominates medium-only prominence. Both candidate rules earn positive profit. The eight deterministic rules have now all been covered.

The numerator defining q_3 is positive, and its denominator exceeds the numerator by $3\mu_C - \mu_h - \mu_m - \mu_\ell = 3(\mu_C - \bar{\mu}) > 0$. Hence $q_3 \in (0, 1)$. Finally,

$$3(\Pi_h^\infty - \Pi_{hm}^\infty) = q_3 m_\ell - (1 - q_3) m_m,$$

which proves both strict comparisons and the equality case. At equality, the prominence frequencies are $1/3$ and $2/3$, so tie-breaking selects high-only prominence. In particular m_h cancels; no ordering of high- and low-state margins has been imposed. □

Under complete information, high-quality purchases occur irrespective of prominence, low-quality purchases never occur, and medium-quality purchases occur precisely under prominence. Positive m_m therefore requires promoting m for profit maximization; the tie-breaking convention selects medium-only prominence. Its CQL outcome instead has valuation μ_m and purchases only in state m, as just proved. The comparison is thus between policies under different consumer models, not an assertion that the coarse learner can implement informed choices.

High-only prominence gives welfare μ_C, because

$$\mu_h - \mu_C = q_3(2\mu_C - \mu_m - \mu_\ell).$$

Its conditional coarse-oracle welfare is $(\mu_h + 2\mu_C)/3$, implemented by $C \succ H \succ O$; the gap is $(\mu_h - \mu_C)/3$. As in Lemma G.3, this ranking cannot be sustained by CQL: its brand purchases occur only in state h, calibrating the brand to $\mu_h > \mu_C$. No prominence gives welfare μ_C, because the competitor is then chosen in every state. With high-and-medium prominence, welfare is $(\mu_h + \mu_m + \mu_C)/3 \leq \mu_C$. The competitor's shares under the two candidate policies are $2(1 - q_3)/3$ and $1/3$, both strictly positive.

G.5 Optimal randomized prominence

Proposition G.5 (Optimal randomized prominence)**.** *Under* (G.1)*, with $m_i \geq 0$, the supremum of limiting profit over fixed stationary prominence rules is*

$$\sup_{\tau \in [0,1]^J} \Pi^\infty(\tau) = \max_{d \in [0,1]^J} \left\{ \frac{1}{J} \sum_i m_i d_i : \sum_i (\mu_i - \mu_O) d_i \geq 0 \right\}. \tag{G.15}$$

The bound can be approached by a sequence of stationary rules that randomize in at most one state. Along this sequence, limiting brand purchases occur only under prominence and the limiting brand valuation approaches μ_O from above. If every margin is strictly positive, no fixed stationary rule attains this supremum. More generally, every stationary rule can be replaced by one randomizing in at most one state with weakly higher limiting profit, although its prominence frequency may change.

Proof. Upper bound. For every fixed rule, let $d_i = \lim_\beta d_i^\beta$, which exists by Lemma G.2. At finite sensitivity, the fixed-point equation gives the exact calibration identity

$$v_\beta \sum_i d_i^\beta = \sum_i \mu_i d_i^\beta. \tag{G.16}$$

If $\sum_i d_i > 0$, positive limiting purchases require the limiting valuation to be at least μ_O. Passing to the limit in (G.16) yields $\sum_i (\mu_i - \mu_O) d_i \geq 0$. If all d_i vanish, this inequality holds as equality. Since $0 \leq d_i \leq 1$, every realized limiting profit is bounded by the compact linear program on the right of (G.15).

Solving the linear program. Put

$$I_+ = \{i : \mu_i \geq \mu_O\}, \quad I_- = \{i : \mu_i < \mu_O\}, \quad S = \sum_{i \in I_+} (\mu_i - \mu_O) > 0, \quad w_i = \mu_O - \mu_i \quad (i \in I_-).$$

Increasing any d_i, $i \in I_+$, weakly increases the objective and preserves feasibility, so an optimizer can be chosen with these coordinates equal to one. The remaining problem is

$$\max \left\{ \sum_{i \in I_-} m_i d_i : 0 \leq d_i \leq 1, \quad \sum_{i \in I_-} w_i d_i \leq S \right\}. \tag{G.17}$$

Because $\bar{\mu} < \mu_O$, the total weight $\sum_{I_-} w_i$ exceeds S. With nonnegative margins, a maximizing vector can be chosen to use the entire budget: any unused budget can be filled in an unsaturated coordinate without reducing the objective, even when that margin is zero.

Order the ratios m_i/w_i non-increasingly, fill coordinates in that order, and use a fraction of the first coordinate whose full weight would exceed the remaining budget. This is optimal. Indeed, moving an equal amount of weight from an included lower-ratio coordinate to an unfilled higher-ratio coordinate weakly increases the objective. Applying these exchanges successively transforms any budget-exhausting feasible vector into the stated allocation without decreasing profit. The construction uses at most one fractional coordinate; ties may give additional optimizers. This establishes the asserted existence statements without claiming that every optimizer has a binding constraint when some margins vanish.

Approximate implementation with randomization in one state. Choose a binding optimizer d^* of the preceding form: every coordinate in I_+ equals one, and at most one coordinate in I_- is fractional. Since $\sum_{i\in I_-} w_i d_i^* = S > 0$, at least one below-outside-option product is included. If a fractional coordinate exists, let j be that state; otherwise choose any $j \in I_-$ with $d_j^* = 1$. In either case, $d_j^* > 0$ and $d_i^* \in \{0, 1\}$ for every $i \neq j$. For $0 < \varepsilon < d_j^*$, set

$$\tau_i^\varepsilon = \begin{cases} d_j^* - \varepsilon, & i = j, \\ d_i^*, & i \neq j. \end{cases} \tag{G.18}$$

Every rule randomizes only in state j, which is fixed along the sequence. Since d^* binds the constraint,

$$\sum_i (\mu_i - \mu_O)\tau_i^\varepsilon = \varepsilon w_j > 0.$$

The rule uses both menus: prominence is certain in every state in I_+, while $\tau_j^\varepsilon < 1$. Its promoted mean satisfies

$$a_\varepsilon = \mu_O + \frac{\varepsilon w_j}{\sum_i \tau_i^\varepsilon} \longrightarrow \mu_O \quad \text{from above.}$$

For all sufficiently small ε, it lies below μ_C. Lemma G.2 then gives limiting sales exactly $d_i^\varepsilon = \tau_i^\varepsilon$. Their profit shortfall from the linear program's value is exactly $\varepsilon m_j/J$, which tends to zero (and is identically zero when $m_j = 0$). Proposition G.1 supplies the unique globally attracting joint equilibrium and almost-sure learning convergence for each of these fixed rules and each finite sensitivity. This proves the lower bound in (G.15) with the stated order of limits.

Non-attainment with positive margins. If all $m_i > 0$, every optimizer has $d_i = 1$ on I_+: increasing such a coordinate strictly improves the objective while preserving feasibility. Every optimizer also binds the constraint, since otherwise an unsaturated coordinate could be increased; the all-one vector is infeasible. Thus its positive sales have purchase-weighted mean exactly μ_O. For any fixed stationary rule attaining this objective, calibration would force its limiting valuation to equal μ_O. Lemma G.2 then forces $a = \mu_O$ and $d_i = \tau_i/2 \leq 1/2$, contradicting $d_i = 1$ on the nonempty set I_+. Hence the value is a supremum rather than a maximum over fixed rules. In particular, simply setting $\tau = d^*$ gives promoted mean μ_O and halves the intended promoted purchases, rather than implementing d^*. The final, policy-by-policy assertion follows from Lemma G.12(ii). □

G.6 The illustrative randomized optima and their welfare

Corollary G.6 (Two and three states under randomized prominence)**.** *In the two-state environment of Lemma G.3, put*

$$t_2^* = \frac{\mu_h - \mu_O}{\mu_O - \mu_\ell} \in (q_2, 1).$$

The unique optimizing sales vector in (G.15) *is* $(1, t_2^*)$*. Its profit supremum is* $(m_h + t_2^* m_\ell)/2$*, approached by* $(\tau_h, \tau_\ell) = (1, t)$ *as* $t \uparrow t_2^*$*.*

In the three-state environment of Proposition G.4, put

$$t_3^* = \frac{\mu_h + \mu_m - 2\mu_O}{\mu_O - \mu_\ell} \in (q_3, 1).$$

The unique optimizing sales vector is $(1, 1, t_3^*)$*. Its profit supremum is* $(m_h + m_m + t_3^* m_\ell)/3$*, approached by* $(\tau_h, \tau_m, \tau_\ell) = (1, 1, t)$ *as* $t \uparrow t_3^*$*. Each supremum strictly exceeds every deterministic policy's profit.*

For every sequence of stationary rules whose limiting profits approach the corresponding supremum, limiting consumer welfare and competitor share converge, respectively, to

$$W_j^* = \mu_C - \frac{\mu_C - \mu_O}{j}(j - 1 + t_j^*), \qquad c_j^* = \frac{1 - t_j^*}{j} > 0, \qquad j = 2, 3. \tag{G.19}$$

In each case the limiting brand valuation approaches μ_O*, and* $W_j^* < \mu_C$*, the no-prominence welfare level.*

Proof. In either example there is a single state below μ_O. Positive margins force every above-outside-option coordinate to be one and the remaining coordinate to exhaust the feedback constraint. This proves uniqueness and the formulas for t_2^*, t_3^*. Positivity of their numerators and the strict unconditional-mean assumption put both in $(0, 1)$. The implementations follow from the same promoted-mean argument as (G.18).

For two states, the function $(\mu_h - z)/(z - \mu_\ell)$ is strictly decreasing on (μ_ℓ, μ_h), so $t_2^* > q_2$. For three states the high-only sales vector $(1, q_3, q_3)$ has weighted mean $\mu_C > \mu_O$. Increasing its medium-state coordinate from q_3 to one adds $(1-q_3)(\mu_m - \mu_O) > 0$ to its strict feedback surplus. The vector $(1, 1, q_3)$ therefore still has positive surplus, so the binding low-state coordinate t_3^* is strictly larger than q_3. Thus $(1, 1, t_3^*)$ weakly increases every coordinate of both deterministic candidate vectors $(1, q_3, q_3)$ and $(1, 1, 0)$, and strictly increases at least one coordinate of each. Strictly positive margins prove the strict profit comparisons.

It remains to establish welfare for *every* approximately optimal sequence, rather than only the displayed implementation. Let τ^n be such a sequence and let d^n be its limiting sales vector. Each d^n is feasible for the compact linear program. Uniqueness of its optimizer implies $d^n \to d^*$: otherwise a convergent subsequence bounded away from d^* would have an equally profitable feasible limit. In particular, $d_h^n \to 1$.

Let A^n be the high-sensitivity limit of the promoted purchase probability for rule τ^n, defining the counterfactual probability when the promoted menu is absent. From $B_\beta \le A_\beta$ we have $d_h^n \le A^n \le 1$, so $A^n \to 1$. The outside share in unpromoted encounters is at most $e^{-\beta(\mu_C - \mu_O)}$, uniformly over scores and rules. In promoted encounters it is $T^n(1 - A^n)$. Consequently the limiting outside shares o^n tend to zero, and the competitor shares converge to $1 - J^{-1}\sum_i d_i^*$.

The limiting calibration identity gives $v^n = (\sum_i \mu_i d_i^n)/(\sum_i d_i^n)$ whenever the denominator is positive; it tends to μ_O, since the optimal feedback constraint binds. Finally,

$$W^\infty(\tau^n) = \frac{1}{J}\sum_i \mu_i d_i^n + \mu_C c^n + \mu_O o^n \longrightarrow \mu_O \frac{1}{J}\sum_i d_i^* + \mu_C \left(1 - \frac{1}{J}\sum_i d_i^*\right).$$

Substituting the two optimizing sales vectors proves (G.19) and all the remaining assertions. □

For J states, every binding maximizing allocation constructed in Proposition G.5 leaves a strictly positive competitor share in its approximate implementation: the all-one vector is infeasible under $\bar{\mu} < \mu_O$, so $J^{-1}\sum_i d_i^* < 1$. Its welfare tends to $\mu_C - (\mu_C - \mu_O)J^{-1}\sum_i d_i^*$. When there are several below-outside-option states, tied margin-to-shortfall ratios can give several optimizing sales vectors and different welfare limits. No general uniqueness claim is needed for the profit proposition.

The positive competitor share in these randomized implementations is distinct from consumer mixing within a menu. Along the approximating rules, $\mu_O < v_H^\infty < \mu_C$, so the brand is purchased under prominence and the competitor without it. It is the platform's state-contingent prominence that randomizes between these encounters. By contrast, the deterministic high-only examples have $v_H^\infty = \mu_C$ and strictly interior brand choice in unpromoted menus. The welfare decline in (G.19) is relative to no prominence. For each displayed approximate implementation, the ranking $C \succ H \succ O$ already implements the limiting choices and is optimal among common-score rankings at that same prominence rule. It is therefore a different comparison from the additional learning loss, holding prominence fixed, in the deterministic illustrations.

G.7 Heterogeneous consumer partitions

Let a fraction $\lambda \in [0,1]$ of consumers use the coarse classes H, C, O, and let the remainder distinguish the J in-house products, using classes $H_1, \dots, H_J, C, O$. Both types learn from their own repeated experiences using harmonic visit-count updates and the maintained logit rule. State probabilities, product payoffs, and margins are the same for both types. The platform knows λ and applies one fixed stationary prominence rule independently of consumer type and individual learning histories. Profits and welfare are population-weighted averages of the two types' long-run outcomes. We retain (G.1) and the order of limits specified above.

For a finest-partition consumer, define

$$a_i^{\mathrm{f}} = \begin{cases} 1 & \mu_i > \mu_O, \\ 1/2 & \mu_i = \mu_O, \\ 0 & \mu_i < \mu_O, \end{cases} \qquad b_i^{\mathrm{f}} = \begin{cases} 1 & \mu_i > \mu_C, \\ 1/2 & \mu_i = \mu_C, \\ 0 & \mu_i < \mu_C. \end{cases} \tag{G.20}$$

These are the high-sensitivity brand-purchase probabilities with and without prominence. Write $\Pi_{\mathrm{c}}^\infty(\tau)$ for the coarse-consumer profit studied above, and put

$$\Pi_{\mathrm{f}}^\infty(\tau) = \frac{1}{J} \sum_i m_i \big[\tau_i a_i^{\mathrm{f}} + (1-\tau_i) b_i^{\mathrm{f}}\big], \qquad \Pi_\lambda^\infty(\tau) = \lambda \Pi_{\mathrm{c}}^\infty(\tau) + (1-\lambda)\Pi_{\mathrm{f}}^\infty(\tau).$$

Let V_{c} be the linear-program value in (G.15), and define

$$V_{\mathrm{f}} = \frac{1}{J}\left(\sum_{\mu_i > \mu_O} m_i + \frac{1}{2} \sum_{\mu_i = \mu_O} m_i \right). \tag{G.21}$$

Proposition G.7 (Randomized prominence with heterogeneous consumers)**.** *After omitting unavailable classes, for every fixed stationary rule and finite sensitivity both consumer types have unique globally asymptotically stable joint equilibria, and their harmonic learning processes converge almost surely. Moreover,*

$$\sup_{\tau \in [0,1]^J} \Pi_\lambda^\infty(\tau) = \lambda V_{\mathrm{c}} + (1-\lambda) V_{\mathrm{f}}. \tag{G.22}$$

The policies in (G.18), constructed from an optimizer chosen as in Proposition G.5, randomize in at most one state and approach this bound simultaneously for every $\lambda \in [0,1]$. For each fixed λ, every stationary rule also admits a replacement randomizing in at most one state with weakly higher limiting population profit; its prominence frequency may change. If all margins are strictly positive and $\lambda > 0$, no fixed rule attains the bound. At $\lambda = 0$ the bound is attained; among its maximizers, the unique rule minimizing the frequency of prominence is

$$\tau_i = \mathbf{1}\{m_i > 0,\ \mu_O \le \mu_i \le \mu_C\}. \tag{G.23}$$

Proof. For finest-partition consumers, the payoff target of every active class is constant: it is μ_i for H_i, μ_C for C, and μ_O for O. Every H_i is available with probability $1/J$; C is omitted only if it is never available. Proposition A.8 therefore gives uniqueness, global asymptotic stability, and almost-sure convergence at every finite sensitivity. At that limit the finest consumer's brand-choice probabilities in state i are

$$\frac{e^{\beta\mu_i}}{e^{\beta\mu_i}+e^{\beta\mu_O}}, \qquad \frac{e^{\beta\mu_i}}{e^{\beta\mu_i}+e^{\beta\mu_O}+e^{\beta\mu_C}}.$$

Their high-sensitivity limits are (G.20). Thus finest-partition learning gives informed choices whenever payoff comparisons are strict; at a tie it gives the corresponding equal-probability logit limit. The martingale averaging argument following (G.6) applies unchanged, so these probabilities also determine almost-sure realized time-average payoffs. The coarse type's convergence follows from Proposition G.1.

Because $a_i^{\mathrm{f}} \geq b_i^{\mathrm{f}}$ and $m_i \geq 0$, maximizing the fine type's profit amounts to setting $\tau_i = 1$ whenever $m_i(a_i^{\mathrm{f}} - b_i^{\mathrm{f}}) > 0$. Its maximal value is $J^{-1}\sum_i m_i a_i^{\mathrm{f}} = V_{\mathrm{f}}$. The strict inequality holds exactly when $m_i > 0$ and $\mu_O \leq \mu_i \leq \mu_C$. All other coordinates leave profit unchanged, so setting them to zero uniquely minimizes $J^{-1}\sum_i \tau_i$. This proves (G.23).

For every common policy, each type's profit is bounded by its individual supremum, giving the upper bound in (G.22). The implementing policies (G.18) set $\tau_i = 1$ on every state with $\mu_i \geq \mu_O$. They therefore attain V_{f} exactly for every fixed ε. Changing prominence on states below μ_O does not affect fine-consumer brand sales, since both probabilities in (G.20) vanish there. Meanwhile, their coarse-consumer profit tends to V_{c} by Proposition G.5. This proves the lower bound with a policy family independent of λ. The policy-by-policy strengthening follows from Lemma G.12(ii); that replacement may depend on λ.

If all margins are positive and $\lambda > 0$, an attaining common policy would have to attain V_{c}, since

$$\lambda V_{\mathrm{c}} + (1-\lambda)V_{\mathrm{f}} - \Pi_\lambda^\infty(\tau) = \lambda\big(V_{\mathrm{c}} - \Pi_{\mathrm{c}}^\infty(\tau)\big) + (1-\lambda)\big(V_{\mathrm{f}} - \Pi_{\mathrm{f}}^\infty(\tau)\big), \tag{G.24}$$

and both terms are nonnegative. This contradicts the non-attainment part of Proposition G.5. The endpoint $\lambda = 0$ was established directly above. □

For any fixed $\lambda > 0$, identity (G.24) also implies that every sequence approaching the mixed-population bound approaches V_{c} for coarse consumers; if $\lambda < 1$, it approaches V_{f} for fine consumers as well. The result identifies a common approximately optimal policy family, not a unique family for general J. Tied margin-to-shortfall ratios can generate several maximizing coarse sales allocations with different welfare, as discussed above. The assumption of a common rule is substantive for welfare: if types were observed and prominence could be personalized, the two individual bounds would still bound total profit, but unnecessary low-product prominence could be withheld from finest-partition consumers.

Corollary G.8 (Heterogeneous consumers in the illustrations)**.** *Maintain the respective two- and three-state assumptions in Lemma G.3 and Proposition G.4, including strictly positive margins. Let q_j and t_j^* be as defined there and in Corollary G.6.*

For every $\lambda > 0$, high-only prominence remains the unique profit-maximizing deterministic rule in the two-state illustration, with profit $(m_h + \lambda q_2 m_\ell)/2$. In the three-state illustration, only high-only and high-and-medium prominence can maximize deterministic profit, with respective values

$$\Pi_{\lambda,h}^\infty = \frac{m_h + \lambda q_3(m_m + m_\ell)}{3}, \qquad \Pi_{\lambda,hm}^\infty = \frac{m_h + m_m}{3}. \tag{G.25}$$

With ties resolved in favor of less frequent prominence, high-only prominence is selected precisely when

$$\lambda q_3(m_m + m_\ell) \geq m_m; \tag{G.26}$$

otherwise high-and-medium prominence is selected. At $\lambda = 0$, the least-prominence profit maximizers are no prominence in the two-state illustration and medium-only prominence in the three-state illustration.

For every fixed $\lambda > 0$, the randomized profit suprema are

$$V_{\lambda,2} = \frac{m_h + \lambda t_2^* m_\ell}{2}, \qquad V_{\lambda,3} = \frac{m_h + m_m + \lambda t_3^* m_\ell}{3}. \tag{G.27}$$

Each strictly exceeds every deterministic rule's profit. Every sequence of stationary rules approaching the respective supremum satisfies

$$\tau^n \longrightarrow \begin{cases} (1, t_2^*), & j = 2, \\ (1, 1, t_3^*), & j = 3. \end{cases} \tag{G.28}$$

Along every such sequence, aggregate brand, competitor, and outside shares converge, respectively, to

$$\frac{j - 1 + \lambda t_j^*}{j}, \qquad \frac{1 - t_j^*}{j}, \qquad \frac{(1-\lambda) t_j^*}{j}, \qquad j = 2, 3. \tag{G.29}$$

Welfare converges to $W_{\lambda,j}^* = \lambda W_j^* + (1-\lambda) W_{\mathrm{f},j}^*$, where W_j^* is given by (G.19) and

$$W_{\mathrm{f},2}^* = \frac{\mu_h + t_2^* \mu_O + (1 - t_2^*)\mu_C}{2}, \qquad W_{\mathrm{f},3}^* = \frac{\mu_h + \mu_m + t_3^* \mu_O + (1 - t_3^*)\mu_C}{3}. \tag{G.30}$$

Relative to no prominence, aggregate welfare falls by

$$\Delta_{\lambda,2} = \frac{(\mu_C - \mu_O)(\lambda + t_2^*)}{2}, \qquad \Delta_{\lambda,3} = \frac{(\mu_C - \mu_O)(2\lambda + t_3^*) + (1-\lambda)(\mu_C - \mu_m)}{3}. \tag{G.31}$$

Proof. In both illustrations, a fine consumer buys the high product regardless of prominence and never buys the low product. With three states, the medium product is purchased exactly under prominence. Thus two-state fine profit is $m_h/2$ under every deterministic rule. Combining this with Lemma G.3 proves the two-state deterministic claim for $\lambda > 0$.

With three states, fine profit equals $(m_h + \tau_m m_m)/3$. Coarse profit under the eight deterministic rules was computed in Proposition G.4. High-and-medium prominence strictly dominates medium-only prominence by the additional coarse high-state sales, worth $\lambda m_h/3 > 0$. It also strictly dominates every rule whose coarse sales vanish: their fine profits are at most $(m_h + m_m)/3$, whereas high-and-medium prominence produces positive coarse sales and the same maximal fine profit. This leaves only the stated pair. Their weighted profits are (G.25), with difference $[\lambda q_3(m_m + m_\ell) - m_m]/3$. The high-only rule has prominence frequency $1/3$, as against $2/3$, so the less-prominence convention selects it at equality. The $\lambda = 0$ conclusions follow from (G.23).

The randomized values follow from (G.22) because $V_{\mathrm{f}} = m_h/2$ or $(m_h + m_m)/3$, respectively. The difference between the two-state randomized bound and high-only profit is $\lambda(t_2^* - q_2)m_\ell/2 > 0$. In three states its difference from high-and-medium profit is $\lambda t_3^* m_\ell/3 > 0$, and from high-only profit is

$$\frac{(1 - \lambda q_3)m_m + \lambda(t_3^* - q_3)m_\ell}{3} > 0.$$

These comparisons cover all deterministic candidates.

Fix $\lambda > 0$ and an approximately optimal sequence τ^n. Equation (G.24) forces its coarse profits to converge to V_c. Let d^n be its limiting coarse sales vector, and v^n the limiting coarse brand valuation. The proof of Corollary G.6 gives $d^n \to d^*$ and $v^n \to \mu_O$, where d^* is the unique optimizer $(1, t_2^*)$ or $(1, 1, t_3^*)$. In particular, $d_h^n \to 1$. For all sufficiently large n, $d_h^n > 1/2$ and $v^n < \mu_C$. The complete classification in Lemma G.2 now forces the promoted mean a_n to satisfy $\mu_O < a_n < \mu_C$: absent or universal prominence and $a_n < \mu_O$ give zero sales; $a_n = \mu_O$ gives $d_h^n \le 1/2$; and $a_n \ge \mu_C$ gives $v^n = \mu_C$. Hence $d^n = \tau^n$ for all sufficiently large n, proving (G.28). This step is needed to identify the fine type's outcomes for every approximately optimal sequence, rather than only for the displayed implementation.

Eventually, coarse consumers therefore buy the brand under prominence and the competitor otherwise. Fine consumers always buy h, buy m under prominence when present, and choose O under low-state prominence or C without it. Taking the policy limits gives the shares in (G.29) and the fine welfare in (G.30). Coarse welfare tends to W_j^* by Corollary G.6; weighting the two types gives $W_{\lambda,j}^*$.

Without prominence, coarse welfare is μ_C and fine welfare is $[\mu_h + (j-1)\mu_C]/j$. Subtracting the preceding limits from their population-weighted average gives (G.31). In particular the fine type's own losses are $t_2^*(\mu_C - \mu_O)/2$ and $[\mu_C - \mu_m + t_3^*(\mu_C - \mu_O)]/3$, both positive. These consumers avoid inferior brand purchases but cannot restore a competitor excluded by the common policy. All sequence conclusions fix $\lambda > 0$; at $\lambda = 0$, profit maximization with the tie-breaking convention selects the different endpoint policies already stated. □

G.8 Prominence caps

Return to the all-coarse benchmark and maintain (G.1). A prominence cap restricts a stationary rule to

$$T(\tau) := \frac{1}{J}\sum_i \tau_i \le \rho, \qquad 0 \le \rho \le 1.$$

The cap concerns the frequency of prominence, rather than purchases. As before, learning first reaches its long-run outcome at fixed sensitivity, sensitivity then increases without bound, and any sequence of different stationary rules is considered only thereafter. Put $c = J\rho$ and $M(d) = J^{-1}\sum_i m_i d_i$. The first program below describes sales supported by purchases only under prominence, including their limits. The second allows purchases when the competitor remains available, with the brand calibrated to μ_C.

Proposition G.9 (Optimal randomized prominence under a cap)**.** *Define the two linear-program values*

$$\begin{aligned} A_\rho := \max_{d\in[0,1]^J} \quad & M(d) \\ \textit{subject to} \quad & \sum_i d_i \le c, \\ & \sum_i (\mu_i - \mu_O) d_i \ge 0, \qquad \sum_i (\mu_i - \mu_C) d_i \le 0. \end{aligned} \tag{G.32}$$

$$\begin{aligned} B_\rho := \max_{d\in[0,1]^J} \quad & M(d) \\ \textit{subject to} \quad & \sum_i (\mu_i - \mu_C) d_i = 0, \\ & \sum_j d_j - (J-c) d_i \le c \quad \textit{for every } i. \end{aligned} \tag{G.33}$$

Then

$$P_\rho := \sup_{T(\tau)\le\rho} \Pi^\infty(\tau) = \max\{A_\rho, B_\rho\}. \tag{G.34}$$

Every feasible allocation in (G.33) *is implemented by setting*

$$q = \min_i d_i, \qquad \tau_i = \frac{d_i - q}{1-q}. \tag{G.35}$$

An allocation feasible for (G.32) *with* $\sum_i(\mu_i-\mu_O)d_i > 0$ *is implemented by* $\tau = d$*; every other feasible allocation is a limit of implementable allocations when* $\rho > 0$*. Consequently, a profit-maximizing stationary rule exists unless* $A_\rho > B_\rho$ *and every optimizer of* (G.32) *satisfies* $\sum_i(\mu_i - \mu_O)d_i = 0$*. In precisely that case, the value is a supremum that is not attained. At* $\rho = 0$*, the sole feasible rule is no prominence and* $P_0 = 0$*.*

Every stationary rule has a replacement with the same prominence frequency and weakly higher limiting profit that randomizes in at most two states. Hence an attained P_ρ *has such a maximizer, and an unattained* P_ρ *can be approached by such rules. The reduction also holds for any fixed mixture of coarse and finest-partition consumers, although* (G.34) *concerns the all-coarse case. The two-state bound is sharp even with a positive finest-partition share.*

Proof. Both programs have nonempty compact feasible sets and continuous linear objectives, so their maxima exist. First suppose $\rho > 0$.

Upper bound. Consider a fixed feasible rule. The zero-sales cases in Lemma G.2 yield zero profit. If its promoted mean satisfies $\mu_O < a \le \mu_C$, its limiting sales are $d = \tau$, which is feasible for (G.32). If $a = \mu_O$, the limiting sales $d = \tau/2$ are also feasible for that program. If $a > \mu_C$, the sales satisfy

$$d_i = q + (1-q)\tau_i, \qquad 0 < q < 1, \qquad \sum_i(\mu_i - \mu_C)d_i = 0.$$

Writing $D = \sum_i d_i$, the cap is equivalent to $D \le c + (J-c)q$. Since $q \le d_i$ for every i and $J - c \ge 0$, the constraints of (G.33) follow. This proves the upper bound in (G.34).

Implementation of the second program. Let d be feasible for (G.33) and take q as in (G.35). Calibration excludes $d = (1, \dots, 1)$, because $\bar\mu < \mu_C$, so $q < 1$. The recovered rule belongs to $[0,1]^J$. The constraint corresponding to an index with $d_i = q$ gives

$$T(\tau) = \frac{D - Jq}{J(1-q)} \le \rho.$$

If $d = 0$, this is no prominence. Otherwise $0 < T(\tau) < 1$: equality $T = 0$ would make every $d_i = q$, and calibration would then force $q = 0$; equality $T = 1$ would force every $d_i = 1$. Moreover,

$$\sum_i(\mu_i - \mu_C)\tau_i = \frac{qJ(\mu_C - \bar\mu)}{1-q}.$$

Thus $q = 0$ gives promoted mean $a = \mu_C$, and Lemma G.2 implements $d = \tau$. If $q > 0$, the promoted mean exceeds μ_C, and substitution in (G.8) gives $q_\tau = q$; the same lemma implements $d_i = q + (1-q)\tau_i$. This proves exact implementation. For a given nonzero sales allocation, the prominence frequency $(D - Jq)/(J(1-q))$ decreases with q, since $D < J$. Hence choosing $q = \min_i d_i$ also minimizes prominence among implementations of that allocation. This conditional statement does not select among different profit-maximizing allocations.

Implementation of the first program. If its outside-option constraint is strict, $d \ne 0$ and the rule $\tau = d$ has promoted mean in $(\mu_O, \mu_C]$. Its frequency is strictly below one, because the all-one

vector violates that constraint under $\bar{\mu} < \mu_O$. Lemma G.2 therefore gives limiting sales exactly d. To handle the boundary, choose states h, ℓ with $\mu_h > \mu_O > \mu_\ell$, which exist under (G.1). A sufficiently small positive vector r, supported on these states, can have

$$\sum_i r_i < c, \qquad \mu_O < \frac{\sum_i \mu_i r_i}{\sum_i r_i} < \mu_C.$$

Indeed, choose its mean strictly between μ_O and $\min\{\mu_C, \mu_h\}$ and scale its coordinates downward. For every feasible d, the vectors $d^\varepsilon = (1-\varepsilon)d + \varepsilon r$ remain feasible, have strict outside-option constraint, and tend to d. Their exact implementations prove the lower bound in (G.34). The learning conclusions of Proposition G.1 apply separately to each implementing rule.

Attainment. If $B_\rho \geq A_\rho$, the second program provides an exact maximizer. If $A_\rho > B_\rho$ and the first program has an optimizer with strict outside-option constraint, $\tau = d$ attains the value. In the remaining case, $P_\rho = A_\rho > B_\rho \geq 0$. An attaining rule could not belong to the competitor-indifference regime, whose profit is at most B_ρ, or have zero sales. It also could not have $a = \mu_O$: then τ itself is feasible for (G.32), whereas sales are $\tau/2$, so profit is at most $A_\rho/2 < P_\rho$. Finally, a rule with $\mu_O < a \leq \mu_C$ would produce an optimizer of the first program with strict outside-option constraint, contrary to the hypothesis. This proves the asserted equivalence.

Randomization in at most two states. Apply Lemma G.12(i) to an exact maximizer or to each rule in a profit-approaching sequence. The replacement preserves prominence frequency, hence feasibility, and weakly raises profit. Its profit remains bounded by P_ρ, proving both assertions. The lemma applies to each fixed consumer mix as well. Sharpness, including with heterogeneous consumers, follows from Proposition G.11.

At $\rho = 0$, the first program has only $d = 0$. The second program requires $D \leq Jd_i$ for every i, hence all coordinates equal D/J; calibration and $\bar{\mu} < \mu_C$ force $D = 0$. Both values therefore agree with the unique feasible policy's zero profit, and that policy uses no randomization. □

Proposition G.10 (Welfare under a prominence cap)**.** *No feasible stationary rule raises limiting consumer welfare above the no-prominence benchmark μ_C. A welfare-neutral profit-maximizing rule exists if and only if $B_\rho \geq A_\rho$.*

Fix $\rho > 0$ with $P_\rho > 0$. Define

$$G_\rho := \max\left\{ \frac{1}{J}\sum_i (\mu_C - \mu_i) d_i : d \text{ is feasible for (G.32)}, \quad M(d) = P_\rho \right\}, \tag{G.36}$$

with $G_\rho = 0$ if the displayed set is empty. Every sequence of feasible stationary rules whose limiting profits tend to P_ρ satisfies

$$\liminf_n W^\infty(\tau^n) \geq \mu_C - G_\rho,$$

and a profit-approaching sequence randomizing in at most two states attains equality with convergent welfare. A welfare-neutral exact maximizer, whenever one exists, can also be chosen with at most two randomized states. Thus welfare converges to the no-prominence benchmark along every profit-approaching sequence if and only if $G_\rho = 0$. In particular, this holds when $B_\rho > A_\rho$, whereas $A_\rho > B_\rho$ implies a welfare loss bounded away from zero along every profit-approaching sequence.

Proof. With no prominence welfare is μ_C. With universal prominence it is μ_O, because $\bar{\mu} < \mu_O$. For $0 < T < 1$, apply Lemma G.2. If $a < \mu_O$, consumers choose the outside option under prominence and the competitor otherwise, giving welfare $\mu_C - T(\mu_C - \mu_O)$. The same expression holds at $a = \mu_O$,

because the conditional expected payoffs of brand purchases and the outside option then coincide. If $\mu_O < a \leq \mu_C$, sales equal $d = \tau$ and

$$W^{\infty}(\tau) = \mu_C - \frac{1}{J}\sum_i (\mu_C - \mu_i)d_i. \tag{G.37}$$

If $a > \mu_C$, consumers never take the outside option in the limit; calibration makes the purchase-weighted brand payoff equal to μ_C, so welfare is again μ_C. These formulas prove the upper bound. They also show that the welfare-neutral rules are exactly no prominence and the positive-frequency rules with $a \geq \mu_C$. Their sales are precisely allocations implemented by the second program. Consequently their maximal profit is B_ρ, proving the characterization of a neutral maximizer.

Now consider a profit-approaching sequence at a fixed cap with $P_\rho > 0$. Eventually it excludes all zero-sales rules and all rules with $a = \mu_O$: as in the preceding proof, the latter have profit at most $A_\rho/2 \leq P_\rho/2$. Every remaining rule is either welfare-neutral with sales feasible for the second program, or has $\mu_O < a \leq \mu_C$ and sales feasible for the first. In any subsequence of the latter type, compactness gives a further subsequence of sales converging to an allocation feasible for (G.32) with profit P_ρ. The continuous identity (G.37) then bounds its welfare limit below by $\mu_C - G_\rho$. Neutral subsequences satisfy the same bound because $G_\rho \geq 0$.

If the set defining G_ρ is nonempty, choose an allocation attaining that maximum and use the strict-constraint approximations from Proposition G.9. Their profits tend to P_ρ, and their welfare tends to $\mu_C - G_\rho$ by (G.37). If the set is empty, $A_\rho < P_\rho = B_\rho$, and a constant sequence of neutral maximizers attains the bound instead. Apply Lemma G.12(i) to each rule in the resulting sequence: its frequency and coarse welfare are unchanged, its profit weakly increases and still tends to P_ρ, and it randomizes in at most two states. The same argument applies to a neutral exact maximizer. This proves the stated sharpness and the if-and-only-if claim, using also $W^{\infty} \leq \mu_C$.

Finally, if $A_\rho > B_\rho$, no optimizer of the first program can have $\sum_i (\mu_C - \mu_i)d_i = 0$: such an allocation would also be feasible for the second program. Compactness of the first program's optimal set therefore gives a strictly positive minimum of this welfare gap. Rules from the second program cannot approach P_ρ, and the preceding subsequence argument bounds every profit-approaching sequence's welfare uniformly below μ_C. When $A_\rho = B_\rho$, a neutral exact maximizer exists, but the condition $G_\rho = 0$ remains necessary for preservation along all approximately optimal sequences. □

A positive cap can preserve welfare. The characterization does not imply that every positive cap fails in every environment. For example, take

$$J = 2, \quad (\mu_h, \mu_\ell) = (3/2, 0), \quad \mu_O = 1, \quad \mu_C = 6/5, \quad (m_h, m_\ell) = (1, 1/100), \quad \rho = 1/10.$$

The first program requires $d_h + d_\ell \leq 1/5$ and $d_\ell \geq d_h/4$. For a given total D, its objective is at most $[D - (99/100)(D/5)]/2$; equality is feasible at $(d_h, d_\ell) = (4/25, 1/25)$. Hence $A_\rho = 401/5000$. In the second program, calibration gives $d_\ell = d_h/4$, and the low-state cap inequality gives $d_h \leq 1/4$. The upper endpoint is feasible, so $B_\rho = 401/3200 > A_\rho$. It is implemented by $\tau = (1/5, 0)$, with unpromoted purchase probability $q = 1/16$. Thus this positive cap admits a neutral profit maximizer, and every profit-approaching sequence preserves welfare. Whether a cap protects welfare depends on the payoff and margin configuration, as well as its size.

G.8.1 A cap that reduces but does not eliminate the welfare loss

The two-program characterization permits an exact solution in a three-state example. Unlike the main-text illustration, both the medium and low products here yield less than the outside option.

Let the three states be equally likely, with

$$\begin{aligned}(\mu_h, \mu_m, \mu_\ell) = (1.202, 0.8, 0.2), \qquad \mu_O = 1, \quad \mu_C = 1.1,\\ (m_h, m_m, m_\ell) = (0.001, 0.001, 1).\end{aligned} \tag{G.38}$$

The platform faces the prominence-frequency constraint $T(\tau) = (\tau_h + \tau_m + \tau_\ell)/3 \leq \rho$. All policies below are stationary. Limiting profits and welfare first take the long-run limit at fixed sensitivity and then let sensitivity tend to infinity, as in Lemma G.2.

Proposition G.11 (The optimal response to a prominence cap)**.** *Under* (G.38)*, define*

$$\rho_* := \frac{167}{400}, \qquad r := \min\{\rho, \rho_*\}.$$

For every $\rho \in [0,1]$, the supremum of limiting platform profit is

$$V(\rho) := \sup_{\tau \in [0,1]^3 : T(\tau) \leq \rho} \Pi^\infty(\tau) = \frac{169}{835}\, r. \tag{G.39}$$

For $\rho > 0$, no fixed rule attains this supremum. Every feasible sequence whose profits approach it converges to the same policy vector,

$$\tau^*(\rho) = \left(\frac{400r}{167},\, 0,\, \frac{101r}{167}\right), \tag{G.40}$$

and its consumer welfare converges to

$$W^*(\rho) = \frac{11}{10} - \frac{r}{10} < \frac{11}{10}. \tag{G.41}$$

One implementing sequence is $\tau^\epsilon = (\tau_h^*, 0, (1-\epsilon)\tau_\ell^*)$ as $\epsilon \downarrow 0$. At $\rho = 0$, no prominence is the only feasible rule; it yields zero platform profit and welfare $11/10$.

If $0 < \rho < \rho_*$, every sequence approaching $V(\rho)$ eventually randomizes in both h and ℓ. Restricting randomization to at most one state therefore leaves a strictly positive profit gap.

For a fixed coarse share $\lambda \in (0,1]$, the population profit supremum is

$$V_\lambda(\rho) = \lambda \frac{169r}{835} + (1-\lambda)\frac{1}{3000}.$$

For $\rho > 0$, the same non-attainment, policy-limit and randomization conclusions hold. Relative to population welfare without prominence,

$$W_\lambda^0 = \lambda \frac{11}{10} + (1-\lambda)\frac{567}{500},$$

every profit-approaching sequence has limiting welfare

$$W_\lambda^0 - r\left[\frac{\lambda}{10} + (1-\lambda)\frac{101}{5010}\right] < W_\lambda^0.$$

At $\lambda = 0$, all stationary rules give the same limiting profit and the frequency tie-break selects no prominence.

Proof. Put $R := 3\rho$. Write (x, y, z) for state-specific brand purchase probabilities and

$$F(x, y, z) := \frac{x + y}{1000} + z$$

for three times platform profit. We solve the two programs in (G.32)–(G.33) and then check fixed-rule implementation and all approximately optimal sequences.

Purchases confined to promoted encounters. The first program requires

$$0 \le x, y, z \le 1, \qquad x + y + z \le R, \qquad D := \frac{101}{500}x - \frac{1}{5}y - \frac{4}{5}z \ge 0, \tag{G.42}$$

as well as the upper promoted-payoff bound $\frac{51}{500}x - \frac{3}{10}y - \frac{9}{10}z \le 0$. For the upper bounds that follow, the latter restriction can be omitted. Let

$$\alpha := \frac{169}{835}, \qquad \eta := \frac{333}{334}, \qquad \delta := \alpha + \frac{\eta}{5} - \frac{1}{1000} > 0.$$

Direct substitution gives the identity

$$F(x, y, z) = \alpha(x + y + z) - \eta D - \delta y \le \alpha R.$$

Also, $D \ge 0$ implies $z \le 101x/400 - y/4$, and therefore

$$F(x, y, z) \le \frac{507}{2000}x - \frac{249}{1000}y \le \frac{507}{2000}.$$

The smaller of these two bounds is attained in this linear program by

$$(x, y, z) = \left(\frac{400r}{167}, 0, \frac{101r}{167}\right).$$

This vector has purchase-weighted payoff exactly $\mu_O = 1$, so it also satisfies the omitted upper bound. Consequently the first program's value is

$$V_A(\rho) = \min\left\{\frac{169}{835}\rho, \frac{169}{2000}\right\} = \frac{169}{835}r. \tag{G.43}$$

Its optimizer is unique. When $0 < \rho \le \rho_*$, equality in the first bound requires $x + y + z = R$, $D = 0$, and $y = 0$, which determine the displayed vector. When $\rho \ge \rho_*$, equality in the second bound requires $x = 1$, $y = 0$, and $z = 101/400$. These conditions agree at $\rho = \rho_*$.

Purchases also possible without prominence. In the second program, calibration at $\mu_C = 1.1$ gives

$$\frac{51}{500}x = \frac{3}{10}y + \frac{9}{10}z, \qquad z = \frac{17}{150}x - \frac{1}{3}y, \qquad F(x, y, z) = \frac{343x - 997y}{3000}.$$

Recover the common nonpromoted purchase probability $q = \min\{x, y, z\}$ as in (G.35). Then $q \le y$ and $x + y + z \le R + (3 - R)q$. Since $0 \le R \le 3$, it follows that

$$\frac{167}{150}x \le R + \left(\frac{7}{3} - R\right)y.$$

Substituting this inequality into the formula for F yields

$$F \le \frac{343R}{3340} - \frac{46449 + 51450R}{501000}\,y \le \frac{343R}{3340}.$$

The bound $x \le 1$ also gives $F \le 343/3000$. The smaller of these two bounds is attained by taking $q = 0$ and

$$(x, y, z) = \left(u, 0, \frac{17u}{150}\right), \qquad u := \min\left\{\frac{450\rho}{167}, 1\right\}.$$

For $\rho > 0$, the corresponding prominence rule has promoted mean exactly μ_C, and Lemma G.2 implements these purchase probabilities. Thus the second program has value

$$V_B(\rho) = \min\left\{\frac{343}{3340}\rho, \frac{343}{9000}\right\} = \frac{343}{3340}\min\left\{\rho, \frac{167}{450}\right\}. \tag{G.44}$$

Because $343/3340 < 169/835$ and $167/450 < 167/400$, $V_B(\rho) < V_A(\rho)$ whenever $\rho > 0$.

Implementation and non-attainment. Fix $\rho > 0$. Reducing the low component of (G.40) gives the feasible rule $\tau^\epsilon = (\tau_h^*, 0, (1-\epsilon)\tau_\ell^*)$. For $0 < \epsilon < 167/303$, its promoted mean is strictly between μ_O and μ_C. Lemma G.2 therefore gives limiting sales $d = \tau^\epsilon$, and its profit tends to $V_A(\rho)$. This proves (G.39), either directly or by the general reduction (G.34).

For completeness, a fixed rule with promoted mean below μ_O has zero limiting sales. A fixed rule with promoted mean exactly μ_O has limiting sales $\tau/2$; its vector τ is feasible for the first program, so its profit is at most $V_A(\rho)/2$. A fixed rule with promoted mean at least μ_C has profit at most $V_B(\rho)$. The endpoint rules with $T = 0$ or $T = 1$ have zero brand sales. Any rule attaining $V_A(\rho)$ would consequently have promoted mean strictly between μ_O and μ_C, with sales equal to τ. Uniqueness in the first program would force $\tau = \tau^*(\rho)$, whose promoted mean is exactly μ_O, a contradiction. The supremum is therefore unattained.

Every approximately optimal sequence. Let τ^n be feasible with $\Pi^\infty(\tau^n) \to V_A(\rho)$. The strict bounds

$$\max\{V_B(\rho), V_A(\rho)/2, 0\} < V_A(\rho)$$

show that all sufficiently large n have promoted means strictly between μ_O and μ_C. Hence their sales equal τ^n, which are feasible vectors in the first program. Compactness of that feasible set and uniqueness of its optimizer imply $\tau^n \to \tau^*(\rho)$: otherwise a subsequence bounded away from the optimizer would have a convergent subsequence whose limit is a different optimizer.

For these policies consumers buy the brand under prominence and the competitor otherwise. Consequently

$$W^\infty(\tau^n) = \frac{1}{3}\sum_{i\in\{h,m,\ell\}} \left[\mu_i \tau_i^n + \mu_C(1 - \tau_i^n)\right].$$

At the limiting vector, $\sum_i \mu_i \tau_i^* = \mu_O \sum_i \tau_i^*$ and $\sum_i \tau_i^*/3 = r$. Taking limits gives $W^*(\rho) = \mu_C - (\mu_C - \mu_O)r$, proving (G.41). At $\rho = 0$ the only feasible rule is $\tau = 0$; since the unconditional brand mean is below μ_O, consumers choose the competitor in the high-sensitivity limit.

Sharpness of the two-state bound. Fix $0 < \rho < \rho_*$. Both $\tau_h^*(\rho) = 400\rho/167$ and $\tau_\ell^*(\rho) = 101\rho/167$ lie strictly between zero and one. The convergence $\tau^n \to \tau^*(\rho)$ just proved therefore forces both coordinates of every profit-approaching sequence to be interior for all sufficiently large n. Consequently,

$$\sup_{\substack{\tau\in[0,1]^3,\ T(\tau)\le\rho \\ |\{i:0<\tau_i<1\}|\le 1}} \Pi^\infty(\tau) < V(\rho).$$

If equality held, a sequence satisfying the restriction would approach $V(\rho)$, contradicting that eventual interiority. The displayed implementing sequence randomizes only in h and ℓ, establishing sharpness of the bound in Proposition G.9.

Heterogeneous consumers. Finest-partition consumers buy h with or without prominence and never buy m or ℓ in the high-sensitivity limit. Thus, for every rule,

$$\Pi_{\mathrm{f}}^{\infty}(\tau) = \frac{1}{3000}, \qquad W_{\mathrm{f}}^{\infty}(\tau) = \frac{\mu_h + 2\mu_C}{3} - \frac{\mu_C - \mu_O}{3}(\tau_m + \tau_\ell) = \frac{567}{500} - \frac{\tau_m + \tau_\ell}{30}.$$

For fixed $\lambda > 0$, population profit approaches its supremum if and only if coarse-consumer profit approaches $V(\rho)$. The policy limit, non-attainment and positive profit gap under one-state randomization therefore carry over directly; the latter gap is multiplied by λ. Substituting (G.40) into the fine welfare formula gives a loss of $101r/5010$. Combining it with the coarse welfare loss $r/10$ proves the population formula. When $\lambda = 0$, profit is constant across rules and no prominence uniquely minimizes frequency. □

For example, at $\rho = 1/3$ the limiting policy and the corresponding profit and welfare limits are

$$\tau^* = \left(\frac{400}{501}, 0, \frac{101}{501}\right), \qquad V(1/3) = \frac{169}{2505} \simeq 0.067465, \qquad W^*(1/3) = \frac{16}{15}.$$

The feasible rule $(0.8, 0, 0.2)$ yields profit approximately 0.066933 and welfare 1.0672, close to these limits. For caps below ρ_*, tightening the cap scales down the profitable combination of high- and low-quality prominence without changing its proportions. Welfare loss falls with the cap but remains positive at every positive cap. Above ρ_*, the cap does not reduce the profit supremum or its limiting welfare loss.

G.9 General policy reduction and bounds on randomization

The preceding randomization bounds follow from the number of policy moments that determine coarse feedback. Put

$$U(\tau) := \frac{1}{J}\sum_i \mu_i \tau_i.$$

For a coarse share $\lambda \in [0,1]$, let $\Pi_{\lambda,\beta}$ denote long-run population profit at finite sensitivity, and write $\Pi_{\lambda,\infty} := \Pi_{\lambda}^{\infty}$ for its fixed-rule high-sensitivity limit. At finite sensitivity the fine type's conditional purchase probabilities are

$$a_{i,\beta}^{\mathrm{f}} = A_\beta(\mu_i), \qquad b_{i,\beta}^{\mathrm{f}} = B_\beta(\mu_i);$$

at $\beta = \infty$ use (G.20). Thus fine-consumer profit is affine in the policy at every sensitivity considered here.

Lemma G.12 (Reduction of stationary prominence policies)**.** *Maintain* (G.1). *Fix a consumer share* $\lambda \in [0,1]$ *and an evaluation sensitivity* $\beta \in [0,\infty]$. *For every stationary rule* τ*:*

(i) There is a stationary rule $\widehat{\tau}$ *randomizing in at most two states such that*

$$T(\widehat{\tau}) = T(\tau), \qquad U(\widehat{\tau}) = U(\tau), \qquad \Pi_{\lambda,\beta}(\widehat{\tau}) \geq \Pi_{\lambda,\beta}(\tau).$$

The replacement preserves coarse consumers' deterministic mean dynamics at every finite sensitivity, and their equilibrium valuations, aggregate choice shares and welfare at every finite sensitivity and in the high-sensitivity limit. In the all-coarse case, one replacement can be chosen to weakly raise profit simultaneously at every sensitivity.

(ii) If prominence frequency need not be preserved, there is a stationary rule $\widetilde{\tau}$ *randomizing in at most one state with* $\Pi_{\lambda,\beta}(\widetilde{\tau}) \geq \Pi_{\lambda,\beta}(\tau)$.

Proof. We use an elementary fact about a box intersected with affine equalities. If the equalities have rank at most r, each vertex has at most r coordinates in $(0,1)$. Otherwise a nonzero perturbation supported on the interior coordinates lies in the null space of the equalities; small positive and negative multiples remain feasible, contradicting extremality. A linear objective on a nonempty compact polytope attains its maximum at a vertex.

(i) Preserving the two moments. Consider the nonempty compact polytope

$$\mathcal{P}(\tau) := \left\{ z \in [0,1]^J : \sum_i z_i = \sum_i \tau_i, \quad \sum_i \mu_i z_i = \sum_i \mu_i \tau_i \right\}. \tag{G.45}$$

For $0 < T < 1$, all its policies have the same menu probabilities and conditional brand means $a = U/T$ and $b = (\bar{\mu} - U)/(1-T)$. Consequently their reduced mean dynamics coincide. At $T = 0$ or $T = 1$ the polytope is a singleton and the claim is immediate. Uniqueness in Proposition G.1 gives the same coarse equilibrium valuation v_β for all $z \in \mathcal{P}(\tau)$. At the evaluation sensitivity write A, B for the corresponding coarse conditional purchase probabilities, using their limits when $\beta = \infty$. Population profit on the polytope is

$$\Pi_{\lambda,\beta}(z) = \frac{1}{J} \sum_i m_i \left\{ \lambda[B + (A-B)z_i] + (1-\lambda)[b^{\mathrm{f}}_{i,\beta} + (a^{\mathrm{f}}_{i,\beta} - b^{\mathrm{f}}_{i,\beta})z_i] \right\}. \tag{G.46}$$

It is affine in z, so an optimal vertex gives the profit inequality with at most two interior coordinates.

The total coarse brand share is $TA + (1-T)B$ and its contribution to welfare is $AU + B(\bar{\mu} - U)$. The competitor and outside-option shares depend only on T and the common equilibrium valuations, by (G.5). All aggregate shares and coarse welfare are therefore unchanged. This argument holds at every finite sensitivity; Lemma G.2 gives the same conclusion in the limit. If $\lambda = 1$, maximize $\sum_i m_i z_i$ on (G.45). Since $A_\beta - B_\beta > 0$ at finite sensitivity and its limit is nonnegative, a single optimal vertex weakly improves profit at all sensitivities.

(ii) One equality without a frequency restriction. First let $\beta < \infty$, fix the original equilibrium brand valuation v, and put $A = A_\beta(v)$ and $B = B_\beta(v)$. On the nonempty polytope

$$\left\{ z \in [0,1]^J : \sum_i (\mu_i - v)[B + (A-B)z_i] = 0 \right\}, \tag{G.47}$$

the same v satisfies the brand calibration equation. Its sampling denominator is positive. Uniqueness makes v the equilibrium valuation under every feasible rule, including $T = 0, 1$ when these are feasible. Formula (G.46) is again the actual population profit. Maximizing it at a vertex gives at most one interior coordinate.

Now let $\beta = \infty$. Endpoint rules with $T = 0$ or $T = 1$ are already deterministic, so retain them. For $0 < T < 1$ and promoted mean $a \le \mu_C$, define

$$k(a) = \begin{cases} 0, & a < \mu_O, \\ 1/2, & a = \mu_O, \\ 1, & \mu_O < a \le \mu_C. \end{cases}$$

On the polytope $\{z \in [0,1]^J : \sum_i (\mu_i - a) z_i = 0\}$, every non-endpoint rule has the same promoted mean and coarse sales $k(a)z_i$. This sales formula also holds at $z = 0$. The other endpoint $z = \mathbf{1}$ is feasible only if $a = \bar{\mu} < \mu_O$, when it too has zero coarse sales. Hence population profit is affine on the entire polytope, and an optimal vertex has at most one interior coordinate. In particular, the boundary probability $1/2$ at $a = \mu_O$ is preserved whenever the replacement uses prominence.

Finally, if $a > \mu_C$, put $H = \sum_i \tau_i(\mu_i - \mu_C) > 0$ and impose the single equality $\sum_i z_i(\mu_i - \mu_C) = H$. Neither endpoint is feasible. Every feasible policy has promoted mean above μ_C and the same non-promoted brand-purchase probability

$$q = \frac{H}{J(\mu_C - \bar{\mu}) + H} \in (0, 1).$$

Its coarse sales are $q + (1-q)z_i$, so population profit is affine once more. An optimal vertex completes the proof. □

Part (i) preserves expected feedback and the resulting mean dynamics, not necessarily the reward distributions or stochastic learning paths. It does not generally preserve finest-partition or population welfare. The profit-improving replacement may depend on consumer composition and the evaluation sensitivity, except for the simultaneous all-coarse conclusion. Part (ii) imposes no preservation of frequency or welfare.

Corollary G.13 (Randomization bounds for optimal policies)**.** *For each fixed consumer share and finite sensitivity, a profit-maximizing stationary rule exists. One can be chosen to randomize in at most one state without a cap, and at most two states under any cap. In the high-sensitivity limit the same bounds hold for attained optima and for sequences approaching a profit supremum. If all consumers use the finest partition, one randomized state suffices even under a cap. In the high-sensitivity limit, the capped two-state bound is sharp for every positive coarse share.*

Proof. At finite sensitivity, the equilibrium brand valuation is continuous in the policy. Indeed, each valuation lies in the compact interval $[\min_i \mu_i, \max_i \mu_i]$. For any convergent policy sequence, every subsequential valuation limit solves the limiting policy's continuous calibration equation, whose sampling denominator is positive and whose solution is unique. Thus all subsequential limits coincide with that policy's equilibrium valuation, including at $T = 0, 1$. Profit is therefore continuous on the policy cube. The cube and each capped subset are compact, so maxima exist. Apply part (ii) of the lemma without a cap and part (i) with a cap. At high sensitivity, apply the same parts to an attained optimizer or separately to each member of a profit-approaching sequence; no exchange of limits is involved.

For a pure finest-partition population, profit is affine in the policy. An optimum over $[0, 1]^J$ with $\sum_i \tau_i \le J\rho$ has at most one interior coordinate at a vertex: if the cap binds, perturb two interior coordinates in opposite directions; if it is slack, perturb one. This holds at finite sensitivity and in the high-sensitivity limit. If an attained optimizer is selected by the frequency tie-break, start from that rule and, where frequency is not already fixed, minimize it on the relevant optimal face before choosing a vertex. The replacement cannot have a lower frequency than the selected optimizer and therefore preserves the tie-break. Finally, Proposition G.11 proves high-sensitivity sharpness for every fixed positive coarse share. □

G.10 Optimal costly stationary prominence

This subsection retains the all-coarse, uncapped specification and (G.1), but charges $c \ge 0$ whenever prominence occurs, irrespective of the ensuing purchase. The platform maximizes

$$N_c(\tau) := \Pi^\infty(\tau) - cT(\tau).$$

Consumer rewards and choice menus are unchanged, so Proposition G.1 and Lemma G.2 continue to apply to each fixed rule. The realized prominence frequency converges almost surely to $T(\tau)$ by the

strong law for the independent state and prominence draws. Consequently subtracting $cT(\tau)$ also gives the limit of realized average net profit, with the same order of limits as above. If a tie-break is needed, lower prominence frequency is preferred among rules with equal net profit.

For $d \in [0,1]^J$, put

$$L_O(d) := \sum_i (\mu_i - \mu_O) d_i, \qquad L_C(d) := \sum_i (\mu_i - \mu_C) d_i, \qquad M(d) := \frac{1}{J} \sum_i m_i d_i,$$

and define

$$\mathcal{D} := \{d \in [0,1]^J : L_O(d) \geq 0,\ L_C(d) \leq 0\}, \qquad \mathcal{E} := \{d \in \mathcal{D} : L_C(d) = 0\}. \tag{G.48}$$

Writing $\bar{d} = J^{-1} \sum_i d_i$ and $\underline{d} = \min_i d_i$, set

$$T_{\min}(d) := \begin{cases} \bar{d}, & L_C(d) < 0, \\ \dfrac{\bar{d} - \underline{d}}{1 - \underline{d}}, & L_C(d) = 0. \end{cases} \tag{G.49}$$

The denominator is positive: $\mathbf{1} \notin \mathcal{D}$, since $L_O(\mathbf{1}) = J(\bar{\mu} - \mu_O) < 0$.

Lemma G.14 (Feasible sales and least prominence frequency)**.** *The closure of the limiting sales vectors of fixed stationary rules is $\mathcal{D}$. For each $d \in \mathcal{D}$,*

$$T_{\min}(d) = \inf_{\{\tau^n:\ d(\tau^n) \to d\}} \liminf_{n \to \infty} T(\tau^n), \tag{G.50}$$

where $d(\tau)$ denotes the fixed-rule high-sensitivity sales vector. The infimum is achieved by a constant sequence if $d = 0$ or $L_O(d) > 0$. In the latter case a least-frequency implementation is

$$\tau_i = \frac{d_i - q}{1 - q}, \qquad q = \begin{cases} 0, & L_C(d) < 0, \\ \underline{d}, & L_C(d) = 0. \end{cases} \tag{G.51}$$

For nonzero d with $L_O(d) = 0$, the infimum is approached by rules whose promoted means lie strictly between μ_O and μ_C.

Proof. For a rule with zero limiting sales, $d = 0 \in \mathcal{D}$ and $T_{\min}(d) = 0 \leq T(\tau)$. Otherwise $0 < T < 1$, and Lemma G.2 gives three possibilities. If $a = \mu_O$, then $d = \tau/2$, so $L_O(d) = 0$, $L_C(d) < 0$, and $T_{\min}(d) = \bar{d} = T/2 \leq T$. If $\mu_O < a \leq \mu_C$, then $d = \tau$ and $d \in \mathcal{D}$; for $a < \mu_C$ one has $T_{\min}(d) = T$. If $a > \mu_C$, then $d_i = q + (1-q)\tau_i$ for $q = q_\tau \in (0,1)$, and calibration gives $L_C(d) = 0$. Both cases with $L_C(d) = 0$ have the same representation, with $q = 0$ when $a = \mu_C$. Necessarily $0 \leq q \leq \underline{d}$ and

$$T = \frac{\bar{d} - q}{1 - q} \geq \frac{\bar{d} - \underline{d}}{1 - \underline{d}} = T_{\min}(d),$$

because the displayed fraction is decreasing in q when $\bar{d} < 1$. Thus every actual sales vector belongs to $\mathcal{D}$ and uses at least $T_{\min}(d)$ prominence.

Conversely, implement $d = 0$ by no prominence. If $L_O(d) > 0$ and $L_C(d) < 0$, take $\tau = d$. It has $0 < T < 1$, and its promoted mean lies strictly between μ_O and μ_C. Hence its sales are d and its frequency is $\bar{d}$. If $d \in \mathcal{E} \setminus \{0\}$, put $q = \underline{d}$ and use (G.51). A positive constant vector cannot belong to $\mathcal{E}$, because $\bar{\mu} < \mu_C$. Thus this policy has $0 < T < 1$. Writing $\Delta := J(\mu_C - \bar{\mu}) > 0$, one obtains

$$\sum_i (\mu_i - \mu_C)\tau_i = \frac{q\Delta}{1 - q}.$$

If $q = 0$, its promoted mean is μ_C and its sales are $\tau = d$. If $q > 0$, its promoted mean exceeds μ_C, and (G.8) gives competitive purchase probability q. Its sales are again d, with frequency $T_{\min}(d)$. Notice that $L_O(d) = (\mu_C - \mu_O)\sum_i d_i > 0$ on $\mathcal{E} \setminus \{0\}$.

It remains to approximate a nonzero d with $L_O(d) = 0$. Choose a state h with $\mu_h > \mu_O$ and put $d^\varepsilon = (1-\varepsilon)d + \varepsilon e_h$, where e_h is its coordinate unit vector. Then

$$L_O(d^\varepsilon) = \varepsilon(\mu_h - \mu_O) > 0, \qquad L_C(d) = -(\mu_C - \mu_O)\sum_i d_i < 0.$$

For all sufficiently small positive ε, $L_C(d^\varepsilon) < 0$. The policies $\tau^\varepsilon = d^\varepsilon$ therefore generate sales $d^\varepsilon \to d$, with $T(\tau^\varepsilon) \to \bar{d} = T_{\min}(d)$. This proves that $\mathcal{D}$ is exactly the closure of feasible sales and constructs sequences attaining the right side of (G.50).

For its reverse inequality, $T_{\min}$ is lower semicontinuous on $\mathcal{D}$. Indeed, $\underline{d}$ is bounded away from one on this compact set. The expression on $\mathcal{E}$ is continuous and does not exceed $\bar{d}$, while outside $\mathcal{E}$ the function equals the continuous function $\bar{d}$. At $\mathcal{E}$ the only possible jump is therefore downward. For any sequence with $d(\tau^n) \to d$, the already established frequency bound gives

$$\liminf_n T(\tau^n) \geq \liminf_n T_{\min}(d(\tau^n)) \geq T_{\min}(d).$$

This proves (G.50). □

Proposition G.15 (Optimal costly stationary prominence)**.** *Under* (G.1)*, with* $m_i \geq 0$ *and* $c \geq 0$*,*

$$V_c := \sup_{\tau \in [0,1]^J} N_c(\tau) = \max_{d \in \mathcal{D}} \{M(d) - cT_{\min}(d)\}. \tag{G.52}$$

If $V_c > 0$*,* V_c *is attained by a stationary rule if and only if some maximizing sales allocation has* $L_O(d) > 0$*. Each such allocation is implemented at value* V_c *by* (G.51)*. If* $V_c = 0$*, no prominence is optimal. An optimal rule, when one exists, or a sequence approaching* V_c *can be chosen to randomize in at most one state.*

Proof. The set $\mathcal{D}$ is nonempty and compact, and $M - cT_{\min}$ is upper semicontinuous by Lemma G.14 and $c \geq 0$. It therefore has a maximum, denoted $F_c^* \geq 0$. Every rule satisfies

$$N_c(\tau) = M(d(\tau)) - cT(\tau) \leq M(d(\tau)) - cT_{\min}(d(\tau)) \leq F_c^*.$$

For any maximizing allocation, the implementation or approximation constructed in that lemma has sales converging to d and frequency converging to $T_{\min}(d)$. Its net profit converges to F_c^*. This proves (G.52). The approximation is over fixed policies after their respective long-run and high-sensitivity limits; no interchange of limits is used.

Attainment. A maximizing d with $L_O(d) > 0$ has an exact least-frequency implementation, which attains V_c. If $V_c = 0$, the zero policy attains it. Conversely, suppose a rule attains a positive V_c. It has positive sales and $0 < T < 1$, so its promoted mean is at least μ_O. It cannot have $a = \mu_O$: in that case its actual sales are $\tau/2$, whereas the vector τ itself belongs to $\mathcal{D}$, satisfies $L_C(\tau) < 0$, and obeys

$$M(\tau) - cT_{\min}(\tau) = M(\tau) - cT(\tau) = N_c(\tau) + \frac{1}{2}M(\tau) > V_c.$$

The last inequality follows from $N_c(\tau) > 0$, which implies $M(\tau) > 0$, and contradicts (G.52). Thus $a > \mu_O$ and its actual sales d have $L_O(d) > 0$. The upper-bound inequalities above must both be equalities, so d is a maximizing allocation. This proves the asserted equivalence, including cases

with zero margins or zero cost. In particular, the exact half-purchase rule at $a = \mu_O$ cannot attain a positive supremum, even though boundary allocations may approximate that value.

One-state randomization. We show more generally that every fixed rule admits a rule randomizing in at most one state with weakly higher net profit. A rule with zero sales has $N_c \leq 0$ and can be replaced by no prominence. For $0 < T < 1$ and $\mu_O \leq a \leq \mu_C$, put $k = 1/2$ if $a = \mu_O$ and $k = 1$ otherwise. On the nonempty compact polytope

$$\mathcal{P}_a := \left\{ z \in [0,1]^J : \sum_i (\mu_i - a) z_i = 0 \right\},$$

every nonzero rule has the same promoted mean a, and hence sales kz. The rule $z = \mathbf{1}$ is infeasible since $a \geq \mu_O > \bar{\mu}$; the zero rule also has sales $kz = 0$. Thus actual net profit throughout $\mathcal{P}_a$ is the linear function $J^{-1} \sum_i (km_i - c) z_i$. It attains its maximum at a vertex.

If instead $a > \mu_C$, hold fixed $H := \sum_i (\mu_i - \mu_C)\tau_i > 0$ and maximize over

$$\mathcal{P}_H := \left\{ z \in [0,1]^J : \sum_i (\mu_i - \mu_C) z_i = H \right\}.$$

Neither endpoint rule belongs to this polytope. Every feasible rule has promoted mean above μ_C and the same competitive purchase probability $q = H/(\Delta + H)$, where $\Delta = J(\mu_C - \bar{\mu}) > 0$. Its actual net profit is affine:

$$\frac{q}{J} \sum_i m_i + \frac{1}{J} \sum_i \big((1-q)m_i - c\big) z_i.$$

Again choose an optimal vertex. A box intersected with one affine equality has at most one coordinate in $(0,1)$ at each vertex: two interior coordinates would permit a nonzero small two-sided perturbation preserving the equality, contradicting extremality. This is the same reduction as in the proof of Lemma G.12, with the affine cost subtracted from the objective. Apply it to an attained optimizer or separately to each rule in a sequence approaching V_c. If an attained optimizer is selected by the frequency tie-break, minimize frequency on the relevant optimal face before choosing a vertex; the same bound holds for a selected optimizer. □

Competitive mixing, deterministic implementation, and welfare. For any nonzero $d \in \mathcal{E}$, the least-frequency implementation has $q = \underline{d}$. If $q > 0$ and every d_i is either q or one, (G.51) is deterministic: it promotes exactly the states with $d_i = 1$. If such a d maximizes (G.52), this deterministic policy is optimal even among unrestricted stationary randomized rules. More generally, every rule sustaining valuation μ_C has no limiting outside-option purchases, and its brand sales satisfy $L_C(d) = 0$. Consequently

$$W^\infty = \frac{1}{J} \sum_i \mu_i d_i + \mu_C (1 - \bar{d}) = \mu_C. \tag{G.53}$$

This equals no-prominence welfare under (G.1). Costly prominence can therefore select a welfare-preserving policy through competitive mixing. Neither deterministic optimality nor monotonicity of optimal consumer welfare in c is asserted for general primitives.

G.10.1 The two-state costly-prominence example

Proposition G.16 (The exact cost range for high-only prominence)**.** *Suppose the two states are equally likely and*

$$(\mu_h, \mu_\ell) = (4, 0), \qquad \mu_O = \frac{5}{2}, \qquad \mu_C = 3, \qquad m_h = m_\ell = 1. \tag{G.54}$$

High-only prominence is optimal among all stationary policies if and only if $c \in [4/9, 8/9]$. *Throughout this closed interval it is the unique optimal policy; it induces sales* $(1, 1/3)$, *competitive purchase probability* $1/3$, *net profit* $2/3 - c/2$, *and consumer welfare* 3. *At* $c = 2/3$ *its net profit is* $1/3$. *The policy* $(1, 1/3)$ *generates the same sales but net profit* $2/9$. *The costless-optimal sequence* $(1, t)$, *with* $t \uparrow 3/5$, *has net profit tending to* $4/15$ *at* $c = 2/3$ *and consumer welfare tending to* $13/5$.

Proof. Write a rule as $\tau = (x, y) \in [0, 1]^2$ and put

$$N_c(x, y) = \Pi^\infty(x, y) - \frac{c}{2}(x + y).$$

No prominence yields zero net profit, whereas universal prominence yields $-c$, since $\bar{\mu} = 2 < \mu_O$. For every other nonzero rule the promoted mean is $a = 4x/(x + y)$. We apply Lemma G.2 to each purchase regime.

First let $a > \mu_C$, or equivalently $x > 3y$. The competitive purchase probability and state-specific sales are

$$q = \frac{x - 3y}{2 + x - 3y}, \qquad d_h = \frac{3(x - y)}{2 + x - 3y}, \qquad d_\ell = \frac{x - y}{2 + x - 3y}.$$

Set $z = (x - y)/(1 - y) \in (0, 1]$. The rule $(z, 0)$ generates exactly the same sales. If $y > 0$, it uses strictly less prominence, because

$$x + y - z = \frac{y(2 - x - y)}{1 - y} > 0.$$

For $c > 0$, any competitive-mixing rule with $y > 0$ is therefore strictly dominated by $(z, 0)$. The latter rule has

$$\begin{aligned} N_c(z, 0) &= \frac{2z}{2 + z} - \frac{c}{2}z, \\ \frac{d}{dz}N_c(z, 0) &= \frac{4}{(2 + z)^2} - \frac{c}{2}, \qquad & \frac{d^2}{dz^2}N_c(z, 0) &= -\frac{8}{(2 + z)^3} < 0. \end{aligned} \tag{G.55}$$

Thus $(1, 0)$ uniquely maximizes net profit within the competitive-mixing regime for $0 < c \leq 8/9$, with value $2/3 - c/2$.

Next let $\mu_O < a \leq \mu_C$, or equivalently $x/3 \leq y < 3x/5$. Purchases occur only under prominence, and for $c < 1$,

$$N_c(x, y) = \frac{1 - c}{2}(x + y) < \frac{4}{5}(1 - c). \tag{G.56}$$

The bound is approached by $(1, t)$ with $t \uparrow 3/5$. At $a = \mu_O$, one has $y = 3x/5$ and $T = (x+y)/2 \leq 4/5$. The promoted purchase probability is $1/2$, so

$$N_c(x, y) = (1/2 - c)T < (1 - c)T \leq \frac{4}{5}(1 - c) \qquad (c < 1).$$

If $a < \mu_O$, sales vanish and net profit is nonpositive.

For $c \in [4/9, 8/9]$, high-only net profit is positive and

$$N_c(1, 0) - \frac{4}{5}(1 - c) = \frac{9c - 4}{30} \geq 0. \tag{G.57}$$

The strict bounds in the nonmixing regimes and uniqueness in the competitive-mixing regime prove global uniqueness. At $c = 4/9$, the competing supremum in (G.56) is not attained. At $c = 8/9$, the

derivative in (G.55) is zero only at $z = 1$ and strictly positive below it. Hence both endpoints belong to the range of unique optimality.

Conversely, if $0 \leq c < 4/9$, the sequence $(1, t)$ with $t \uparrow 3/5$ has net profit tending to a value strictly above $N_c(1, 0)$, and some member strictly improves on high-only prominence. If $c > 8/9$, the derivative of $N_c(z, 0)$ at $z = 1$ is negative, so reducing high-state prominence slightly strictly raises net profit. This proves the if-and-only-if cost range.

For $(1, 0)$, $q = 1/3$, sales are $(1, 1/3)$, gross profit is $2/3$, and $T = 1/2$. At $c = 2/3$, net profit is therefore $1/3$. The rule $(1, 1/3)$ has promoted mean 3, the same sales, and $T = 2/3$, giving net profit $2/9$. Under either rule welfare is

$$W^\infty = \frac{1}{2}\,4 + \frac{1}{2}\left(\frac{1}{3}\,0 + \frac{2}{3}\,3\right) = 3,$$

also following from (G.53). Without prominence, the brand valuation is $2 < \mu_O$, so all limiting purchases go to the competitor and welfare is also 3.

Finally, for $1/3 < t < 3/5$, the rule $(1, t)$ has promoted mean strictly between μ_O and μ_C. Its gross profit and prominence frequency both equal $(1 + t)/2$, while

$$W^\infty(1, t) = 2 + \frac{3}{2}(1 - t).$$

At $c = 2/3$, its net profit tends to $4/15$ and its welfare to $13/5$. At $c = 0$, its gross profit tends to the global supremum $4/5$: direct-purchase rules have profit below $4/5$, competitive-mixing rules have profit at most $2/3$ by the same-sales replacement and (G.55), exact outside-option-boundary rules have profit at most $2/5$, and all other rules have zero sales. □

G.10.2 Costly prominence in the three-state illustration

The next result revisits the three-state illustration with equal margins. Costs can change which products receive prominence, while making a deterministic rule uniquely optimal even among randomized rules.

Proposition G.17 (High-only prominence with three states)**.** *Suppose* (G.1) *and* (G.11) *hold, and* $m_h = m_m = m_\ell = m > 0$. *Let* q_3 *be given by* (G.12). *Then* $0 < q_3 < 1/2$, *and high-only prominence* $(1, 0, 0)$ *uniquely maximizes* N_c *over all stationary rules* $\tau \in [0, 1]^3$ *whenever*

$$m \leq c \leq \bar{c}_3 := m(1 - q_3)(1 + 2q_3). \tag{G.58}$$

Throughout this closed interval, the optimal rule induces sales $(1, q_3, q_3)$, *brand valuation* μ_C, *and*

$$N_c(1, 0, 0) = \frac{m(1 + 2q_3) - c}{3}, \qquad W^\infty(1, 0, 0) = \mu_C. \tag{G.59}$$

If $c > \bar{c}_3$, *high-only prominence is not optimal.*

The interval in (G.58) has positive length and includes $c = m$, at which optimal net profit is $2mq_3/3 > 0$. Its lower bound is sufficient, not asserted to be necessary. In contrast, when prominence is costless, Proposition G.4 selects high-and-medium prominence uniquely among deterministic rules, since $q_3 < 1/2$. Thus costs can make concentrating prominence on the best product optimal even when all states have the same margin.

Proof. Write

$$a_h := \mu_h - \mu_C, \qquad a_m := \mu_C - \mu_m, \qquad a_\ell := \mu_C - \mu_\ell, \qquad \Delta := a_m + a_\ell - a_h = 3(\mu_C - \bar{\mu}) > 0.$$

The three-state assumptions imply $0 < a_h \leq a_m < a_\ell$, so $q_3 = a_h/(a_m + a_\ell) < 1/2$. In particular, $\bar{c}_3 - m = mq_3(1 - 2q_3) > 0$.

Fix $c \geq m$. For a stationary rule τ, let $S := \tau_h + \tau_m + \tau_\ell = 3T(\tau)$ and

$$H := a_h\tau_h - a_m\tau_m - a_\ell\tau_\ell.$$

First suppose $H > 0$. This is precisely the competitive-mixing regime: the promoted mean exceeds μ_C, and neither no prominence nor universal prominence belongs to this regime. By (G.8), its competitive purchase probability is $q(H) = H/(\Delta + H)$. Since promoted purchases occur with probability one, sales satisfy $d_i = q(H) + (1 - q(H))\tau_i$, giving

$$3N_c(\tau) = 3mq(H) + \big[m(1 - q(H)) - c\big]S. \tag{G.60}$$

The coefficient of S is strictly negative, including when $c = m$. For a fixed H, therefore, net profit is maximized by minimizing prominence frequency. Because $0 < H \leq a_h$, the rule $\widehat{\tau} = (H/a_h, 0, 0)$ is feasible, has the same H, and satisfies

$$S - \frac{H}{a_h} = \left(1 + \frac{a_m}{a_h}\right)\tau_m + \left(1 + \frac{a_\ell}{a_h}\right)\tau_\ell \geq 0,$$

with equality only if $\tau_m = \tau_\ell = 0$. Thus every rule in this regime that promotes either lower-quality product is strictly dominated by a rule promoting only the high-quality product.

It remains to choose its prominence probability $x \in (0, 1]$. Writing $n_c(x) := N_c(x, 0, 0)$ and extending continuously to $n_c(0) = 0$,

$$\begin{aligned} 3n_c(x) &= \frac{m(\Delta + 3a_h)x}{\Delta + a_h x} - cx, \\ 3n_c'(x) &= \frac{m\Delta(\Delta + 3a_h)}{(\Delta + a_h x)^2} - c, \\ 3n_c''(x) &= -\frac{2ma_h\Delta(\Delta + 3a_h)}{(\Delta + a_h x)^3} < 0. \end{aligned} \tag{G.61}$$

Since $a_h/(\Delta + a_h) = q_3$,

$$3n_c'(1) = m(1 - q_3)(1 + 2q_3) - c = \bar{c}_3 - c.$$

Strict concavity implies $n_c'(x) > n_c'(1) \geq 0$ for $x < 1$ whenever $c \leq \bar{c}_3$. Hence $x = 1$ is the unique maximizer throughout (G.58), including its upper endpoint.

To verify global optimality, consider the remaining purchase regimes. For $0 < T < 1$ and promoted mean at most μ_C, Lemma G.2 gives sales 0, $\tau/2$, or τ. Gross profit is consequently at most mT, and $N_c \leq (m - c)T \leq 0$. This also covers the exact outside-option boundary. No prominence has net profit zero, while universal prominence has zero limiting sales and net profit $-c$, since $\bar{\mu} < \mu_O$. High-only net profit, by contrast, is strictly positive over the entire stated interval: at its upper endpoint,

$$3n_{\bar{c}_3}(1) = m(1 + 2q_3) - \bar{c}_3 = mq_3(1 + 2q_3) > 0.$$

The preceding strict comparisons establish uniqueness among all stationary rules without invoking the frequency tie-break. If $c > \bar{c}_3$, then $n_c'(1) < 0$, so reducing x slightly below one strictly improves net profit. This proves the last assertion.

Finally, Lemma G.2 gives the high-only sales $(1, q_3, q_3)$ and brand valuation μ_C. Its frequency is $1/3$, yielding the net profit in (G.59). Equation (G.53) gives welfare μ_C, equal to its no-prominence level. The costless deterministic comparison follows from $3(\Pi^{\infty}_{hm} - \Pi^{\infty}_{h}) = m(1 - 2q_3) > 0$ and Proposition G.4. □

G.11 The comparative prediction about experience

The main-text prediction concerns the direction of the learning response, not an additional equilibrium-selection claim. Conditional on the same pre-update valuation, visit count, and selected class H, two realized rewards $r' > r$ produce next valuations differing by $(r' - r)/(\kappa_H + N(H) + 1) > 0$. Holding rival valuations and the subsequent menu fixed, the logit probability of choosing H is increasing in this valuation (strictly for $\beta > 0$). Thus a more favorable brand experience raises its next choice probability in a competitive encounter as well as in a promoted one. The policy and equilibrium results above establish how stationary prominence can sustain this transmission across product states; they do not assert that every stochastic sample path is monotone after an experience.

H Bayesian learning and coarse misspecification

This section connects an argmax counterpart of harmonic coarse learning to Fudenberg, Lanzani, and Strack (2021), hereafter FLS. We impose finite reward support and singleton support, but allow different reward supports across class–menu pairs. The resulting Bayesian embedding is exact along every realized history. Its stability conclusions concern argmax learning and suitably concentrated initial beliefs; they are distinct from the fixed-finite-sensitivity logit results proved above.

H.1 A finite-outcome Bayesian embedding

Let $Y \subset \mathbb{R}$ be a finite reward set and let $\mathsf{R}_{s\omega} \in \Delta(Y)$ be the conditional reward law after selecting $s \in \omega$, obtained by integrating out the fixed within-class implementation. Conditional reward draws have these stationary laws and are independent across dates. As elsewhere, Ω^+ is the finite set of supported class menus, $p(\omega) > 0$ for $\omega \in \Omega^+$, and menus are iid. Throughout this section assume

$$p(\{s\}) > 0 \qquad (s \in \mathcal{S}). \tag{H.1}$$

The decision maker's subjective parameter is $q = (q_s)_{s\in\mathcal{S}} \in Q := \prod_s \Delta(Y)$. She regards rewards from class s as iid with law q_s, independently of the realized menu, and has a product-Dirichlet prior

$$\nu = \bigotimes_{s\in\mathcal{S}} \mathrm{Dirichlet}((a_{sy})_{y\in Y}), \qquad a_{sy} > 0. \tag{H.2}$$

Put $\kappa_s := \sum_y a_{sy}$ and $v_0(s) := \kappa_s^{-1} \sum_y y a_{sy}$.[1] If $N_k(s, y)$ counts observations of reward y from class s strictly before date k, the posterior parameter is $(a_{sy} + N_k(s, y))_y$, and its mean payoff is

$$v_k(s) = \frac{\kappa_s v_0(s) + \sum_{t<k} \mathbf{1}\{s_t = s\} r_t}{\kappa_s + N_k(s)}. \tag{H.3}$$

Thus Bayesian updating coincides exactly with the maintained harmonic recursion. Here choices maximize posterior mean payoff in the current menu, with a fixed measurable tie-breaking rule,

[1] Any finite initial valuations and positive pseudo-counts can be represented after enlarging Y, if necessary, by two hypothetical rewards strictly bracketing those valuations and the actual reward support. Assigning these additional rewards zero objective probability leaves the observed process unchanged and is allowed by the weaker support condition used below.

rather than using logit. To embed this process in the finite-action model of FLS, define the finite set of feasible menu plans

$$\Phi := \{\phi : \Omega^+ \to \mathcal{S} : \phi(\omega) \in \omega \text{ for all } \omega\}.$$

A plan is an auxiliary action, and its observed outcome is only the realized pair (ω, y). Fix any full-support subjective menu law $\widetilde{p} \in \Delta(\Omega^+)$. The objective and subjective outcome laws associated with plan ϕ are

$$P^*_\phi(\omega, y) = p(\omega)\mathsf{R}_{\phi(\omega),\omega}(y), \qquad P^q_\phi(\omega, y) = \widetilde{p}(\omega)q_{\phi(\omega)}(y), \tag{H.4}$$

and utility is y. Because every $\widetilde{p}(\omega)$ is positive, maximizing a plan's expected payoff is equivalent to maximizing the posterior mean separately in each menu. Selecting the counterfactual plan before the menu is drawn therefore implements the same myopic choices as selecting after observing it. The likelihood factors $\widetilde{p}(\omega)$ cancel when updating q, so neither choices nor class valuations require knowledge of the true menu probabilities. Plans are an analytical device: only one distribution per class is learned, and no counterfactual rewards are observed.

For any feasible behavioral rule $\sigma = (\sigma(s \mid \omega))$, define

$$\begin{aligned} D_s(\sigma) &:= \sum_{\omega \ni s} p(\omega)\sigma(s \mid \omega), & q^\sigma_s &:= \frac{\sum_{\omega \ni s} p(\omega)\sigma(s \mid \omega)\mathsf{R}_{s\omega}}{D_s(\sigma)}, \\ v^\sigma_s &:= \sum_{y \in Y} y q^\sigma_s(y). \end{aligned} \tag{H.5}$$

Singleton support gives $D_s(\sigma) \geq p(\{s\}) > 0$. The pair (v^σ, σ) is a VE exactly when

$$\operatorname{supp} \sigma(\cdot \mid \omega) \subseteq \arg\max_{s \in \omega} v^\sigma_s \qquad (\omega \in \Omega^+). \tag{H.6}$$

For a pure plan ϕ, write q^ϕ and v^ϕ for these objects under its deterministic behavioral rule. Call it *menu-wise strict* when $v^\phi_{\phi(\omega)} > v^\phi_s$ for every supported menu and every available $s \neq \phi(\omega)$. No ordering between winners that never compete is needed here; in particular, the strict VE definition used earlier implies the strictness required below.

Proposition H.1 (Argmax learning and Berk–Nash equilibrium)**.** *Under the finite-outcome assumptions and* (H.1)–(H.2)*:*

(i) A behavioral rule σ satisfies (H.6) *if and only if any distribution over feasible plans with menu-wise marginals σ is a mixed Berk–Nash equilibrium of the embedded problem. Every such equilibrium is* strongly uniform*: every plan in its support is optimal at every KL-minimizing model. In particular, any unperturbed argmax plan that is eventually followed with positive probability must be a VE plan.*

(ii) If ϕ is a menu-wise strict VE, then for every $\epsilon \in (0, 1)$ there is $\delta > 0$ such that every product-Dirichlet prior of the form (H.2) *satisfying*

$$\nu\Big(\max_s \|q_s - q^\phi_s\|_\infty < \delta\Big) > 1 - \delta \tag{H.7}$$

induces, with probability greater than $1-\epsilon$, eventual adherence to ϕ in every menu and $v_k \to v^\phi$.

(iii) Let (v^σ, σ) be any VE. For every $\tau > 0$ there is a finite iid recommendation signal and a known payoff bonus taking values in $\{0, \tau\}$ such that the following holds. For every $\epsilon \in (0, 1)$

there is $\delta > 0$ for which every prior (H.2) satisfying (H.7) with q^σ in place of q^ϕ induces, with probability greater than $1-\epsilon$, convergence $v_k \to v^\sigma$ and

$$\frac{\sum_{k<K} \mathbf{1}\{\omega_k = \omega,\ s_k = s\}}{\sum_{k<K} \mathbf{1}\{\omega_k = \omega\}} \longrightarrow \sigma(s \mid \omega) \quad \textit{for every } \omega \in \Omega^+ \textit{ and } s \in \omega. \tag{H.8}$$

Valuations in this assertion continue to update the base reward y; the known bonus is separately included in choice utility.

The concentration neighborhoods in (ii) and (iii) contain positive product-Dirichlet priors. In (iii), the required neighborhood may depend on the perturbation size τ.

Proof. Regularity and exact updating. For a realized history consistent with the fixed policy, its likelihood as a function of q is

$$\prod_{t<k} \widetilde{p}(\omega_t) q_{s_t}(r_t).$$

Its factors involving q separate by class, giving Dirichlet conjugacy and (H.3). The subjective model family is the image $\Theta = \{(P^q_\phi)_{\phi\in\Phi} : q \in Q\}$, with the prior induced by (H.2). Singleton menus make the parameterization injective, since every plan chooses s in $\{s\}$ and $\widetilde{p}(\{s\}) > 0$.

We verify the weaker support condition stated in Section B.2 of the FLS supplemental material: every model in the prior support can be approximated by models assigning positive probability to every objectively possible outcome. For $\eta > 0$, replace every q_s by $(1-\eta)q_s + \eta\,\mathrm{Unif}(Y)$. These models belong to Θ, converge to the original model, and assign positive probability to every pair (ω, y). This is FLS Assumption 1(i′). It permits zero probabilities in the original model and in the objective reward laws. As established there, the results used here remain valid under this condition; the limit-point theorem is invoked only for the myopic case, with economic discount factor zero.

For completeness, the Dirichlet prior also satisfies the required subexponential thickness condition. Let $m = |Y| \geq 2$ and $0 < r \leq 1$. For every $q^0_s \in \Delta(Y)$, a relative simplex box around $(1-r/2)q^0_s + (r/2)\,\mathrm{Unif}(Y)$ has volume at least $C_m r^{m-1}$, lies within sup-norm distance r of q^0_s, and has all coordinates at least $r/(4m)$; the constants can be chosen uniformly in q^0_s. On this box the Dirichlet density is bounded below by a positive constant times $r^{\sum_y (a_{sy}-1)_+}$. Consequently the product prior assigns at least Cr^d to every radius-r parameter ball, where

$$d = |\mathcal{S}|(m-1) + \sum_{s,y} (a_{sy} - 1)_+.$$

The map $q \mapsto (P^q_\phi)_\phi$ is 1-Lipschitz in the sup norm, so model-law balls inherit this lower bound. In particular, $C(K/n)^d e^n \to \infty$ for every $K > 0$, as required by FLS Assumption 1(ii). The case $m = 1$ is immediate. Thus the embedding satisfies the hypotheses of the FLS results being applied, without truncating the simplex or requiring common objective reward support.

KL minimization and part (i). Let $\alpha \in \Delta(\Phi)$ have marginals σ. Such a distribution always exists; for example,

$$\alpha(\phi) = \prod_{\omega\in\Omega^+} \sigma(\phi(\omega) \mid \omega). \tag{H.9}$$

With the usual extended-value convention for KL divergence,

$$\begin{aligned} \sum_\phi \alpha(\phi) D_{\mathrm{KL}}(P^*_\phi \| P^q_\phi) &= D_{\mathrm{KL}}(p\|\widetilde{p}) + \sum_{\omega, s\in\omega} p(\omega)\sigma(s \mid \omega) D_{\mathrm{KL}}(\mathsf{R}_{s\omega}\|q_s) \\ &= C_\sigma + \sum_s D_s(\sigma) D_{\mathrm{KL}}(q^\sigma_s \| q_s), \end{aligned} \tag{H.10}$$

where C_σ is finite and independent of q. Every coefficient $D_s(\sigma)$ is positive, so the unique minimizer is q^σ. This conclusion remains valid when its coordinates lie on simplex boundaries. All beliefs supported on the minimizer set therefore induce the same valuation vector v^σ.

A plan is optimal at these beliefs exactly when it selects a maximizing valuation in every supported menu, because every subjective menu weight is positive. Hence every plan in the support of α is optimal if and only if (H.6) holds. Uniqueness of the KL minimizer also makes the equilibrium strongly uniform in the sense of FLS Definition 13. For pure plans, their Theorem 1 then implies the final claim in (i): an eventual limit plan must be a uniform Berk–Nash equilibrium, hence a VE plan in this embedding. This does not assert convergence of the argmax process from arbitrary initial beliefs.

Strict stability, part (ii). When ϕ is menu-wise strict, it is the unique optimal plan at q^ϕ and therefore a uniformly strict Berk–Nash equilibrium. FLS Theorem 2 gives, for each $\epsilon > 0$, a concentration neighborhood of its KL-minimizing outcome law such that every admissible initial belief in that neighborhood generates eventual adherence to ϕ with probability greater than $1 - \epsilon$. Since the parameter-to-law map is continuous, choosing δ sufficiently small in (H.7) places the induced prior in that neighborhood. The result applies to every fixed measurable optimal tie-breaking rule.

On the adherence event, all observations after a finite date are generated by the fixed plan ϕ. The strong law for these iid menu–reward draws, together with $D_s(\phi) > 0$, makes each class's empirical reward distribution converge to q_s^ϕ. Finite initial histories and prior pseudocounts have vanishing weight, so (H.3) gives $v_k \to v^\phi$. These strong-law statements hold on an event of probability one and hence also on the positive-probability adherence event; no independence conditional on that event is being assumed.

Purification by finite recommendations, part (iii). Fix the distribution α in (H.9), and draw a recommendation $z \in Z := \operatorname{supp} \alpha$ iid with law α, independently of the menu and reward draws. The recommended class in menu ω is $z(\omega)$. Give class s the known additional utility $\tau \mathbf{1}\{s = z(\omega)\}$, so the decision maker chooses

$$s_k \in \arg\max_{s \in \omega_k} \{v_k(s) + \tau \mathbf{1}\{s = z_k(\omega_k)\}\}. \tag{H.11}$$

The reward entered in (H.3) is the base reward $r_k = y_k$. Equivalently, the decision maker subtracts the known bonus from total realized utility before updating the unknown class payoff.

Apply the finite-action embedding again, now with actions all maps $F : Z \times \Omega^+ \to \mathcal{S}$ satisfying $F(z, \omega) \in \omega$. An observed outcome is (z, ω, y), with laws

$$P_F^*(z, \omega, y) = \alpha(z) p(\omega) \mathsf{R}_{F(z,\omega),\omega}(y),$$
$$P_F^q(z, \omega, y) = \alpha(z) \widetilde{p}(\omega) q_{F(z,\omega)}(y).$$

The utility of this outcome is $y + \tau \mathbf{1}\{F(z, \omega) = z(\omega)\}$. Myopic optimization over F is separable across recommendation–menu pairs, so this ex ante construction implements exactly (H.11). The posterior for q is unchanged by the independent known signal probabilities and still obeys (H.3).

Consider $F^*(z, \omega) = z(\omega)$. Its induced menu-wise frequencies are σ, so (H.10) gives its unique KL minimizer as q^σ. At this model, the recommended class already maximizes base expected payoff by (H.6); its bonus then makes it uniquely optimal. Every other contingent plan differs at some pair (z, ω) and loses expected utility by at least

$$\tau \min_{z \in Z} \alpha(z) \min_{\omega \in \Omega^+} \widetilde{p}(\omega) > 0.$$

Thus F^* is uniformly strict. The dense-positivity and prior-thickness arguments above apply unchanged to this finite enlarged outcome space. FLS Theorem 2 yields, from sufficiently concentrated product-Dirichlet priors, eventual adherence to F^* with probability greater than $1 - \epsilon$.

On that event, the strong law for the iid pairs (z_k, ω_k) gives (H.8), and the class-specific reward strong law gives $v_k \to v^\sigma$. Singleton support again makes every class's limiting sampling frequency positive. The conclusion holds for each fixed $\tau > 0$; no exchange of the long-run and vanishing-perturbation limits is used.

Finally, the concentration neighborhoods in (ii) and (iii) are nonempty within the stipulated prior family. Given the relevant target q^*, fix $b_y > 0$ and take $a_{sy} = Lq^*_s(y) + b_y$. As $L \to \infty$, the Dirichlet means converge to q^*_s and their variances vanish, so their product concentrates at q^*. All parameters remain strictly positive, even if the target distribution has zero coordinates. □

Scope of the connection. The stability neighborhood controls posterior concentration, not only initial posterior mean valuations. Part (ii) is the strict-stability implication needed here; no converse is asserted within the restricted family of product-Dirichlet priors. Part (iii) is a finite-contingent-plan application of the same FLS theorem, in the spirit of their purification result for strongly uniform mixed equilibria. It constructs signals and bonuses specifically for the target VE, and its admissible neighborhood may shrink as $\tau \downarrow 0$. It establishes neither stability of that mixed VE under unperturbed argmax learning nor approximation under the particular logit family. The logit refinement, local ODE stability, and global convergence results of the main analysis therefore remain separate conclusions. The finite-outcome assumption here is also additional to the maintained moment assumptions used elsewhere.

I Convergence in balanced binary constant-sum decision trees

Suppose every supported menu is binary, write $p_{sr} = p_{rs}$ for the probability of $\{s, r\}$, and suppose its payoffs sum to $2c$, independently of the pair. Center rewards and valuations at c and write

$$\pi_s(\{s, r\}) - c = a_{sr} = -a_{rs}, \qquad B_{sr} = p_{sr} a_{sr}, \qquad P_s = \sum_r p_{sr} > 0.$$

Uniform menu-wise choice gives every class the same conditional mean precisely when $\sum_r B_{sr} = 0$ for every s. Indeed the common mean must be c, since the ex ante payoff of uniform choice is c and $\sum_s P_s/2 = 1$. In what follows valuations are centered, so the candidate SVE is $v = 0$.

Proposition I.1 (Global convergence under harmonic visits)**.** *At every finite sensitivity the unique SVE is $v = 0$, with $x_s = P_s/2$. It is globally asymptotically stable under the joint dynamics. Harmonic visit-count learning converges to it almost surely under the maintained second-moment assumption. The normalized equilibrium is locally exponentially stable at every finite sensitivity, but normalized global convergence need not hold.*

Proof. For $\beta > 0$, let $t_s = e^{\beta v_s}$. The centered unconditional reward contribution is $R_s(v) = \sum_r B_{sr} t_s/(t_s + t_r)$. Antisymmetry and the zero row sums give the exact cancellation

$$\begin{aligned} \sum_s \sinh(\beta v_s) R_s(v) &= \sum_{s<r} B_{sr} \frac{\sinh(\beta v_s) t_s - \sinh(\beta v_r) t_r}{t_s + t_r} \\ &= \frac{1}{2} \sum_{s<r} B_{sr}(t_s - t_r) = 0. \end{aligned} \tag{I.1}$$

Consider the nonnegative function

$$L(v, x) = \sum_s x_s \big(\cosh(\beta v_s) - 1\big).$$

Using $\dot{v}_s = (R_s - \Lambda_s v_s)/x_s$ and $\dot{x}_s = \Lambda_s - x_s$, its derivative is

$$\dot{L} = -L + \sum_s \Lambda_s(v)\{\cosh(\beta v_s) - 1 - \beta v_s \sinh(\beta v_s)\} \qquad \leq -L.$$

The inequality follows because $\cosh u - 1 - u \sinh u \leq 0$: its derivative is $-u \cosh u$, and its value at zero is zero. Thus $L(t) \leq e^{-t} L(0)$. Valuations are bounded and frequencies stay uniformly positive along each trajectory. Since $\cosh u - 1 \geq u^2/2$, this proves $v(t) = O(e^{-t/2})$. The frequency filter then gives $x(t) - P/2 = O(e^{-t/2})$. The same estimates in a fixed neighborhood prove Lyapunov stability. At any equilibrium $\dot{L} = 0$ forces $v = 0$, proving uniqueness. If $\beta = 0$, centered g is zero and the constant-target argument applies. Almost-sure convergence follows from the mean-field approximation and the globally asymptotically stable equilibrium.

To check normalized local stability, differentiation at zero gives

$$Dg(0) = -\frac{\beta}{2} Q^{-1} B, \qquad Q = \operatorname{diag}(P_s).$$

The diagonal terms vanish because $B\mathbf{1} = 0$. The matrix $Q^{-1}B$ is similar to the real skew-symmetric matrix $Q^{-1/2} B Q^{-1/2}$. Hence all eigenvalues of $Dg(0) - I$ have real part -1. The joint spectral correspondence yields the same nontrivial eigenvalues and the additional stable frequency directions. The following explicit example proves the last assertion. □

I.1 A normalized periodic orbit despite a stable unique SVE

Consider three binary menus with centered payoffs and probabilities

menu	probability	first payoff	second payoff
$\{1,2\}$	$1/28$	1	-1
$\{1,3\}$	$20/28$	$-1/20$	$1/20$
$\{2,3\}$	$7/28$	$1/7$	$-1/7$

Here $B_{12} = B_{23} = -B_{13} = 1/28$, so $B\mathbf{1} = 0$. The unique equilibrium has normalized eigenvalues $-1, -1 + i\beta/18, -1 - i\beta/18$. We next prove analytically that, for every sufficiently large finite β, a nonconstant periodic orbit also exists.

Set $X = \beta(v_1 - v_3)$, $Y = \beta(v_2 - v_3)$, $a = e^X$, and $b = e^Y$. The three conditional means are

$$g_1 = \frac{1-b}{21a + 20b + 1}, \qquad g_2 = \frac{a-1}{7a + 8b + 1}, \qquad g_3 = \frac{b-a}{7a + 20b + 27}.$$

With time $\tau = \beta t$ and $\epsilon = 1/\beta$, the gap system is

$$(X, Y)' = H(X, Y) - \epsilon(X, Y), \qquad H = (g_1 - g_3, g_2 - g_3).$$

Its Jacobian at zero, before subtracting ϵI, is

$$DH(0) = \begin{pmatrix} 1/54 & -8/189 \\ 35/432 & -1/54 \end{pmatrix}.$$

The change $u = X$, $w = -X/3 + 16Y/21$ makes the linear part $(u', w') = (-w/18, u/18)$ at $\epsilon = 0$. Write its nonlinear terms as (f, g). Direct differentiation at the origin gives

$$\begin{aligned} (f_{uu}, f_{uw}, f_{ww}) &= (11/576, 11/1728, -53/5184), \\ (g_{uu}, g_{uw}, g_{ww}) &= (-25/1728, -23/576, -5/1728), \\ (f_{uuu}, f_{uww}, g_{uuw}, g_{www}) &= (-7/3072, -757/82944, 77/3072, -145/82944). \end{aligned}$$

With $\omega = 1/18$, the cubic radial coefficient is

$$a_1 = \frac{f_{uuu} + f_{uww} + g_{uuw} + g_{www}}{16} + \frac{f_{uw}(f_{uu} + f_{ww}) - g_{uw}(g_{uu} + g_{ww}) - f_{uu}g_{uu} + f_{ww}g_{ww}}{16\omega} = \frac{247}{663552} > 0.$$

The eigenvalues cross the imaginary axis with real part $-\epsilon$. The nondegenerate subcritical Hopf theorem therefore produces a small repelling periodic gap orbit for every sufficiently small $\epsilon > 0$. Equivalently it exists for all sufficiently large finite β. The remaining normalized equation is the stable linear filter $\dot{v}_3 = g_3(X(t), Y(t)) - v_3$. A periodic forcing has exactly one periodic solution, so the gap orbit lifts to a periodic orbit in all three valuations. This establishes failure of normalized global convergence.

References

Benaïm, M. (1999). Dynamics of stochastic approximation algorithms. *Séminaire de Probabilités* XXXIII, 1–68. https://www.numdam.org/item/SPS_1999__33__1_0/.

Borkar, V. S. (1997). Stochastic approximation with two time scales. *Systems & Control Letters* 29(5), 291–294. https://doi.org/10.1016/S0167-6911(97)90015-3.

Brandière, O., and M. Duflo (1996). Les algorithmes stochastiques contournent-ils les pièges? *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 32(3), 395–427. https://www.numdam.org/item/AIHPB_1996__32_3_395_0/.

Fudenberg, D., G. Lanzani, and P. Strack (2021). Limit points of endogenous misspecified learning. *Econometrica* 89(3), 1065–1098. https://doi.org/10.3982/ECTA18508.

Kuznetsov, Y.A. (2013). *Elements of Applied Bifurcation Theory*. Applied Mathematical Sciences. Springer New York

van den Dries, L. (1998). *Tame Topology and O-minimal Structures*. London Mathematical Society Lecture Note Series 248. Cambridge University Press.

van den Dries, L., A. Macintyre, and D. Marker (1994). The elementary theory of restricted analytic fields with exponentiation. *Annals of Mathematics* 140(1), 183–205.