

Clearing Conditional Commitments: Privacy, Specification, and Custody in a Clearinghouse for Interdependent Action*

Timothy J. Miano[†]

June 2026 — preliminary and incomplete

Abstract

Latent willingness of the form “I will, if enough of the right others will” has no standard clearing institution. We model a *conditional-commitment clearinghouse*: a platform that authenticates participants (accuracy q), verifies that their mutual conditions are jointly satisfiable (consensus integrity c), routes compatible counterparts to one another, and chooses three design dials — custody of principal γ , specification timing σ , and disclosure granularity k — before activating the coalition at a fixed point. Our central result is that the welfare-maximizing clearinghouse is *minimal on all three dials at once*: it holds no principal, fixes terms late, and discloses progress only through a k -anonymized aggregate. The result is stated with its boundary: it holds in a bounded high-authenticity, intermediate-integrity region, reverses on each axis outside it, and the three minimalisms are coupled in welfare magnitude rather than additively separable. Three further results discipline the disclosure claim. First, the privacy dial’s screening cost is a genuine information-leakage object — the Shannon mutual information between coalition type and the k -anonymized support count — and full disclosure does not perfectly screen. Second, the masking advantage survives a refinement-free selection: under Carlsson–van Damme risk dominance no rationality parameter appears anywhere in the model, the revealed game’s contested-regime collapse is exact, and masked momentum is sustained — proven at the deadline rung and for the smallest coalitions, conjecture-gated (and grid-verified) at interior rungs. Third, partial masking is *constrained-optimal for a structural reason*: a single disclosure dial necessarily bundles “this coalition is real” with “you are not pivotal,” so — in a one-round theorem extended by a dynamic calibration — the interior optimum exists even where the unconstrained ideal would be full type-disclosure with zero pivotality leakage. We state three formal propositions with complete proofs, enumerate the open conjectures, and propose an administrable safe-harbor rule keyed to the mechanism’s structure.

JEL: C72, D82, D47, H41. **Keywords:** conditional commitment, coordination, information design, k -anonymity, market design, assurance contracts.

*Preliminary working draft (v1, June 2026). Comments welcome. The numerical companion — a chain of nine verified “toy model” memos, the solver scripts, and the adversarial verification logs — is available from the author. All errors are mine.

[†]Affiliate, Massachusetts Institute of Technology; and Founder & CEO, Proffer Futures, Inc. Contact: timmiano@mit.edu. *Disclosure:* the author holds equity in Proffer Futures, Inc., which operates the ifwish-list platform discussed in Section 9.

1 Introduction

1.1 The vocabulary, fixed first

The broad problem family is *interdependent action*: situations where each participant’s willingness to act depends on who else will act, in what number, under what terms, and with what visibility. Collective action, collective consumption, assurance, focal-point coordination, and matching and scheduling are subcases — the family is broad, and we will not pretend any one branch is the genus.

The economic object this paper introduces is the *conditional commitment*: “I will, if a qualifying composition of others will.” Not a pledge (a scalar added to a total), but a node with edges — a private commitment whose condition refers to the *composition* of the committing set: rosters, roles, types, compound thresholds.

The institution that makes the object clearable is the *conditional-commitment clearinghouse*: private conditional commitments are collected, checked for joint satisfiability, and activated at the moment the edge-constraints are simultaneously satisfied — a fixed point whose output is common knowledge of a feasible coalition. It clears commitments, not prices. Markets clear goods; this institution clears conditional commitments — and in doing so makes previously unclearable interdependent action actionable.

1.2 The friction: why the clearinghouse never existed

Conditional willingness has always existed. No institution cleared it because four barriers stood at once, and removing any three does not open the institution: credibility (“I would” is cheap talk; signals are forgeable), compatibility (mutual conditions may not be jointly satisfiable — is there a real coalition, or sixty unrelated “ifs”?), discoverability (compatible counterparts cannot find each other), and legality (turning activation into value movement can trigger custody, transmission, or liability regimes). The institution’s non-existence was overdetermined: each wall alone is fatal. The thesis is that the four walls are now jointly removable for the first time — credibility and discoverability by algorithmic inference at scale, compatibility by a mechanism-design move, legality by a legal-architecture move.

1.3 The contribution

Three results, in increasing order of surprise.

First, a clearable primitive. We formalize the conditional commitment as a graph object — nodes, compositional edge-constraints, a fixed-point activation event — and show that a pledge-threshold mechanism is its scalar degenerate case (Section 3).

Second, one central result, three minimalisms. In a single welfare objective over custody γ , specification σ , and disclosure k , the joint optimum is the minimal corner on all three axes together — no custody, fuzzy-early specification, partial masking — in a bounded high- q , intermediate- c region, reversing off it on each axis (Section 4). The efficient clearinghouse clears *because* it withholds.

Third, the disclosure dial decomposed. The privacy dial controls two channels with opposite signs — type/screening (favoring disclosure) and pivotality/protection (favoring masking) — and because one physical dial bundles them, partial masking is constrained-optimal even

where the unconstrained ideal is full type-disclosure (Sections 5–6). The screening cost of masking is quantified by an actual information measure, and the protection advantage survives a refinement-free equilibrium selection (with the momentum bound conjecture-gated at interior rungs, as the proposition statement discloses). Proposition 1 establishes the interior masking band in integrity; Proposition 2 the refinement-free selection; Proposition 3 the bundled-dial structure of partial masking.

The displacement claim in one line: prior work either fixes the good, conditions on a sum, or treats common knowledge as exogenous; we endogenize the good, the coalition membership, and the disclosure boundary, and then clear relational conditional commitments.

2 Related literature

Six lineages bracket the contribution; none contains it.

Collective action and provision points. Olson (1965) names the free-rider problem but offers no clearing institution. Bagnoli and Lipman (1989) and the assurance-contract mechanisms of Tabarrok (1998) fix the good ex ante and condition on a *scalar* threshold; we endogenize the good’s terms and make the condition *relational* (compositional edge-constraints). Marx and Matthews (2000) is the direct dynamic ancestor — an observable running total as the Markov coordinating state in voluntary contribution — and supplies the completion structure, but contains no disclosure-design result. Halac et al. (2020) reach unique implementation through contingent *payments* to heterogeneous investors over a scalar threshold: the transfer route to the destination our disclosure route reaches.

Coordination and common knowledge. Schelling’s focal points and Chwe (2001) explain when coordination succeeds given common knowledge; we *manufacture* common knowledge as a service, with endogenous membership. The global-games selection machinery we draw on is Carlsson and van Damme (1993) and Frankel et al. (2003). We cite the latter’s §6.4 for the criterion’s pedigree, not as a blanket license: where the stage game has strategic complementarities, their §6.4/Theorem 4 identifies the Laplacian (uniform-belief) action as the noise-independent selection; at the pivotal (diagonal) states that drive our result the stage game is a volunteer’s dilemma — commit and wait are strategic *substitutes* there — which falls outside the licensed class, so at those states we adopt the Laplacian action as the *definition* of stagewise risk dominance, cross-validated by an independent λ -free potential-maximizer refinement (Proposition 2 and its Appendix B scoping).

Information design. Morris and Shin (2002) and Cornand and Heinemann (2008) establish that public disclosure precision can flip welfare; Li et al. (2023) studies obfuscation directly. In all three, precision is a Gaussian noise parameter on a fundamental — not a quantified privacy-mechanism leakage measure, and there is no coalition screening rate. Kamenica and Gentzkow (2011), Bergemann and Morris (2019), Morris et al. (2024), and Inostroza and Pavan (2025) supply the persuasion, BCE, and adversarial-selection toolkit over payoff states: lineage, not preemption. Vosooghi (2017) is nearest by setting — information design inside coalition formation — but the lever is beliefs about a scalar social-cost state, not a privacy mechanism; its published dynamic extension (Vosooghi and Caparrós, 2022) has a third-party

sender choosing threshold revelation of a scalar fundamental to a single coalition, with no privacy mechanism over participant structure and no screening rate. Halac et al. (2021) is the closest privacy-as-instrument neighbor (rank uncertainty defeating strategic uncertainty in team production), with no quantified leakage primitive. Chen and Hu (2023) is the matching-side precedent for stability being information-structure-relative, with pairwise types and no disclosure welfare sign-flip.

Crowdfunding and matching. The crowdfunding design literature (Strausz, 2017; Ellman and Hurkens, 2019; Belleflamme et al., 2014) studies all-or-nothing thresholds and moral hazard over a *fixed* good with a scalar target; matching with contracts (Hatfield and Milgrom, 2005) clears relational structure but over fixed contract terms with no disclosure design. Both condition on what we endogenize.

Market design. The program of Roth and Peranson (1999); Roth (2002) — designing clearing institutions for market-like environments that did not previously clear — is the lineage we claim membership in. Posner and Weyl (2018) is nearest in ambition, distinct in object: we clear non-transferable conditional commitments, not assets or votes.

The new ingredient, absent from each lineage by construction, is a *quantified privacy-leakage primitive* (the mutual information between coalition type and a k -anonymized support count) *driving a coalition screening rate*, joined to a difficulty-keyed masking advantage. That conjunction is the contribution.

3 The environment

N agents; one latent outcome whose terms are not yet fixed. Each agent i holds a private type and a conditional commitment $c_i =$ “act if composition S_i of others commits.” A resolver authenticates nodes with accuracy q . A router surfaces the outcome to compatible agents. Coalitions are of two types, good (deliverable) and bad (illusory), with prior π_b on bad; consensus integrity c indexes the probability that the posted conditional structure is jointly satisfiable.

The platform chooses three dials:

- **Custody** $\gamma \in [0, 1]$ — the fraction of principal held at commitment time, carrying a trust-liability cost $c_{\text{cust}}\gamma$ and deadweight d_γ , against a possible commitment-device benefit β_γ that can lower the effective activation cutoff.
- **Specification** $\sigma \in [0, 1]$ — how early the outcome’s terms harden. Precise-early kills latent commitments via the compatibility factor $\text{comp}(\sigma)$; fuzzy-too-long adds indeterminacy cost $\text{indet}(\sigma)$. The outcome’s terms may be an *endogenous output* co-resolved by the committed set: scheduling-by-coalition is the clean instance, where the date is selected by the coalition rather than fixed before the coalition exists.
- **Disclosure** $k \in [0, 1]$ — the granularity of progress disclosure, implemented as a k -anonymity dial on the support count (additive blur $r(k) = \text{round}((1 - k)N)$).

Activation occurs when the edge-constraints are simultaneously satisfiable — the fixed point — after which campaign terms harden and value, where present, moves on the recipient’s own rail (the platform holds nothing). The welfare objective of the unified numerical model is

$$W(\gamma, \sigma, k; q, c) = \text{comp}(\sigma) (1 - \rho_{\text{eff}}(k, \gamma, c)) (V - \text{indet}(\sigma)) (1 - d_\gamma \gamma) - (1 - \text{comp}(\sigma)) (1 - s(k)) L - c_{\text{cust}} \gamma, \quad (1)$$

where ρ_{eff} is the effective unraveling risk and $s(k)$ the bad-coalition screening rate. We solve for when an activatable coalition forms — the fixed point exists *and is selected* — and for the dial setting that maximizes welfare: coalitions that should form, do; spurious ones do not.

Reduced-form status. The unified objective (1), the revealed game’s pivotality blend, and the bundled-dial channel coupling are disciplined reduced-form primitives, not objects derived from a Bayesian extensive form; the appendices state this where each is used, and each appendix’s local model is self-contained (local symbols are local). The conclusions most robust to this level of abstraction are the structural ones — the two-channel decomposition, the bundling constraint, the regime splits; the quantitative surfaces are measurement targets (Section 10).

4 The central result: minimal is optimal

Theorem 1 (informal; numerically established). *In a bounded high-authenticity, intermediate-integrity region, and under a bounded custody-benefit, the welfare-maximizing clearinghouse reduces custody, specification, and disclosure together: the joint optimum $(\gamma^*, \sigma^*, k^*)$ of (1) is $\gamma^* = 0$ (no custody), $\sigma^* = 0$ (fuzzy-early), and $k^* < 1$ (partial masking). The optimum reverses off the region on each axis. The three dials are coupled in welfare magnitude — quasi-separable in threshold structure, not additively separable — with specification \times disclosure the leading interaction.*

Each apparent weakness does real work. The clearinghouse clears precisely because it withholds: holding principal adds a trust liability the clearing does not require; forcing terms early kills latent commitments before the coalition can form; revealing granular pivotal state can unravel the conditional consensus it is meant to advertise.

Theorem 1 is established numerically (a restricted static disclosure facet is the part proven analytically, as Proposition 1 below); four qualifiers travel with it. *Region, not everywhere:* the joint-all-minimal bands are $c \in [0.66, 0.72]$ at $q = 0.95$ and $c \in [0.64, 0.74]$ at $q = 1.0$, and the region requires $q \gtrsim 0.95$; the masking band in c is an interior band — transparency wins below it (screening dominates) and, for $q < 1$, above it (at $q = 1$ with the reference parameters the upper closure fails and masking persists toward $c = 1$, Appendix A). *Coupled, not separable:* the three-dimensional joint argmax equals the product of one-dimensional argmaxes in only 221 of 600 parameter draws; the welfare cost of the separable policy is typically small (median 0.43%) with a real tail ($p_{95} = 92\%$). *Custody-minimalism is the bounded claim:* $\gamma^* = 0$ holds only when the commitment-device benefit is small ($\beta_\gamma \geq 0.25$ flips it), the qualifier [Diamond \(1984\)](#) already suggests. *Minimal disclosure means partial masking:* the optimal dial is an interior k -anonymity level, not a blackout; and the *dynamic* (live progress-bar) version of the masking advantage is a finite- N effect that vanishes in the continuum — the durable large-population result is the static coordination/screening logic. The dynamic effect is real for small campaigns, which is the regime the live platform actually runs.

5 The disclosure dial decomposed

The single dial k controls two information channels with opposite signs.

5.1 The type/screening channel

Disclosure carries a signal about coalition *type* — good (deliverable) versus bad (illusory). The channel’s precision is grounded in an actual information measure: the support count $m \sim \text{Binom}(N, p_\psi)$ with $p_G > p_B$ is published under k -anonymity, and precision is the Shannon mutual information $I(\text{type}; m_{\text{obs}})$. The numerical findings, each surviving cross-mechanism re-implementation: $I(k)$ is monotone (data processing) and *saturating* ($I(1) = 0.62 \text{ bits} < H = 0.97$ at the reference calibration — the exact count does not perfectly reveal type, because the binomial supports overlap); therefore full disclosure does not perfectly screen, $s(1) < 1$, *derived* from the support gap rather than posited, with $s(1) \rightarrow 1$ recovered in the large-gap perfect-separation limit; and the screening rate $s(k)$ is non-monotone under a pessimistic prior. One scope restriction: at $k = 1$ every privacy mechanism collapses to the exact count, so the endpoint is pinned by (p_G, p_B) alone, and the curve’s interior shape is mechanism-dependent. The grounding closes the endpoint, not the curve; (p_G, p_B) is itself the next-deeper primitive.

5.2 The pivotality/protection channel

Disclosure also reveals *pivotal structure* — who is and is not needed. The two channels live in complementary excludability regimes: the type channel has bite only at $\xi < 1 - \kappa/V$, where the activation payoff exceeds the free-ride payoff; the protection channel operates in the complementary regime where free-riding pays ($\xi > 1 - \kappa/V$). There, in a contested market, revealing pivotality lets inframarginal movers free-ride and collapses clearing; masking forces every committer to act as if pivotal, sustaining the coalition.

5.3 The static band: a theorem

The static interplay of screening and unraveling produces the central disclosure comparative static: partial masking is optimal on a nonempty set contained strictly inside the interior of the integrity range — a single band in every tested calibration, with connectedness the one open conjecture. Figure 1 displays it; Proposition 1 proves it.

The platform chooses a disclosure granularity $k \in [0, 1]$ ($k = 1$ full disclosure, $k = 0$ full masking). A coalition is good with probability qc and bad with probability $1 - qc$, where $q \in (0, 1]$ is authentication accuracy and $c \in (0, 1)$ is consensus integrity. A good coalition has robustness $r_g \sim F_c$ on $[0, 1]$ and delivers surplus $V > 0$ unless it unravels, which occurs iff $r_g < \hat{r}(k)$, where $\hat{r}(k)$ is the strategic activation cutoff induced by disclosure level k ; a bad coalition imposes loss $L > 0$ unless screened, which happens with probability $s(k)$. Welfare is

$$W(k; q, c) = qcV(1 - F_c(\tau k)) - (1 - qc)L(1 - k). \quad (\text{W})$$

Let $K^*(c) := \arg \max_{k \in [0, 1]} W(k; q, c)$ (the dependence on q is suppressed except in part (e)). We write “full disclosure is uniquely optimal at c ” for $K^*(c) = \{1\}$ and “masking is optimal at c ” for $1 \notin K^*(c)$. Define the *masking set*

$$M := \{c \in (0, 1) : 1 \notin K^*(c)\},$$

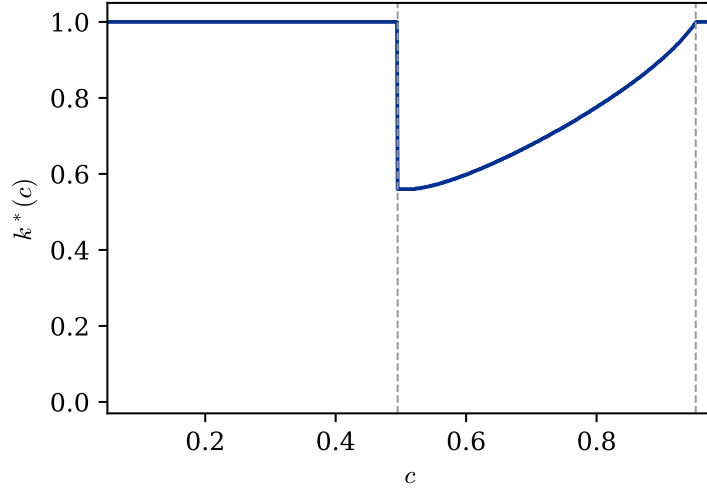


Figure 1: The interior masking band. Optimal disclosure $k^*(c)$ in the static model at $q = 0.95$ ($V = 1$, $L = 1.2$, $\tau = 0.7$, Beta integrity family with concentration $\eta = 8$). Full disclosure is optimal at both integrity extremes; partial masking on the interior band (dashed edges). The lower transition is a discontinuous global-max switch; the upper transition is continuous (Appendix A).

and the *marginal-balance function* (minus the right-hand slope of W at $k = 1$)

$$\varphi(c) := qcV\tau f_c(\tau) - (1 - qc)L. \quad (\varphi)$$

$\varphi(c) > 0$ says: at full disclosure, the marginal unraveling loss $qcV\tau f_c(\tau)$ exceeds the marginal screening gain $(1 - qc)L$.

Assumption 1.

(A1) (Payoffs.) $V, L > 0$, $\tau \in (0, 1)$, $q \in (0, 1]$, $c \in (0, 1)$; welfare is **(W)**.

(A2) (Linear channels — restriction.) $\hat{r}(k) = \tau k$, $s(k) = k$, hence the unraveling risk is $\rho(k, c) = F_c(\tau k)$. This is exactly the specification of the numerical ground-truth scripts `v3.py` / `v3_shape.py`.

(A3) (Family regularity.) For each $c \in (0, 1)$, F_c is a CDF on $[0, 1]$ with density f_c strictly positive and continuous on $(0, 1)$, and $(r, c) \mapsto f_c(r)$ is continuous on $(0, 1) \times (0, 1)$.

(A4) (Fragile-end thinness at the cutoff.) There exists $\bar{k} \in (0, 1)$ with

$$\lim_{c \rightarrow 0^+} c \sup_{r \in [\tau \bar{k}, \tau]} f_c(r) = 0.$$

(At very low integrity, the robustness density of good coalitions carries vanishing c -weighted mass near the transparency cutoff τ .)

(A5) (Robust-end thinness below the cutoff.) $\lim_{c \rightarrow 1^-} S(c) = 0$, where $S(c) := \sup_{r \in (0, \tau]} f_c(r)$, and **either** (i) $q < 1$, **or** (ii)

$$\limsup_{c \rightarrow 1^-} \frac{cV\tau S(c)}{(1-c)L} < 1.$$

(At very high integrity, good coalitions are too robust to unravel: the density on $[0, \tau]$ vanishes — fast enough, relative to $1 - c$, if $q = 1$.)

(A6) (Interior bite — non-vacuity.) There exists $c_m \in (0, 1)$ with $\varphi(c_m) > 0$, i.e.

$$q c_m V\tau f_{c_m}(\tau) > (1 - q c_m) L.$$

(A7) (Beta specialization, for parts (d).) $F_c = \text{Beta}(c\eta, (1-c)\eta)$ on $[0, 1]$ with fixed concentration $\eta > 2$ (mean c ; the family of `v2/v3` and of `v3_shape.py`). Define

$$\bar{c} := \frac{1 + \tau(\eta - 2)}{\eta} \in \left(\frac{1}{\eta}, 1\right) \quad (\text{the concavity threshold}),$$

the smallest c at which the Beta mode $(c\eta - 1)/(\eta - 2)$ reaches τ .

Proposition 1 (Interior masking band in c). *Assume A1–A3.*

(a) (Existence.) *For every $c \in (0, 1)$, $W(\cdot; c)$ is continuous on $[0, 1]$ and $K^*(c) \neq \emptyset$.*

Assume in addition A4–A6. Then:

(b) (Full disclosure at both extremes.) *There exist $0 < c_1 \leq c_2 < 1$ such that $K^*(c) = \{1\}$ for every $c \in (0, c_1)$ and every $c \in (c_2, 1)$.*

(c) (Masking in the middle; the band.) *$1 \notin K^*(c_m)$, so $M \neq \emptyset$. Consequently*

$$c_{\text{lo}} := \inf M \quad \text{and} \quad c_{\text{hi}} := \sup M$$

satisfy $0 < c_1 \leq c_{\text{lo}} \leq c_m \leq c_{\text{hi}} \leq c_2 < 1$; full disclosure is optimal ($1 \in K^(c)$) for every $c \in (0, c_{\text{lo}}) \cup (c_{\text{hi}}, 1)$, uniquely so on $(0, c_1) \cup (c_2, 1)$; and masking is strictly optimal at $c_m \in [c_{\text{lo}}, c_{\text{hi}}]$. In particular the masking region is contained in a band strictly interior to $(0, 1)$ and is nonempty.*

Assume in addition A7. Then:

(d) (Sharp upper threshold and FOC characterization on the upper shoulder.)

(d1) *$c \mapsto \log[qcV\tau f_c(\tau) / ((1 - qc)L)]$ is strictly concave on $(0, 1)$; hence $\{c : \varphi(c) > 0\}$ is an open interval (c_-, c_+) , nonempty by A6, with $c_- < c_m < c_+$.*

(d2) *For every $c \in [\bar{c}, 1)$, $W(\cdot; c)$ is strictly concave on $[0, 1]$, $K^*(c)$ is a singleton $\{k^*(c)\}$, $k^*(c) > 0$, and*

$$k^*(c) < 1 \quad \iff \quad \varphi(c) > 0;$$

when $\varphi(c) > 0$, $k^(c) \in (0, 1)$ is the unique solution of the first-order condition*

$$qcV\tau f_c(\tau k^*) = (1 - qc)L \quad (\text{marginal unraveling loss} = \text{marginal screening gain}).$$

(d3) If moreover $c_+ > \bar{c}$ (true at the script parameters $(V, L, \tau, \eta) = (1, 1.2, 0.7, 8)$ with $q = 0.95$; Remark A.2), then $c_{\text{hi}} = c_+ < 1$, the interval $(\max(\bar{c}, c_-), c_+) \subseteq M$, k^* is continuous on $(\max(\bar{c}, c_-), c_+)$, and $k^*(c) \rightarrow 1$ as $c \uparrow c_{\text{hi}}$: the upper transition is a continuous closing of the band, characterized exactly by $\varphi(c_{\text{hi}}) = 0$.

(e) (Comparative statics in q ; needs only A1–A3.) For $q < q' \leq 1$:

- $M(q) \subseteq M(q')$: the masking set expands with authentication accuracy; hence, whenever $M(q) \neq \emptyset$ (so that the thresholds $c_{\text{lo}}(q), c_{\text{hi}}(q)$ are defined — and then $M(q') \neq \emptyset$ too, by the inclusion), $c_{\text{lo}}(\cdot)$ is non-increasing and $c_{\text{hi}}(\cdot)$ non-decreasing in q . (When $M(q) = \emptyset$ the thresholds are undefined — $\inf \emptyset / \sup \emptyset$ — and the clause is vacuous.)
- $\max K^*(c; q') \leq \min K^*(c; q)$ for every c : every optimal disclosure level under q' is weakly below every optimal level under q . In particular, wherever K^* is a singleton, $k^*(c; q)$ is non-increasing in q (more masking as authenticity rises).

Conjecture 1 (Connectedness of M). M is an interval, i.e. $M = (c_{\text{lo}}, c_{\text{hi}})$ exactly.

The proof (Appendix A) characterizes the band sharply for the Beta family: the first-order-condition violation set is an exact interval, via strict log-concavity of the marginal-balance ratio (a trigamma inequality); the upper threshold is the root of an explicit equation with $k^*(c) \rightarrow 1$ continuously; and Topkis comparative statics give that raising authentication accuracy q expands the masking set and lowers every optimal disclosure level. Two scope restrictions: at $q = 1$ with the reference parameters the upper closure fails — masking persists toward $c = 1$, consistent with the numerics — so the $c_{\text{hi}} < 1$ half is scoped to $q < 1$; and connectedness of the masking set is Conjecture 1 (the lower transition occurs while the corner first-order condition still holds, so no FOC characterizes it).

6 Selection and the structure of partial masking

Two results complete the disclosure story: the masking advantage needs no equilibrium refinement parameter, and partial masking has two structural sources.

6.1 A refinement-free selection

Earlier numerical versions selected the dynamic equilibrium by a logit rule with rationality parameter λ — and the masking advantage inverted at high λ , leaving the result refinement-dependent. The resolution replaces logit with Carlsson–van Damme risk dominance: a p_{opp} -consistent Laplacian belief, uniform over opponents' commit *propensity*, routed through the binomial movement kernel. At diagonal (volunteer's-dilemma) states, where complementarities fail, the Laplacian action is adopted as the *definition* of stagewise risk dominance rather than invoked as an FMP theorem (§2). No rationality parameter appears anywhere in the model, so no inversion is possible.

Notation. At support count m write $g(m) := K - m$ (gap) and $n(m) := N - m$ (uncommitted agents). $A[m, h] \in \{0, 1\}$ is the action selected by the masked (λ -free CvD/Laplacian) rule at state (m, h) ; π_{m} and π_{r} are the clearing probabilities $P_c[0, T]$ of the masked and revealed games, and $\text{Adv} := \pi_{\text{m}} - \pi_{\text{r}}$. $q_j(n) := \int_0^1 \text{Binom}(n - 1, p\alpha)(j) d\alpha$ is the Laplacian co-mover law. *DI* (*diagonal ignition*) is the property that the masked rule selects commit at every diagonal state:

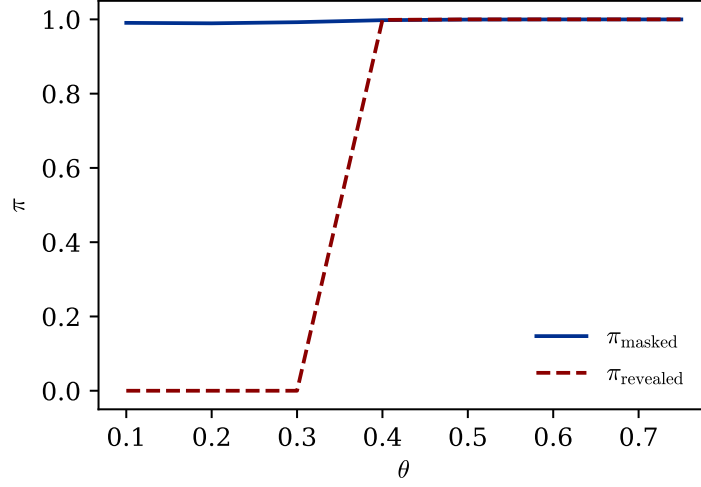


Figure 2: Masked versus revealed clearing probability across market difficulty θ , at the reference calibration. The revealed game collapses exactly ($\pi_r = 0$) in the contested regime; masked momentum is sustained throughout. The masking advantage at $\theta \leq 0.30$ is $+0.99$.

$A[K - g, g] = 1$ for $g = 1, \dots, K$. The momentum bound is $\widehat{P}(\gamma) := \prod_{i=1}^{\gamma} (1 - (1 - p)^{N-K+i})$, with $\widehat{P}(0) := 1$.

Assumptions.

- **(A1) (primitives).** $N \geq 2$, $2 \leq K \leq N$, $T \geq K$, $p \in (0, 1)$, $V > \kappa > 0$, $L_p \geq 0$, $\rho \geq 0$, $\xi \in [0, 1]$.
- **(A2) (kernel).** The dynamic game and profile-consistent arrays of the setup (appendix, §0.1).
- **(A3) (masked rule).** The exact-integral CvD/Laplacian stage selection (appendix, §0.2). At diagonal volunteer's-dilemma states, where strategic complementarities fail, the Laplacian action is adopted as the *definition* of stagewise risk dominance, not invoked as an FMP §6.4 theorem (see the appendix scoping remark).
- **(A4) (revealed rule).** The pivotality-blend reduced form (appendix, §0.3).
- **(A5) (free-riding pays).** $\xi V > V - \kappa$, i.e. $\xi > 1 - \kappa/V$.
- **(A6) (contested regime).** $K - 1 \geq \max(1, pN)$.
- **(A7) (deadline ignition).** $\xi V \cdot (1 - q_0(N - K + 1)) < V - \kappa$, where $q_0(n) := (1 - (1 - p)^n)/(np)$.
- **(A8) (rung-2 ignition, used only for the $K = 2$ part).** With $q_j := q_j(N)$ and $\widehat{P}(1) := 1 - (1 - p)^{N-1}$:
 - (C2)** $q_0 \cdot [(V - \kappa + L_p) \cdot \widehat{P}(1) - L_p + \rho] + q_1 \cdot \rho > (\xi V - (V - \kappa)) \cdot (1 - q_0 - q_1)$;
 - (C3)** $L_p \cdot q_0 + \xi V \cdot (1 - q_0 - q_1) \geq (V - \kappa) \cdot (1 - q_0)$.

A5–A6 are the **contested-regime** hypotheses (the θ -threshold of part (iv)); A7 is the **momentum-ignition** hypothesis (the θ -threshold of part (ii)). With the v9 grid map $K(\theta) = \lceil (1 - \theta) \cdot N \rceil$, A6 reads: θ lies below the threshold at which the initial gap exceeds the expected per-period mover count (at the v9 calibration, where $pN = 6$ is an integer: $\theta < 1 - p = 0.4$, i.e. exactly $\theta \in \{0.10, 0.20, 0.30\}$ on the grid).

Proposition 2 (λ -free selection: masked momentum and the contested-regime advantage). *Under A1–A3 the masked game and under A1, A2, A4 the revealed game are well-defined, and no rationality parameter appears anywhere; for fixed primitives Adv is a single number (no λ -inversion is possible). Moreover:*

(i) (Momentum, conditional form.) *If DI holds, then $\pi_m \geq \widehat{P}(K) > 0$.*

(ii) (Ignition at the last rung; θ -threshold.) *Under A7, the masked rule commits at the deadline-pivotal diagonal state: $A[K - 1, 1] = 1$. Since $q_0(n)$ is strictly decreasing in n , A7 is downward-closed in θ : if it holds at threshold K it holds at every $K' \geq K$, i.e. for every harder (lower- θ) market.*

(iii) (Momentum, unconditional for $K = 2$.) *Let $K = 2$ and assume A5, A7, A8. Then DI holds — $A[1, 1] = 1$ and $A[0, 2] = 1$ — and therefore*

$$\pi_m \geq \widehat{P}(2) = (1 - (1 - p)^{N-1})(1 - (1 - p)^N) > 0.$$

The hypothesis set is non-vacuous: $(N, K, T, p) = (10, 2, 10, 1/10)$ with the v9 payoffs $(V, \kappa, L_p, \rho, \xi) = (3, 1, 3, 1/20, 4/5)$ satisfies A1, A5, A6, A7, A8 simultaneously (exact margins in the appendix proof, Step 5).

(iv) (Revealed collapse in the contested regime.) *Under A1, A2, A4, A5, A6: $\pi_r = 0$ exactly. In particular at the v9 calibration $(N, p) = (10, 3/5)$, A6 holds iff $K \geq 7$, i.e. iff $\theta \in \{0.10, 0.20, 0.30\}$ on the v9 grid.*

(v) (Masking advantage.) *Under the hypotheses of (iv) together with DI (or, for $K = 2$, with A7–A8 instead of DI): $\text{Adv} = \pi_m \geq \widehat{P}(K) > 0$ — strictly positive on the contested regime. Globally (without A5–A6), under DI: $\text{Adv} \geq \widehat{P}(K) - 1$. (No claim of $\text{Adv} \geq 0$ outside the contested regime is made: the ground-truth computation exhibits $\text{Adv} \in [-8 \cdot 10^{-4}, 0)$ on the easy regime; see Remark R4 in the appendix.)*

Conjecture 2 (Diagonal ignition at the v9 calibration). *At $(N, T, p, V, \kappa, L_p, \rho, \xi) = (10, 10, 3/5, 3, 1, 3, 1/20, 4/5)$ and every $K \in \{3, \dots, 9\}$ of the v9 θ -grid, DI holds at every diagonal rung (and, additionally, the masked rule waits at every below-diagonal state). This is verified pointwise by the ground-truth solver at all seven grid points; parts (i) and (v) of Proposition 2 then apply at the calibration. What would close it: a proof of the diagonal stage inequality at rungs $g \geq 2$, which requires an upper bound on the on-cone uncommitted value V_u strictly below ξV by a margin of order $\xi V \cdot (1 - P_c) + \rho$; the crude bound $V_u \leq \xi V$ used in the present technique yields a sufficient condition that is strictly negative at the calibration, while the true (verified) deepest-rung stage margin is +0.417. See Remark R3 in the appendix.*

The contested-regime collapse — part (iv) — is exact and fully proven: when free-riding pays ($\xi > 1 - \kappa/V$) and the market is contested, the pivotality weight classifies every mover as inframarginal, revelation triggers the free-ride action, and $\pi_r = 0$ identically. The momentum

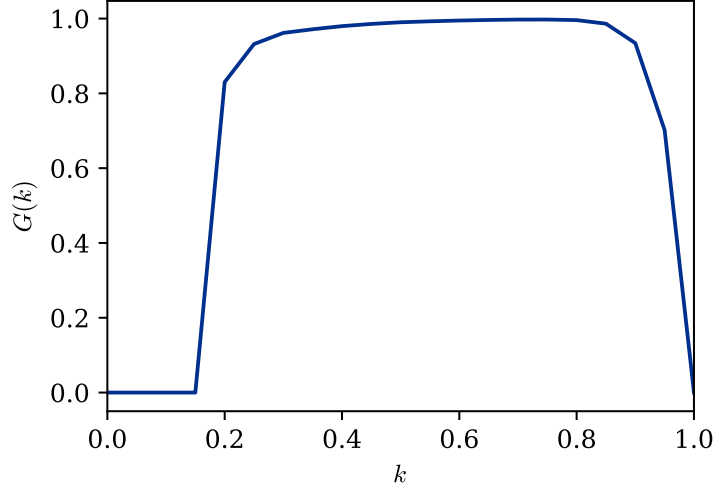


Figure 3: The bundled-dial inverted-U. Activation gain $G(k)$ at the reference contested calibration ($\theta = 0.30$, $\xi = 0.667$, $T = 10$): both endpoints are dead — full masking cannot ignite momentum, full disclosure fully unravels — and the interior peak near $k = 0.70$ carries the entire welfare range.

floor is proven conditional on diagonal ignition (property DI), which is itself proven outright at the deadline rung and unconditionally for $K = 2$; at interior rungs for $K \geq 3$ it is Conjecture 2, verified pointwise at every calibration grid point. The precise reading of the advantage: strictly positive when contested, near-tie elsewhere (the numerics give -8×10^{-4} at $\theta = 0.40$ — “ ≥ 0 everywhere” is false and we do not claim it).

6.2 The bundled dial

Why is the optimal masking *partial*? With both channels live in the dynamic game on the single bundled dial, a genuine interior optimum exists (Figure 3).

Assumption 2. In the activation game $\Gamma_T(k_{\text{type}}, k_{\text{piv}})$ (setup M1–M6 in the appendix; bundled dial $k_{\text{type}} = k_{\text{piv}} = k$): (A0) regularity: $N \geq 2$; $K \in \{2, \dots, N\}$; $p_{\text{opp}} \in (0, 1)$; $0 < \kappa < V$; $L > 0$; $L_p > 0$; $\rho \geq 0$; $\pi_b \in (0, 1)$; $\xi \in (0, 1]$. (A1) pessimistic pooling: $0 \leq \mathbb{E} < \kappa$, where $\mathbb{E} := (1 - \pi_b)V - \pi_b L$. (A2) free-ride regime: $\varepsilon := \xi V - (V - \kappa) > 0$, i.e. $\xi > 1 - \kappa/V$, so the gate $\chi = 1$. (A3) thin root field: $K - 1 \geq p_{\text{opp}}N$, so $\text{piv}(0) = 0$. (A4) pivotal ignition at full type-news: with $X \sim \text{Binom}(N - 1, p_{\text{opp}})$ and $\beta_1 = \Pr(X = K - 1)$, $\beta_+ = \Pr(X \geq K)$, $\beta_- = \Pr(X \leq K - 2)$: $(V - \kappa)\beta_1 > \varepsilon\beta_+ + L_p\beta_-$. (A5) small waiting friction: $\rho(T - 1) < \min\{L_p, \kappa - \mathbb{E}\}$ (vacuous at $T = 1$).

Proposition 3 (bundled-dial interior optimum; decoupled-dial corner).

- (i) **Bundled dial: both corners dead, a live interior, every optimum interior.** Endpoints (any horizon). Under A0, A1, A2, A3, A5, for every $T \geq 1$, the bundled game satisfies

$$G(0) = 0, \quad s(0) = 1, \quad G(1) = 0, \quad s(1) = 1, \quad \text{hence} \quad W(0) = W(1) = 0.$$

($k = 0$: pooled news cannot ignite momentum — the unique stage equilibrium is $q = 0$ at every state; $k = 1$: full pivotality leakage makes the empty state absorbing — every root-state mover free-rides.) Interior (one round). If in addition A4 holds and $T = 1$, there is a threshold $\bar{k} \in (0, 1)$ — the unique root of the affine root-state differential D_1 of the appendix ignition lemma — such that for every $k \in [\bar{k}, 1)$:

$$G(k) = \Pr(\text{Binom}(N, p_{\text{opp}}(1 - k)) \geq K) > 0, \quad s(k) = 1, \\ W(k) = (1 - \pi_b) V G(k) > 0.$$

Moreover $W : [0, 1] \rightarrow \mathbb{R}$ is upper semicontinuous, attains its maximum, and **every maximizer lies in the open interval** $(0, 1)$.

(ii) **Decoupled dials: the optimum is the corner.** Under A0–A5 and $T = 1$, the decoupled welfare satisfies, for all $(k_{\text{type}}, k_{\text{piv}}) \in [0, 1]^2$,

$$W(k_{\text{type}}, k_{\text{piv}}) \leq W(1, 0) = (1 - \pi_b) V \Pr(\text{Binom}(N, p_{\text{opp}}) \geq K) > 0,$$

and the argmax set is **exactly** $[\bar{k}, 1] \times \{0\}$ (with \bar{k} from part (i)): every optimum has zero pivotality leakage and full ignition, and $(k_{\text{type}}, k_{\text{piv}}) = (1, 0)$ attains the optimum. In particular every point with $k_{\text{piv}} > 0$ is strictly suboptimal.

(iii) **The interior is a constrained optimum created by bundling.** Under A0–A5 and $T = 1$: $\sup_{k \in [0, 1]} W(k, k) \leq W(1, 0)$, and the bundled diagonal never attains $W(1, 0)$ (every diagonal point has $a_{\text{eff}} < 1$). The bundled optimum is interior (part (i)) while the unconstrained two-dial optimum is the corner (part (ii)); the interior k^* is the constrained optimum of one dial forced to move both channels together — not a type-versus-pivotality balance.

Conjecture 3 (general-horizon interior). Under A0–A3, A5, and a strengthened ignition condition, for every $T \geq 1$ there exists $k \in (0, 1)$ with $G(k) > 0$ in the bundled game. (Proved at $T = 1$; open for $T \geq 2$.)

Observation 1 (positive leakage can strictly help at $T \geq 2$). There exist instances satisfying A0–A5, horizons $T \geq 2$, and dials $k_{\text{type}} \in [0, 1]$, $k_{\text{piv}} > 0$ such that $W(k_{\text{type}}, k_{\text{piv}}) > W(k_{\text{type}}, 0)$ with $s \equiv 1$ at both points: positive pivotality leakage is strictly welfare-improving, through the good-coalition continuation/waiting-option channel. At $T = 1$ this is impossible (Proposition 3(ii)), so the phenomenon is strictly dynamic. (Verified instance in the appendix.)

Conjecture 4 (narrowed residual conjecture: decoupled argmax retains $k_{\text{piv}} = 0$). Under A0–A5, for every $T \geq 1$ the decoupled welfare attains its maximum at some point with $k_{\text{piv}} = 0$ — with no claim that $(1, 0)$ attains it, and no pointwise dominance claim in k_{piv} . (Theorem at $T = 1$ by Proposition 3(ii); open for $T \geq 2$.)

The original general-horizon decoupled-corner conjecture — that under A0–A5 the decoupled argmax has $k_{\text{piv}} = 0$ and is attained at $(1, 0)$ for every $T \geq 1$ — was falsified in verification by an explicit A0–A5 counterexample at $T \in \{2, 5, 10\}$; the falsification record, counterexample parameters, and replication notes appear in the appendix (Remark C.4).

The decoupled-dial result — part (iii) — is what makes the interior structural rather than incidental: with two separate dials, the optimum is the corner $(k_{\text{type}}, k_{\text{piv}}) = (1, 0)$, full type-disclosure with zero pivotality leakage. The bundled interior is therefore a *constrained* optimum

created by the yoking — one physical dial cannot certify a coalition’s reality without leaking its pivotal structure — and never a type-versus-pivotality balance. Partial masking thus has two distinct sources: the integrity-distribution shape (Proposition 1) and the one-dial coupling (Proposition 3).

Horizon scope: parts (ii)–(iii) are $T = 1$ theorems; the endpoint results hold at every horizon. At $T \geq 2$ the original general-horizon decoupled-corner conjecture was *falsified* in adversarial verification — within the proposition’s own assumptions, the corner can fail to be attained, and positive pivotality leakage can strictly improve welfare at fixed type-disclosure (Observation 1): a leakier future cuts the waiting option and re-ignites mid-states through the continuation channel. Whether the decoupled argmax always retains $k_{\text{piv}} = 0$ is open (Conjecture 4). The leakage-helps phenomenon is itself a new, empirically probeable prediction.

7 Boundary: where the theorem reverses

The theorem is conditional on the walls actually being down.

q collapses. Authentication is weak; masked disclosure now shields *manufactured* commitments rather than protecting a real coalition. “Minimal disclosure” flips from efficient to fraud-enabling. Below the located threshold (the joint region requires $q \gtrsim 0.95$), transparency dominates.

c collapses. The conditional structure is illusory — nodes are real but their conditions are not jointly satisfiable; clearing produces no activatable outcome. The masking band is interior in c (Proposition 1): below it, screening dominates and transparency wins; above it, coalitions are too robust to unravel and disclosure is harmless.

The contested regime is absent. The protection channel matters only where the static selection frontier binds and free-riding pays ($\xi > 1 - \kappa/V$): a good coalition that exists but would unravel if its pivotal structure were exposed. Where completion is easy, masking is neutral (the near-tie region of Proposition 2).

Endogeneity over-extends. If outcome terms stay fuzzy past the binding moment, the campaign instrument is indeterminate; precision must harden as commitment strength rises (the interior σ^* under high indeterminacy cost).

The taxonomy over-expands. “Interdependent action” becomes a slogan unless every example reduces to a conditional-commitment graph — nodes, edges, constraints, fixed point. The family is broad; the primitive must stay narrow.

8 Mechanism and an administrable rule

Mechanism. A no-custody clearinghouse for conditional commitments with (a) authenticated nodes, (b) verifiable conditional structure, (c) endogenous outcome terms with a hardening schedule, (d) partial disclosure (an interior k -anonymity level) on granular state, (e) value riding the recipient’s own rail, and (f) non-transferable participant commitments.

Normative rule. We propose model-motivated design criteria for a limited safe harbor for a privacy-preserving conditional-commitment clearinghouse — a proposed administrable rule, not a legal conclusion, and pending counsel review. The four criteria: non-transferable participant commitments; no custody of principal; an auditable node-authentication process;

and small- N disclosure controls. The rule regulates the *clearing of commitments* — not raw wishes, raw data, platform funds, or transferable financial instruments — and is keyed to the conditional-commitment structure, not the campaign’s vertical. The theory motivates each prong: the authentication prong is the q -collapse boundary; the disclosure-controls prong is the small- N regime where the dynamic protection effect (and its abuse surface) lives; the no-custody and non-transferability prongs keep the cleared object a coordination device rather than a financial instrument.

9 Implementation: a constructive implementability case

The empirical spine is not yet a dataset — it is a live platform’s release history, readable as a constructive case that the institution is implementable. Implementability is all this section claims: a release history cannot validate optimality, welfare, or the selection results, and the calibration quantities are unmeasured (Section 10). The platform (ifwishlist, live June 2026) did not ship unrelated features; it removed the four walls one at a time: authenticated counterpart resolution (credibility, q); a conditional-commitment grammar with compound and roster thresholds (compatibility, c); semantic routing (discoverability); a no-custody rail (legality); and the k -anonymity dial with a specification-hardening schedule — an architecture consistent with the model’s recommended design.

10 Empirical agenda

The theory chain has largely converged; the marginal value has moved to measurement. Four quantities, all currently unmeasured, each first-order for a specific result:

Quantity	Governs	Disciplines
support gap $p_G - p_B$	the screening cap $s(1)$, large-gap reversal	§5
excludability ξ	which channel regime governs	§5–6
campaign-size distribution	the scale of the finite- N protection effect	§4
π_b and the integrity distribution	band location; interior k^*	Prop. 1, 3

A further theory iteration is warranted only when a measured quantity makes a specific model question first-order. Observation 1 adds a sharp new item: the leakage-helps regime is directly probeable on live activation data.

11 Scope, open questions, and the verification record

Every numerical claim in this paper comes from a chain of solver scripts under fixed seeds, and every analytical result was drafted against those scripts as ground truth and then attacked by three separately instantiated adversarial verification passes — a step-by-step logic audit, a counterexample hunt running randomized parameter sweeps inside the stated assumption sets, and a scope audit of claim–proposition alignment — each run with no access to the others’ findings, with scripts and logs archived alongside the numerical companion. Roughly 3,500 randomized counterexample instances were run against the stated assumption sets; no proved statement was refuted. The process corrected the first draft of every artifact it touched —

including four sections of this paper — and falsified one of our own conjectures, producing Observation 1. We report the open residue explicitly:

1. **Conjecture 1** (connectedness of the masking band): the lower transition is a discontinuous global-max switch with no first-order characterization.
2. **Conjecture 2** (diagonal ignition at interior rungs): closing it needs a simultaneous induction bounding the on-cone waiting value.
3. **Conjecture 3** (the $T > 1$ interior): numerically robust; the band grows with horizon.
4. **Conjecture 4** (decoupled argmax retains zero leakage): open after refinement; the pointwise version is false (Observation 1).
5. Deriving (p_G, p_B) from a deeper compatibility primitive; correlated signals; an endogenous adverse-selection prior π_b .

Two further scope notes. The legality-wall claim inherits open predicates pending counsel review (the no-custody rail’s money-transmission posture; charitable-solicitation regimes). And the live system does not yet instrument q or c ; both are required for the central theorem to bite on data.

12 Conclusion

AI does not make the engine intelligent. It removes two of the four walls — credibility and discoverability — that kept interdependent willingness from ever clearing. The mechanism removes the third (compatibility); the legal architecture removes the fourth. With all four down, “I would, if...” becomes a clearable economic object — and the efficient clearinghouse is the one that holds nothing, fixes nothing early, and hides the granular.

References

- Mark Bagnoli and Barton L. Lipman. Provision of public goods: Fully implementing the core through private contributions. *The Review of Economic Studies*, 56(4):583–601, 1989. doi: 10.2307/2297502.
- Paul Belleflamme, Thomas Lambert, and Armin Schwienbacher. Crowdfunding: Tapping the right crowd. *Journal of Business Venturing*, 29(5):585–609, 2014. doi: 10.1016/j.jbusvent.2013.07.003.
- Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019. doi: 10.1257/jel.20181489.
- Hans Carlsson and Eric van Damme. Global games and equilibrium selection. *Econometrica*, 61(5):989–1018, 1993. doi: 10.2307/2951491.
- Yi-Chun Chen and Gaoji Hu. A theory of stability in matching with incomplete information. *American Economic Journal: Microeconomics*, 15(1):288–322, 2023. doi: 10.1257/mic.20200411.

- Michael Suk-Young Chwe. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton University Press, Princeton, NJ, 2001. Paperback reissue 2013, ISBN 9780691158280.
- Camille Cornand and Frank Heinemann. Optimal degree of public information dissemination. *The Economic Journal*, 118(528):718–742, 2008. doi: 10.1111/j.1468-0297.2008.02139.x.
- Douglas W. Diamond. Financial intermediation and delegated monitoring. *The Review of Economic Studies*, 51(3):393–414, 1984. doi: 10.2307/2297430.
- Matthew Ellman and Sjaak Hurkens. Optimal crowdfunding design. *Journal of Economic Theory*, 184:104939, 2019. doi: 10.1016/j.jet.2019.104939. Article 104939.
- David M. Frankel, Stephen Morris, and Ady Pauzner. Equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory*, 108(1):1–44, 2003. doi: 10.1016/S0022-0531(02)00018-2.
- Marina Halac, Ilan Kremer, and Eyal Winter. Raising capital from heterogeneous investors. *American Economic Review*, 110(3):889–921, 2020. doi: 10.1257/aer.20190234.
- Marina Halac, Elliot Lipnowski, and Daniel Rappoport. Rank uncertainty in organizations. *American Economic Review*, 111(3):757–786, 2021. doi: 10.1257/aer.20200555.
- John William Hatfield and Paul R. Milgrom. Matching with contracts. *American Economic Review*, 95(4):913–935, 2005. doi: 10.1257/0002828054825466.
- Nicolas Inostroza and Alessandro Pavan. Adversarial coordination and public information design. *Theoretical Economics*, 20(2):763–813, 2025. doi: 10.3982/TE5768. Open access at econtheory.org. Previously circulated as “Persuasion in Global Games with Application to Stress Testing.”
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011. doi: 10.1257/aer.101.6.2590.
- Fei Li, Yangbo Song, and Mofei Zhao. Global manipulation by local obfuscation. *Journal of Economic Theory*, 207:105575, 2023. doi: 10.1016/j.jet.2022.105575.
- Leslie M. Marx and Steven A. Matthews. Dynamic voluntary contribution to a public project. *The Review of Economic Studies*, 67(2):327–358, 2000. doi: 10.1111/1467-937X.00134.
- Stephen Morris and Hyun Song Shin. Social value of public information. *American Economic Review*, 92(5):1521–1534, 2002. doi: 10.1257/000282802762024610.
- Stephen Morris, Daisuke Oyama, and Satoru Takahashi. Implementation via information design in binary-action supermodular games. *Econometrica*, 92(3):775–813, 2024. doi: 10.3982/ECTA19149.
- Mancur Olson. *The Logic of Collective Action: Public Goods and the Theory of Groups*, volume 124 of *Harvard Economic Studies*. Harvard University Press, Cambridge, MA, 1965. ISBN 978-0674537507. Revised printing 1971, ISBN 978-0674537514.

- Eric A. Posner and E. Glen Weyl. *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton University Press, Princeton, NJ, 2018. ISBN 978-0691177502.
- Alvin E. Roth. The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica*, 70(4):1341–1378, 2002. doi: 10.1111/1468-0262.00335.
- Alvin E. Roth and Elliott Peranson. The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review*, 89(4): 748–780, 1999. doi: 10.1257/aer.89.4.748.
- Roland Strausz. A theory of crowdfunding: A mechanism design approach with demand uncertainty and moral hazard. *American Economic Review*, 107(6):1430–1476, 2017. doi: 10.1257/aer.20151700.
- Alexander Tabarrok. The private provision of public goods via dominant assurance contracts. *Public Choice*, 96(3–4):345–362, 1998. doi: 10.1023/A:1004957109535.
- Sareh Vosooghi. Information design in coalition formation games. Nota di Lavoro 28.2017, Fondazione Eni Enrico Mattei (FEEM), 2017. Working paper; SSRN 2987779. Related published extension: Vosooghi & Caparrós (2022), *European Economic Review* 143:104042.
- Sareh Vosooghi and Alejandro Caparrós. Information disclosure and dynamic climate agreements: Shall the IPCC reveal it all? *European Economic Review*, 143:104042, 2022. doi: 10.1016/j.euroecorev.2022.104042.

A Proof of Proposition 1

Provenance. This appendix formalizes the v2 c -axis reversal / v3 “ k^* -valley” claim. Ground truth for the model is `docs/lit/conditional-commitment/runs/numcheck/v3.py` and `v3_shape.py`; the headline anchor sequence is produced by `v7_unify.py` §0 (see Remarks A.2 and A.3 for its exact parameters). The proof below stands without them.

Setup and assumptions

Model (restricted variant of toy model v3, matching the scripts exactly). A platform chooses a disclosure granularity $k \in [0, 1]$ ($k = 1$ full disclosure, $k = 0$ full masking). A coalition is *good* with probability qc and *bad* with probability $1 - qc$, where $q \in (0, 1]$ is authentication accuracy and $c \in (0, 1)$ is consensus integrity. A good coalition has robustness $r_g \sim F_c$ on $[0, 1]$ and delivers surplus $V > 0$ unless it unravels; unraveling occurs iff $r_g < \hat{r}(k)$, where $\hat{r}(k)$ is the strategic activation cutoff induced by disclosure level k , so the unraveling risk is $\rho(k, c) = F_c(\hat{r}(k))$. A bad coalition imposes loss $L > 0$ unless screened, which happens with probability $s(k)$.

Restriction (linear channels). Throughout we take the channels in the exact form used by the numerical scripts:

$$\hat{r}(k) = \tau k \quad \text{with } \tau \in (0, 1) \quad (\text{v2 Lemma 1 transparency cutoff at } k = 1), \quad s(k) = k.$$

(Toy model v3 allows general C^1 increasing channels with these endpoints; we prove the result for the linear specification only. See Remark A.1.)

Welfare:

$$W(k; q, c) = qcV(1 - F_c(\tau k)) - (1 - qc)L(1 - k). \quad (\text{W})$$

Let $K^*(c) := \arg \max_{k \in [0, 1]} W(k; q, c)$ (the dependence on q is suppressed except in part (e)). We write “full disclosure is uniquely optimal at c ” for $K^*(c) = \{1\}$ and “masking is optimal at c ” for $1 \notin K^*(c)$. Define the *masking set*

$$M := \{c \in (0, 1) : 1 \notin K^*(c)\},$$

and the *marginal-balance function* (minus the right-hand slope of W at $k = 1$)

$$\varphi(c) := qcV\tau f_c(\tau) - (1 - qc)L. \quad (\varphi)$$

$\varphi(c) > 0$ says: at full disclosure, the marginal unraveling loss $qcV\tau f_c(\tau)$ exceeds the marginal screening gain $(1 - qc)L$.

Assumptions.

(A1) (payoffs). $V, L > 0$, $\tau \in (0, 1)$, $q \in (0, 1]$, $c \in (0, 1)$; welfare is (W).

(A2) (linear channels — restriction). $\hat{r}(k) = \tau k$, $s(k) = k$, hence $\rho(k, c) = F_c(\tau k)$. This is exactly the specification of `v3.py` / `v3_shape.py`.

(A3) (family regularity). For each $c \in (0, 1)$, F_c is a CDF on $[0, 1]$ with density f_c strictly positive and continuous on $(0, 1)$, and $(r, c) \mapsto f_c(r)$ is continuous on $(0, 1) \times (0, 1)$.

(A4) (fragile-end thinness at the cutoff). There exists $\bar{k} \in (0, 1)$ with

$$\lim_{c \rightarrow 0^+} c \sup_{r \in [\tau \bar{k}, \tau]} f_c(r) = 0.$$

(At very low integrity, the robustness density of good coalitions carries vanishing c -weighted mass near the transparency cutoff τ .)

(A5) (robust-end thinness below the cutoff). $\lim_{c \rightarrow 1^-} S(c) = 0$, where $S(c) := \sup_{r \in (0, \tau]} f_c(r)$, and **either** (i) $q < 1$, **or** (ii)

$$\limsup_{c \rightarrow 1^-} \frac{cV\tau S(c)}{(1-c)L} < 1.$$

(At very high integrity, good coalitions are too robust to unravel: the density on $[0, \tau]$ vanishes — fast enough, relative to $1 - c$, if $q = 1$.)

(A6) (interior bite — non-vacuity). There exists $c_m \in (0, 1)$ with $\varphi(c_m) > 0$, i.e.

$$q c_m V \tau f_{c_m}(\tau) > (1 - q c_m) L.$$

(A7) (Beta specialization, for parts (d)). $F_c = \text{Beta}(c\eta, (1 - c)\eta)$ on $[0, 1]$ with fixed concentration $\eta > 2$ (mean c ; the family of `v2/v3` and of `v3_shape.py`). Define

$$\bar{c} := \frac{1 + \tau(\eta - 2)}{\eta} \in \left(\frac{1}{\eta}, 1\right) \quad (\text{the concavity threshold}),$$

the smallest c at which the Beta mode $(c\eta - 1)/(\eta - 2)$ reaches τ .

Note: **first-order stochastic dominance of $\{F_c\}$ in c is nowhere used**; what drives the band is the behavior of the density at and below the cutoff τ (A4, A5, A6). FOSD with fixed support is the natural way such behavior arises (Remark A.4). Lemma A.4 below proves that the Beta family satisfies A3, A4, and the first part of A5.

Proof of Proposition 1

Throughout fix $q \in (0, 1]$ and suppress it except where varied. From (W), since F_c is continuous (A3 gives a density), $W(\cdot; c)$ is continuous on the compact $[0, 1]$; $K^*(c) \neq \emptyset$. This proves part (a). For $k \in (0, 1]$, $W(\cdot; c)$ is differentiable with

$$W'(k; c) = (1 - qc)L - qcV\tau f_c(\tau k), \tag{D}$$

continuous in k on $(0, 1]$ (one-sided at $k = 1$) because f_c is continuous on $(0, 1)$ and $0 < \tau k \leq \tau < 1$. Since F_c is C^1 on $(0, 1)$, $W(\cdot; c)$ is absolutely continuous on $[0, 1]$ and for $0 \leq k \leq 1$,

$$W(1; c) - W(k; c) = (1 - qc)L(1 - k) - qcV[F_c(\tau) - F_c(\tau k)]. \tag{*}$$

((*) is also immediate by direct substitution into (W).) Note $W'(1; c) = -\varphi(c)$.

Lemma A.1 (Low integrity: full disclosure uniquely optimal). *Under A1–A4 there exists $c_1 > 0$ such that for all $c \in (0, c_1)$: $W(1; c) > W(k; c)$ for every $k \in [0, 1)$, i.e. $K^*(c) = \{1\}$.*

Proof. Let \bar{k} be as in A4 and write $\varepsilon(c) := c \sup_{r \in [\tau\bar{k}, \tau]} f_c(r) \rightarrow 0$ (A4).

Case $k \in [\bar{k}, 1)$. Then $\tau k \geq \tau\bar{k}$, so

$$F_c(\tau) - F_c(\tau k) = \int_{\tau k}^{\tau} f_c(r) dr \leq \tau(1-k) \sup_{r \in [\tau\bar{k}, \tau]} f_c(r),$$

using that the integrand is bounded by the sup over $[\tau\bar{k}, \tau] \supseteq [\tau k, \tau]$ and the interval has length $\tau(1-k)$. Substituting into (\star):

$$W(1; c) - W(k; c) \geq (1-k)[(1-qc)L - qV\tau\varepsilon(c)].$$

Since $qc \leq q \leq 1$ and $c \rightarrow 0$, $(1-qc)L \rightarrow L > 0$ while $qV\tau\varepsilon(c) \rightarrow 0$; choose $c'_1 > 0$ such that for $c < c'_1$ the bracket is strictly positive. Then $W(1) > W(k)$ for all $k \in [\bar{k}, 1)$.

Case $k \in [0, \bar{k}]$. Bound $F_c(\tau) - F_c(\tau k) \leq 1$ in (\star):

$$W(1; c) - W(k; c) \geq (1-qc)L(1-\bar{k}) - qcV,$$

which is strictly positive once $qcV < (1-qc)L(1-\bar{k})$; since the left side $\rightarrow 0$ and the right side $\rightarrow L(1-\bar{k}) > 0$ as $c \rightarrow 0$, there is $c''_1 > 0$ below which it holds.

Take $c_1 := \min(c'_1, c''_1)$. □

Lemma A.2 (High integrity: full disclosure uniquely optimal). *Under A1–A3 and A5 there exists $c_2 < 1$ such that for all $c \in (c_2, 1)$: $W'(k; c) > 0$ for every $k \in (0, 1]$, hence $W(\cdot; c)$ is strictly increasing on $[0, 1]$ and $K^*(c) = \{1\}$.*

Proof. From (D), for every $k \in (0, 1]$, $W'(k; c) \geq (1-qc)L - qcV\tau S(c)$ with $S(c) = \sup_{(0, \tau]} f_c$.

Case A5(i) ($q < 1$). $(1-qc)L \geq (1-q)L > 0$ for all c , while $qcV\tau S(c) \leq V\tau S(c) \rightarrow 0$ as $c \rightarrow 1$ (A5). So there is $c_2 < 1$ past which the lower bound is strictly positive.

Case A5(ii). $(1-qc)L \geq (1-c)L$ (since $q \leq 1$). By A5(ii) there are $\delta > 0$ and $c_2 < 1$ with $cV\tau S(c) \leq (1-\delta)(1-c)L$ for $c \in (c_2, 1)$; then $W'(k; c) \geq (1-c)L - q(1-\delta)(1-c)L \geq \delta(1-c)L > 0$.

A continuous function with strictly positive derivative on $(0, 1]$ is strictly increasing on $[0, 1]$ (absolute continuity, noted above), so the unique maximizer is $k = 1$. □

Lemma A.3 (Intermediate integrity: masking strictly optimal). *Under A1–A3 and A6: $1 \notin K^*(c_m)$.*

Proof. By (D) and A6, $W'(1; c_m) = -\varphi(c_m) < 0$. $W'(\cdot; c_m)$ is continuous on $(0, 1]$, so there is $\delta \in (0, 1)$ with $W' < 0$ on $[1-\delta, 1]$; integrating, $W(1-\delta; c_m) - W(1; c_m) = -\int_{1-\delta}^1 W' > 0$. Hence $k = 1$ is not a maximizer. □

Proof of parts (b) and (c). Part (b) is Lemmas A.1 and A.2. For (c): Lemma A.3 gives $c_m \in M$, so $M \neq \emptyset$ and $c_{10} = \inf M \leq c_m \leq \sup M = c_{hi}$. Lemma A.1 gives $M \cap (0, c_1) = \emptyset$, so $c_{10} \geq c_1 > 0$; Lemma A.2 gives $M \cap (c_2, 1) = \emptyset$, so $c_{hi} \leq c_2 < 1$. If $c < c_{10}$ then $c \notin M$ (by definition of the infimum no element of M lies below c_{10}), i.e. $1 \in K^*(c)$; symmetrically for $c > c_{hi}$. Uniqueness on $(0, c_1) \cup (c_2, 1)$ is the uniqueness in Lemmas A.1–A.2. □

Lemma A.4 (The Beta family satisfies A3, A4, and the first part of A5). *Let $F_c = \text{Beta}(a, b)$ with $a = c\eta$, $b = (1 - c)\eta$, $\eta > 2$ (A7). Then A3 holds; A4 holds for every $\bar{k} \in (0, 1)$, with $\sup_{[\tau\bar{k}, \tau]} f_c(r) \rightarrow 0$ (not merely $c \cdot \sup \rightarrow 0$); and $S(c) = \sup_{(0, \tau]} f_c(r) \rightarrow 0$ as $c \rightarrow 1^-$. Moreover the A5(ii) ratio has the explicit limit*

$$\limsup_{c \rightarrow 1^-} \frac{cV\tau S(c)}{(1-c)L} = \frac{\eta V\tau^\eta}{(1-\tau)L}. \quad (\text{A5}')$$

Proof. Write $f_c(r) = r^{a-1}(1-r)^{b-1}/B(a, b)$, $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(\eta)$ (using $a + b = \eta$). Joint continuity and positivity on $(0, 1)^2$ is standard (continuity of Γ on $(0, \infty)$ and of powers); A3 holds.

A4. Fix $\bar{k} \in (0, 1)$ and take $c < \min(1/\eta, 1/2)$, so $a < 1$ and $b > \eta/2 \geq 1$. For $r \in [\tau\bar{k}, \tau]$: $r^{a-1} \leq (\tau\bar{k})^{a-1} \leq (\tau\bar{k})^{-1}$ (exponent negative, $r \geq \tau\bar{k}$, and $(\tau\bar{k})^{a-1} \leq (\tau\bar{k})^{-1}$ since $\tau\bar{k} < 1$ and $a > 0$), and $(1-r)^{b-1} \leq 1$ ($b \geq 1$, $1-r \leq 1$). So $\sup_{[\tau\bar{k}, \tau]} f_c \leq (\tau\bar{k})^{-1}/B(a, b)$. Lower-bound B : with $m := \min_{x \in [1, 2]} \Gamma(x) > 0$, $\Gamma(a) = \Gamma(a+1)/a \geq m/a$ for $a \in (0, 1)$; and $\Gamma(b)/\Gamma(\eta) \geq \gamma_0 := \min_{x \in [\eta/2, \eta]} \Gamma(x)/\Gamma(\eta) > 0$ for $c \leq 1/2$. Hence $B(a, b) \geq m\gamma_0/(c\eta)$ and

$$\sup_{[\tau\bar{k}, \tau]} f_c(r) \leq \frac{c\eta}{m\gamma_0\tau\bar{k}} \rightarrow 0 \quad \text{as } c \rightarrow 0^+.$$

A5 first part and (A5'). Take c close enough to 1 that $a > 1$ and $b < 1$. On $(0, \tau]$: $(a-1)/r - (b-1)/(1-r) > 0$ (both terms positive), so f_c is strictly increasing there and $S(c) = f_c(\tau)$. By the same Γ bounds with the roles of a, b swapped, $B(a, b) \geq m\gamma_1/((1-c)\eta)$ with $\gamma_1 := \min_{x \in [\eta/2, \eta]} \Gamma(x)/\Gamma(\eta) > 0$, so $S(c) \leq (1-c)\eta\tau^{a-1}(1-\tau)^{b-1}/(m\gamma_1) \rightarrow 0$ (the powers stay bounded: $\tau^{a-1} \leq 1$ for $a \geq 1$, $(1-\tau)^{b-1} \leq (1-\tau)^{-1}$). For the exact limit: as $c \rightarrow 1^-$, $\Gamma(b) \sim 1/b$ ($\Gamma(b) = \Gamma(b+1)/b$, $\Gamma(b+1) \rightarrow 1$), $\Gamma(a) \rightarrow \Gamma(\eta)$, so $B(a, b) \sim 1/b = 1/((1-c)\eta)$ and

$$S(c) = f_c(\tau) \sim (1-c)\eta \frac{\tau^{\eta-1}}{1-\tau},$$

whence $cV\tau S(c)/((1-c)L) \rightarrow \eta V\tau^\eta/((1-\tau)L)$, which is (A5'). \square

(Thus under A7 with $q < 1$, A5(i) applies with no further condition; if $q = 1$, A5(ii) holds iff $\eta V\tau^\eta < (1-\tau)L$. See Remark A.3.)

Lemma A.5 (Beta: strict log-concavity of the marginal-balance ratio — part (d1)). *Under A7, $g(c) := \log[qcV\tau f_c(\tau)/((1-qc)L)]$ is strictly concave on $(0, 1)$.*

Proof. With $a = c\eta$, $b = (1-c)\eta$:

$$g(c) = \text{const} + \log c + (c\eta - 1) \log \tau + ((1-c)\eta - 1) \log(1-\tau) - \log B(c\eta, (1-c)\eta) - \log(1-qc).$$

The two power terms are affine in c (second derivative 0). Using $\log B(a, b) = \log \Gamma(c\eta) + \log \Gamma((1-c)\eta) - \log \Gamma(\eta)$ and $\frac{d^2}{dx^2} \log \Gamma(x) = \psi_1(x)$ (trigamma), the chain rule gives $\frac{d^2}{dc^2} \log B = \eta^2 \psi_1(c\eta) + \eta^2 \psi_1((1-c)\eta)$. Also $\frac{d^2}{dc^2} [-\log(1-qc)] = q^2/(1-qc)^2$ and $\frac{d^2}{dc^2} \log c = -1/c^2$. Hence

$$g''(c) = -\frac{1}{c^2} - \eta^2 \psi_1(c\eta) - \eta^2 \psi_1((1-c)\eta) + \frac{q^2}{(1-qc)^2}.$$

Two elementary facts close the sign. First, $\psi_1(x) = \sum_{n \geq 0} (x+n)^{-2} > x^{-2}$ for $x > 0$ (the $n = 0$ term alone), so $\eta^2 \psi_1((1-c)\eta) > \eta^2 / ((1-c)\eta)^2 = 1/(1-c)^2$. Second, $1 - qc \geq q(1-c)$ for $q \leq 1$ (equivalent to $1 - q \geq 0$), so $q^2 / (1 - qc)^2 \leq q^2 / (q^2(1-c)^2) = 1/(1-c)^2$. Therefore

$$g''(c) < -\frac{1}{c^2} - \eta^2 \psi_1(c\eta) - \frac{1}{(1-c)^2} + \frac{1}{(1-c)^2} = -\frac{1}{c^2} - \eta^2 \psi_1(c\eta) < 0. \quad \square$$

Since $\varphi(c) > 0 \iff g(c) > 0$ and g is strictly concave and continuous, $\{\varphi > 0\}$ is the intersection of $(0, 1)$ with an open interval; by A6 it is nonempty and contains c_m . Write it (c_-, c_+) . This proves part (d1).

Lemma A.6 (Beta: strict concavity of W in k on the upper shoulder — part (d2)). *Under A7, for every $c \in [\bar{c}, 1)$: f_c is strictly increasing on $(0, \tau)$ with $f_c(0^+) = 0$; consequently $W(\cdot; c)$ is strictly concave on $[0, 1]$, $K^*(c) = \{k^*(c)\}$ is a singleton with $k^*(c) > 0$, and $k^*(c) < 1 \iff \varphi(c) > 0$, in which case $k^*(c)$ solves the FOC $qcV\tau f_c(\tau k^*) = (1 - qc)L$.*

Proof. For $c \geq \bar{c}$, $a = c\eta \geq 1 + \tau(\eta - 2) > 1$ (definition of \bar{c} ; $\eta > 2$). On $(0, 1)$,

$$\frac{f'_c(r)}{f_c(r)} = \frac{a-1}{r} - \frac{b-1}{1-r}.$$

If $b \leq 1$, both terms are positive on $(0, \tau]$, so $f'_c > 0$ there. If $b > 1$, $f'_c(r) > 0$ iff $r < m_c := (a-1)/(a+b-2) = (c\eta-1)/(\eta-2)$, and $c \geq \bar{c} \iff m_c \geq \tau$; so $f'_c > 0$ on $(0, \tau)$ (at $r = \tau = m_c$ the derivative is 0, harmless). Also $f_c(0^+) = 0$ because $a-1 > 0$ makes $r^{a-1} \rightarrow 0$ while $(1-r)^{b-1}/B$ stays bounded near 0.

By (D), $W'(k; c) = (1 - qc)L - qcV\tau f_c(\tau k)$ is then strictly decreasing in k on $(0, 1]$ wherever $\tau k < \tau$ and non-increasing at the endpoint, with $W'(0^+; c) = (1 - qc)L > 0$. A continuous function on $[0, 1]$ whose derivative exists and is strictly decreasing on $(0, 1)$ is strictly concave; its maximizer is unique. Since $W'(0^+) > 0$, $k^*(c) > 0$. If $\varphi(c) \leq 0$, then $W'(1) = -\varphi(c) \geq 0$ and by strict monotonicity of W' , $W'(k) > 0$ for all $k < 1$: W is strictly increasing, $k^*(c) = 1$. If $\varphi(c) > 0$, then $W'(1) < 0 < W'(0^+)$, and by continuity and strict monotonicity of W' on $(0, 1]$ there is a unique root $k^* \in (0, 1)$, the unique maximizer of the strictly concave W ; the root condition is exactly the stated FOC. \square

Lemma A.7 (Sharp upper threshold and continuity at the upper edge — part (d3)). *Under A7, A5(i) or (ii), and A6, suppose $c_+ > \bar{c}$. Then $c_{hi} = c_+ < 1$, $(\max(\bar{c}, c_-), c_+) \subseteq M$, the maximizer $k^*(\cdot)$ is continuous on $(\max(\bar{c}, c_-), c_+)$, and $k^*(c) \rightarrow 1$ as $c \uparrow c_+$.*

Proof. Location of the threshold. By Lemma A.2, for $c \in (c_2, 1)$ we have $W'(1; c) > 0$, i.e. $\varphi(c) < 0$; hence $c_+ \leq c_2 < 1$. By Lemma A.6, for $c \in [\bar{c}, 1)$: $c \in M \iff \varphi(c) > 0 \iff c \in (c_-, c_+)$ (Lemma A.5). Since $c_+ > \bar{c}$, the set $M \cap [\bar{c}, 1)$ equals $[\bar{c}, 1) \cap (c_-, c_+)$, which contains the nonempty interval $(\max(\bar{c}, c_-), c_+)$; in particular it contains points arbitrarily close to c_+ from below and no point $\geq c_+$. Any other element of M lies in $(0, \bar{c}) \subset (0, c_+)$. Hence $\sup M = c_+$, i.e. $c_{hi} = c_+$.

Interior solution and continuity. For $c \in (\max(\bar{c}, c_-), c_+)$, Lemma A.6 gives the unique interior maximizer $k^*(c) \in (0, 1)$ solving $f_c(\tau k^*(c)) = \lambda(c) := (1 - qc)L/(qcV\tau)$. Continuity of k^* on this interval: if $c_n \rightarrow c$ in the interval and (by compactness, along a subsequence) $k^*(c_n) \rightarrow \tilde{k} \in [0, 1]$, two boundary cases are excluded. (i) $\tilde{k} = 0$ is impossible: for c in a compact neighborhood $[\bar{c}, c_+] \subset (0, 1)$, $a = c\eta \geq 1 + \alpha$ with $\alpha := \tau(\eta - 2) > 0$, and $B(a, b)$ is bounded

below by a positive constant (continuity and positivity of B on the compact parameter set), while $(1-r)^{b-1} \leq (1-\tau)^{b-1}$ is bounded above for $r \leq \tau$; hence $f_c(r) \leq C r^\alpha$ uniformly on that neighborhood, so $f_{c_n}(\tau k^*(c_n)) \rightarrow 0$ if $k^*(c_n) \rightarrow 0$, contradicting $f_{c_n}(\tau k^*(c_n)) = \lambda(c_n) \geq \lambda(c_+) > 0$ (λ is strictly decreasing in c , and $1 - qc_+ > 0$ because $c_+ < 1$ and $q \leq 1$). (ii) Otherwise $\tilde{k} \in (0, 1]$ and joint continuity of f (A3, point $(\tau \tilde{k}, c)$ with $\tau \tilde{k} \in (0, \tau] \subset (0, 1)$) gives $f_c(\tau \tilde{k}) = \lambda(c)$; since f_c is strictly increasing on $(0, \tau]$ (Lemma A.6) the solution is unique, so $\tilde{k} = k^*(c)$: every convergent subsequence has the same limit, proving continuity.

Closing of the band. Let $c_n \uparrow c_+$. As above, along any subsequence $k^*(c_n) \rightarrow \tilde{k} \in (0, 1]$ with $f_{c_+}(\tau \tilde{k}) = \lambda(c_+)$. But $\varphi(c_+) = 0$ (continuity of φ and c_+ is the right endpoint of $\{\varphi > 0\}$) means $f_{c_+}(\tau) = \lambda(c_+)$. Strict monotonicity of f_{c_+} on $(0, \tau]$ ($c_+ > \bar{c}$) forces $\tau \tilde{k} = \tau$, i.e. $\tilde{k} = 1$. Hence $k^*(c) \rightarrow 1$ as $c \uparrow c_{\text{hi}}$. \square

Lemma A.8 (Comparative statics in q — part (e)). *Under A1–A3, for $q < q' \leq 1$: (i) $M(q) \subseteq M(q')$; (ii) for every c , $\max K^*(c; q') \leq \min K^*(c; q)$.*

Proof. (i) Let $c \in M(q)$: there is $k < 1$ with $W(k; q, c) > W(1; q, c)$, which by (\star) is

$$qcV[F_c(\tau) - F_c(\tau k)] > (1 - qc)L(1 - k).$$

The left side is strictly increasing in q (the bracket is > 0 since $f_c > 0$ and $k < 1$) and the right side strictly decreasing in q (derivative $-cL(1 - k) < 0$). So the inequality holds at q' , giving $c \in M(q')$. The threshold monotonicities follow, when $M(q) \neq \emptyset$, from inf over a larger nonempty set being smaller and sup larger.

(ii) W has strictly decreasing differences in $(k; q)$: for $k < k'$,

$$\begin{aligned} W(k'; q, c) - W(k; q, c) &= -qcV[F_c(\tau k') - F_c(\tau k)] + (1 - qc)L(k' - k), \\ \frac{\partial}{\partial q}[W(k'; q, c) - W(k; q, c)] &= -cV[F_c(\tau k') - F_c(\tau k)] - cL(k' - k) < 0, \end{aligned}$$

both terms strictly negative (F_c strictly increasing on $(0, 1)$ by $f_c > 0$). Now the standard Topkis-type argument: take $k \in K^*(c; q)$, $k' \in K^*(c; q')$ and suppose $k' > k$. Optimality at q gives $W(k'; q) - W(k; q) \leq 0$; strictly decreasing differences give $W(k'; q') - W(k; q') < W(k'; q) - W(k; q) \leq 0$, contradicting optimality of k' at q' . Hence every element of $K^*(c; q')$ is \leq every element of $K^*(c; q)$. \square

Parts (a)–(e) of Proposition 1 are now proved: (a) the opening paragraph of the proof; (b) Lemmas A.1–A.2; (c) Lemma A.3 plus the assembly in the proof of parts (b) and (c); (d1) Lemma A.5; (d2) Lemma A.6; (d3) Lemma A.7; (e) Lemma A.8; and Lemma A.4 certifies that under A7 the family-level assumptions A3, A4 hold automatically and A5 holds with $q < 1$ (or via (A5') when $q = 1$). \square

Remarks

Remark A.1 (What was restricted relative to the numerics/memos). (i) *Linear channels* $\hat{r}(k) = \tau k$, $s(k) = k$ (A2). This is exactly the specification of the ground-truth scripts `v3.py` and `v3_shape.py`, so nothing is lost relative to the numerical claim; toy model `v3`'s prose allows general C^1 increasing channels, for which Lemmas A.1–A.3 (the band) generalize directly (replace the $\tau(1 - k)$ bounds by $\hat{r}(1) - \hat{r}(k) \leq \hat{r}'_{\max}(1 - k)$ and require s' bounded below), but

parts (d1)–(d3) use linearity through φ and the Beta algebra. (ii) *Beta family* only for the sharp-threshold parts (d); the band itself, (a)–(c), holds for any family satisfying A3–A6. (iii) *Static problem*: k is chosen once; no dynamics, no finite- N progress-bar effects (anchor paper §4 qualifier 4 — the durable large-population claim is exactly the static one proved here).

Remark A.2 (Mapping to the numerical anchors). *Provenance of the headline anchor (re-verified by re-running the scripts)*. The sequence $k^*(c) = 1, 1, 1, 0.62, 0.68, 0.75, 0.85$ is produced by `v7_unify.py` (section 0, the disclosure-facet row `argmax_axis("k, ·)`) at $q = 1.0$, on the c -grid $\{0.1, 0.3, 0.5, 0.6, 0.7, 0.8, 0.9\}$, with $(V, L, \tau, \eta) = (1, 1.2, 0.7, 8)$ and custody/specification pinned at $\gamma = \sigma = 0$ — which leaves a c -dependent effective surplus $V - \text{indet}(0, c)$ in place of V ($= 0.826$ at $c = 0.6$; `v7`'s fixed- $V' = 0.826$ `v3-check` row differs only in the tail, $\dots, 0.77, 0.87$). It is **not** a `v3_shape.py` output: that script's closest row ($q = 1.0$, same (V, L, τ, η) , $V = 1$ unshifted, c -grid $0.05, \dots, 0.95$) is $1, 1, 1, 0.51, 0.60, 0.74, 0.91$ — the same valley shape, different numbers. Crucially, $q = 1.0$ at these parameters is exactly where the proposition's upper threshold does **not** bite: A5(ii) requires $\eta V \tau^\eta / (1 - \tau) < L$, and $8 \cdot 0.7^8 / 0.3 \approx 1.537 > 1.2 = L$ (still $\approx 1.270 > 1.2$ with `v7`'s effective $V' = 0.826$). So at the anchor's own parameter point $c_{\text{hi}} < 1$ is provably false in the model, the numerics agree (k^* stays interior at high c : anchor $k^*(0.9) = 0.85$; in this paper's exact model $k^*(0.95) \approx 0.91$ at $q = 1$), and the headline anchor illustrates the band's existence (b)–(c) and the FOC shape (d2) — **not** part (d3)'s strict interiority of the upper threshold, which at $q = 1$ requires A5(ii) and fails here (Remark A.3).

The upper-threshold half, scoped to $q < 1$, is exhibited at $q = 0.95$ (`v3_shape.py`, same (V, L, τ, η) , this paper's exact model): $\bar{c} = (1 + 0.7 \cdot 6) / 8 = 0.65$; numerically $\{\varphi > 0\} \approx (0.539, 0.952)$ and $M \approx (0.495, 0.952)$. So $c_+ \approx 0.952 > \bar{c} = 0.65$ — the hypothesis of (d3) holds — and indeed $c_{\text{hi}} = c_+$ to grid precision, with $k^*(c) \rightarrow 1$ continuously from below (k^* rises $0.56, 0.60, 0.68, 0.78, 0.90$ across $c = 0.5, \dots, 0.9$). The lower edge is **discontinuous**: M begins at $c \approx 0.495$ while $\varphi > 0$ only from $c \approx 0.539$ — the global maximizer jumps from the corner $k = 1$ to an interior local maximum while $W'(1) > 0$ still holds. This is why Proposition 1 characterizes c_{hi} by the FOC ($\varphi(c_{\text{hi}}) = 0$) but c_{lo} only as $\inf M$: at the lower edge no first-order condition at the corner can locate the threshold. The q -anchor $k^*(q) = 1.00 \rightarrow 0.68$ over $q = 0.4, \dots, 1.0$ at $c = 0.7$ is again `v7_unify.py` (section 8, same `indet` shift); in this paper's exact model the row is $1.00, 1.00, 0.87, 0.74, 0.65$ — the same monotone decline, part (e)'s monotonicity. A6 at $q = 0.95$ and these parameters: $\varphi(0.7) \approx +0.70 > 0$. ✓

Remark A.3 (The $q = 1$ upper edge). Under A7 with $q = 1$, A5(ii) reduces by (A5') to $\eta V \tau^\eta < (1 - \tau)L$. At the script parameters $\eta V \tau^\eta / (1 - \tau) \approx 1.537 > L = 1.2$, so the condition **fails**: at $q = 1$ our proof does not deliver $c_{\text{hi}} < 1$, and the numerics agree — masking persists for all high c tested ($k^* \approx 0.91$ at $c = 0.95$ in this paper's exact model; $k^*(0.9) = 0.85$ in the headline anchor row, which is itself computed at $q = 1.0$, Remark A.2). To state it plainly: at $q = 1$ with the anchor parameters, $c_{\text{hi}} < 1$ is false in the model — masking persists for all c below 1 once the band opens — and that is consistent with the numerics, not contradicted by them. The strict interiority of the upper threshold is therefore unconditional for $q < 1$ and parameter-conditional at $q = 1$; since the headline anchor sits at $q = 1$ with parameters where A5(ii) fails, the $c_{\text{hi}} < 1$ half of the proposition is scoped to $q < 1$ there and is **not** claimed at the anchor point. This is a boundary of the claim, not a gap in the proof.

Remark A.4 (Role of FOSD). First-order stochastic dominance of $\{F_c\}$ in c is *not* an assumption of Proposition 1. The band is driven by the c -profile of the density at and below the unraveling cutoff τ : thin at low c (A4 — under fixed-support FOSD, low- c mass sits near 0, far below

τ), thin at high c (A5 — mass sits above τ), thick somewhere in the middle (A6 — typical robustness near τ). FOSD with fixed support is the economic mechanism that makes A4–A6 natural, and the Beta (mean- c) family instantiates it; but a non-FOSD family with the same cutoff-density profile would produce the same band. The v3 memo’s triangular counterexample ($F_c(r) = r^2/r_c^2$ on support $[0, r_c]$, $r_c = r_{\max}c$) is excluded by **A3**, **not A4**: its density $f_c(r) = 2r/r_c^2$ is identically 0 on $(r_c, 1)$, so A3’s strict positivity on $(0, 1)$ fails (and $(r, c) \mapsto f_c(r)$ jumps from $2/r_c$ to 0 at the moving support edge $r = r_c$, so A3’s joint continuity fails too). The A4 limit, by contrast, **holds** — vacuously: for $c < \tau\bar{k}/r_{\max}$ the entire support $[0, r_{\max}c]$ lies below the window $[\tau\bar{k}, \tau]$, so $\sup_{[\tau\bar{k}, \tau]} f_c = 0$ and $c \cdot \sup_{[\tau\bar{k}, \tau]} f_c = 0 \rightarrow 0$. (An earlier version of this remark misattributed the exclusion to A4.) The moving support $r_c \propto c$ is what inverts the within-family shape $k^*(c)$, and it is precisely what A3’s full-support positivity rules out.

Conjectures. Conjecture 1 in the main text (connectedness of M) states that M is an interval, i.e. $M = (c_{\text{lo}}, c_{\text{hi}})$ exactly. Proved here: $M \subseteq [c_{\text{lo}}, c_{\text{hi}}] \subset (0, 1)$ with masking at c_m and, under (d3), the entire upper shoulder $(\max(\bar{c}, c_-), c_+) \subseteq M$. Not proved: that M has no gaps below \bar{c} . Numerically M is an interval at all tested parameters (grid check, `v3_shape.py` family). What would close it: quasi-concavity in c of $c \mapsto \sup_{k < 1} [W(k; c) - W(1; c)]$; the natural attack (Prékopa / joint log-concavity of $(c, r) \mapsto f_c(r)$) fails because $\log f_c(r)$ contains bilinear terms $c\eta \log r$ whose Hessian is indefinite, and a sup of quasi-concave functions need not be quasi-concave. A workable route may be the variation-diminishing property of the Beta kernel (sign-regularity of $(c, r) \mapsto f_c(r)$), which would bound the sign changes of $c \mapsto W(k; c) - W(1; c)$ uniformly in k .

Conjecture 5 (Monotone reopening). *On $(\max(\bar{c}, c_-), c_{\text{hi}})$, $k^*(c)$ is strictly increasing (the band closes monotonically).*

The FOC gives $dk^*/dc \propto \partial/\partial c [qcVf_c(\tau k)] - (-qL)$ -type expressions whose sign requires $\partial f_c(r)/\partial c$ at $r = \tau k^*$, not signed in general. Numerically robust at all tested parameters.

Remark A.5 (Interpretation of the FOC). Part (d2)’s first-order condition $qcV\tau f_c(\tau k^*) = (1 - qc)L$ is the v3 memo’s “marginal unraveling loss = marginal screening gain” line, with the linear-channel substitution made explicit; $\varphi(c)$ is this balance evaluated at full disclosure, and the band is precisely the integrity range where, at $k = 1$, protecting good coalitions at the margin is worth more than the last unit of screening.

B Proof of Proposition 2

Setup

Provenance. This appendix formalizes toy-model v9 §2 (S10). Ground truth for the game, the CvD belief, and the advantage decomposition: `runs/numcheck/v9_selection_lambdafree.py` (v9.S10.2, the p_{opp} -consistent CvD belief) and the adversarial re-implementation `runs/numcheck/v9_verify_s10`. Selection license: Frankel–Morris–Pauzner (2003, *JET* 108(1):1–44), §6.4 / Thm 4 — scope verified and qualified in Setup §0.4 and Remark R5. No step of the proof appeals to the numerics; script output is used only in the Remarks to map anchors.

0.1 The game (exact formalization of the script kernel). Fix integers $N \geq 2$ (agents), $1 \leq K \leq N$ (clearing threshold), $T \geq 1$ (horizon), and reals $p \in (0, 1)$ (move probability), $V > \kappa > 0$ (cleared value, commitment cost), $L_p \geq 0$ (stranded-commitment loss), $\rho \geq 0$ (per-period decision cost), $\xi \in [0, 1]$ (non-excludable share). Write $g(m) := K - m$ (the *gap*) and $n(m) := N - m$ (uncommitted agents) at support count m .

State and transitions. The state is (m, h) : m agents committed, h periods remaining. In a period at state (m, h) with $m < K$, each of the $n = n(m)$ uncommitted agents independently receives a move opportunity with probability p and, if moving, plays the common selected commit propensity $a(m, h) \in [0, 1]$; the realized number of new commits is $j \sim \text{Binom}(n, p \cdot a(m, h))$ and the next state is $(m + j, h - 1)$. The campaign *clears* when $m \geq K$ (absorbing); it *fails* if h reaches 0 with $m < K$.

Terminal payoffs (per agent, as in v6/v7/v9): commit & clear $V - \kappa$; commit & fail $-L_p$; wait & clear ξV ; wait & fail 0. Each period in which an agent (counterfactually) evaluates her move costs ρ (this flow enters both action values identically and cancels in every comparison below; it survives only inside continuation values).

Profile-consistent arrays. Given a selected propensity table $a(\cdot, \cdot)$, define by the script’s `propagate` recursion (backward induction in h ; all references are to row $h - 1$, so the recursion is well-founded):

- $P_c[m, h]$ — clearing probability: $P_c[m, h] = 1$ for $m \geq K$; $P_c[m, 0] = 0$ for $m < K$; and for $m < K$, $h \geq 1$, with $b = p \cdot a(m, h)$,

$$P_c[m, h] = \sum_{j=0}^n \text{Binom}(n, b)(j) \cdot (\mathbf{1}\{m + j \geq K\} + \mathbf{1}\{m + j < K\} \cdot P_c[m + j, h - 1]).$$

- $V_u[m, h]$ — value of an uncommitted agent under the selected profile: $V_u[\cdot, 0] = 0$, and for $m < K$, $h \geq 1$, $V_u[m, h] = a \cdot v_c + (1 - a) \cdot v_w$, where v_c, v_w are the commit/wait values computed against co-movers $j \sim \text{Binom}(n - 1, b)$ and continuations $P_c[\cdot, h - 1]$, $V_u[\cdot, h - 1]$ (script lines: `propagate`). In the masked game $a \in \{0, 1\}$, so $V_u[m, h]$ is the selected branch — v_c if $a = 1$, v_w if $a = 0$ — evaluated at the selected-profile co-mover law $\text{Binom}(n - 1, b)$. The action a is the argmax of the §0.2 Laplacian comparison, not of this profile-law pair, so $V_u[m, h]$ can lie strictly below $\max(v_c, v_w)$ (e.g. at the v9 calibration, $\theta = 0.30$, state $(6, 1)$: commit is selected while the profile-law values are $v_c = 1.950 < v_w = 2.196$). No lemma uses a $V_u = \max$ identity; Lemma B.5 propagates the selected branch directly.

The outcome of interest is $\pi := P_c[0, T]$.

0.2 The masked selection (λ -free CvD/Laplacian rule). At each (m, h) with $m < K$, $h \geq 1$, the masked mover knows only (m, h) and evaluates commit-vs-wait under the p -consistent Carlsson–van Damme (Laplacian) belief: the others’ commit propensity is $\alpha \sim \text{Uniform}[0, 1]$ and the realized co-mover count is $j \sim \text{Binom}(n - 1, p\alpha)$. The stage values are (script `laplacian_vals`, idealized to the exact integral; see Remark R1 on the script’s 51-point quadrature):

$$v_c(m, h) = -\rho + \sum_j q_j(n) \cdot U_c(j), \quad v_w(m, h) = -\rho + \sum_j q_j(n) \cdot U_w(j),$$

$$U_c(j) = \begin{cases} V - \kappa & \text{if } m + 1 + j \geq K, \\ (V - \kappa) \cdot P_c[m + 1 + j, h - 1] - L_p \cdot (1 - P_c[m + 1 + j, h - 1]) & \text{otherwise,} \end{cases}$$

$$U_w(j) = \begin{cases} \xi V & \text{if } m + j \geq K, \\ 0 & \text{if } m + j < K \text{ and } h - 1 = 0, \\ V_u[m + j, h - 1] & \text{otherwise,} \end{cases}$$

where $q_j(n) := \int_0^1 \text{Binom}(n - 1, p\alpha)(j) d\alpha$ is the Laplacian co-mover law. The masked rule selects $a(m, h) = 1$ iff $v_c > v_w$, else 0 (ties \rightarrow wait, as in the script). Continuations P_c, V_u are then propagated under the *selected* propensity (§0.1) — the Laplacian doubt applies to the current period’s co-movers only; continuation values are model-consistent. Write $A[m, h]$ for the selected action and $\pi_m := P_c[0, T]$ under this rule (π_m , the *masked* clearing probability, abbreviates the draft’s π_{masked}).

0.3 The revealed selection (pivotality revelation, v7 reduced form). In the revealed partition the mover additionally learns her decisiveness. As in v7/v9 this is operationalized (not derived from an extensive form — see Remark R2) by the blend (script `solve_revealed_rd`): at (m, h) with gap $g = K - m$,

- **pivotal action** $a_{\text{marg}} = \mathbf{1}\{V - \kappa > \text{wait_cont}\}$, $\text{wait_cont} = V_u[m, h - 1]$ (0 if $h = 1$);
- **inframarginal action** $a_{\text{inf}} = 0$ iff free-riding pays, i.e. iff $\xi V > V - \kappa$ (else $a_{\text{inf}} = 1$);
- **pivotality weight** $w(m, h) = \text{clip}(1 - (g - 1) / \max(1, pm), [0, 1])$;
- selected propensity $a(m, h) = w \cdot a_{\text{marg}} + (1 - w) \cdot a_{\text{inf}}$, propagated by the same kernel.

Write $\pi_r := P_c[0, T]$ under this rule (π_r , the *revealed* clearing probability, abbreviates the draft’s π_{rev}). No object in §0.1–§0.3 contains a rationality parameter: the selection is λ -free by construction, and for fixed primitives each of π_m, π_r is a single number.

0.4 Selection license (FMP §6.4), verified and qualified. Each stage decision in §0.2 is a symmetric, **two-action**, n -player game. Frankel–Morris–Pauzner (2003) §6.4 characterizes the noise-independent global-game selection in symmetric two-action many-player games as the action that is a best response to a **uniform belief over the opponents’ aggregate action** — the Laplacian criterion used verbatim in §0.2, made p -consistent by routing the uniform propensity through the $\text{Binom}(n - 1, p\alpha)$ movement kernel. Their conditions, checked for this class:

- *Own-action quasiconcavity*: automatic — with a two-point action set every payoff function is own-action quasiconcave, so FMP Thm 4’s quasiconcavity hypothesis is vacuous.
- *LP-maximizer*: FMP §6.4 (via Lemma 2 / Morris–Ui p -dominance) identifies the Laplacian action as the LP-maximizer **under strategic complementarities**. Our stage differentials $U_c(j) - U_w(j)$ are *not* globally increasing in j : at diagonal states (Def. §0.6) the stage game is volunteer’s-dilemma-like (substitutes). Where complementarities hold, the Laplacian action is the FMP-licensed noise-independent selection; where they fail, we adopt the Laplacian action as the *definition* of stagewise risk dominance (the CvD vanishing-noise criterion), exactly as the ground-truth script does. This is a refinement *choice*, cross-validated by an independent λ -free refinement (potential-maximizer, v9_verify_s10.py) that reproduces the same regime split. See Remark R5.

0.5 Assumptions.

- **(A1) (primitives)**. $N \geq 2$, $2 \leq K \leq N$, $T \geq K$, $p \in (0, 1)$, $V > \kappa > 0$, $L_p \geq 0$, $\rho \geq 0$, $\xi \in [0, 1]$.
- **(A2) (kernel)**. The dynamic game and profile-consistent arrays of §0.1.
- **(A3) (masked rule)**. The exact-integral CvD/Laplacian stage selection of §0.2.
- **(A4) (revealed rule)**. The pivotality-blend reduced form of §0.3.
- **(A5) (free-riding pays)**. $\xi V > V - \kappa$, i.e. $\xi > 1 - \kappa/V$.
- **(A6) (contested regime)**. $K - 1 \geq \max(1, pN)$.
- **(A7) (deadline ignition)**. $\xi V \cdot (1 - q_0(N - K + 1)) < V - \kappa$, where $q_0(n) := (1 - (1 - p)^n)/(np)$.
- **(A8) (rung-2 ignition, used only for the $K = 2$ part)**. With $q_j := q_j(N)$ of Lemma B.1 and $\hat{P}(1) := 1 - (1 - p)^{N-1}$:
 - (C2)** $q_0 \cdot [(V - \kappa + L_p) \cdot \hat{P}(1) - L_p + \rho] + q_1 \cdot \rho > (\xi V - (V - \kappa)) \cdot (1 - q_0 - q_1)$;
 - (C3)** $L_p \cdot q_0 + \xi V \cdot (1 - q_0 - q_1) \geq (V - \kappa) \cdot (1 - q_0)$.

A5–A6 are the **contested-regime** hypotheses (the θ -threshold of part (iv)); A7 is the **momentum-ignition** hypothesis (the θ -threshold of part (ii)). With the v9 grid map $K(\theta) = \lceil (1 - \theta) \cdot N \rceil$, A6 reads: θ lies below the threshold at which the initial gap exceeds the expected per-period mover count (at the v9 calibration, where $pN = 6$ is an integer: $\theta < 1 - p = 0.4$, i.e. exactly $\theta \in \{0.10, 0.20, 0.30\}$ on the grid).

0.6 Definitions.

- **Diagonal / cone**. State (m, h) with $m < K$ is *on the diagonal* if $h = g(m)$, *in the cone* if $h \geq g(m)$, *below the diagonal* if $h < g(m)$.
- **DI (diagonal ignition)**. The masked rule selects commit at every diagonal state: $A[K - g, g] = 1$ for $g = 1, \dots, K$. (Well-defined since $T \geq K$ by A1.)

- **Momentum bound.** $\widehat{P}(\gamma) := \prod_{i=1}^{\gamma} (1 - (1-p)^{N-K+i})$ for $\gamma \geq 0$ ($\widehat{P}(0) := 1$). Each factor lies in $(0, 1)$, so \widehat{P} is strictly decreasing in γ and $\widehat{P}(K) > 0$.
- **Advantage.** $\text{Adv} := \pi_m - \pi_r$.

Proposition 2 (restated). *Under A1–A3 the masked game and under A1, A2, A4 the revealed game are well-defined, and no rationality parameter appears anywhere; for fixed primitives Adv is a single number (no λ -inversion is possible). Moreover:*

- (i) (Momentum, conditional form.) *If DI holds, then $\pi_m \geq \widehat{P}(K) > 0$.*
- (ii) (Ignition at the last rung; θ -threshold.) *Under A7, the masked rule commits at the deadline-pivotal diagonal state: $A[K-1, 1] = 1$. Since $q_0(n)$ is strictly decreasing in n , A7 is downward-closed in θ : if it holds at threshold K it holds at every $K' \geq K$, i.e. for every harder (lower- θ) market.*
- (iii) (Momentum, unconditional for $K = 2$.) *Let $K = 2$ and assume A5, A7, A8. Then DI holds — $A[1, 1] = 1$ and $A[0, 2] = 1$ — and therefore $\pi_m \geq \widehat{P}(2) = (1 - (1-p)^{N-1})(1 - (1-p)^N) > 0$. The hypothesis set is non-vacuous: $(N, K, T, p) = (10, 2, 10, 1/10)$ with the v9 payoffs $(V, \kappa, L_p, \rho, \xi) = (3, 1, 3, 1/20, 4/5)$ satisfies A1, A5, A6, A7, A8 simultaneously (exact margins in the proof, Step 5).*
- (iv) (Revealed collapse in the contested regime.) *Under A1, A2, A4, A5, A6: $\pi_r = 0$ exactly. In particular at the v9 calibration $(N, p) = (10, 3/5)$, A6 holds iff $K \geq 7$, i.e. iff $\theta \in \{0.10, 0.20, 0.30\}$ on the v9 grid.*
- (v) (Masking advantage.) *Under the hypotheses of (iv) together with DI (or, for $K = 2$, with A7–A8 instead of DI): $\text{Adv} = \pi_m \geq \widehat{P}(K) > 0$ — strictly positive on the contested regime. Globally (without A5–A6), under DI: $\text{Adv} \geq \widehat{P}(K) - 1$. (No claim of $\text{Adv} \geq 0$ outside the contested regime is made: the ground-truth computation exhibits $\text{Adv} \in [-8 \cdot 10^{-4}, 0)$ on the easy regime; see Remark R4.)*

The diagonal-ignition conjecture (Conjecture 2) accompanies the proposition; see Remark R3 for its precise status and obstruction.

Proof of Proposition 2

Throughout, $\Delta_{\text{FR}} := \xi V - (V - \kappa)$ (> 0 under A5) is the free-ride premium, and $q_j(n)$ is the Laplacian co-mover law of §0.2 at a state with n uncommitted agents.

Lemma B.1 (Laplacian belief identity). *For $0 \leq j \leq n-1$,*

$$q_j(n) = \frac{1}{np} \cdot \Pr(\text{Binom}(n, p) \geq j+1).$$

Consequently $q_0(n) = (1 - (1-p)^n)/(np)$, $\sum_{j=0}^{n-1} q_j(n) = 1$, and $q_j(n)$ is strictly decreasing in j .

Proof. Substituting $b = p\alpha$,

$$q_j(n) = \int_0^1 \binom{n-1}{j} (p\alpha)^j (1-p\alpha)^{n-1-j} d\alpha = \frac{1}{p} \int_0^p \binom{n-1}{j} b^j (1-b)^{n-1-j} db.$$

The integrand is $1/n$ times the Beta($j+1, n-j$) density, so the integral equals $\frac{1}{np} \cdot \Pr(\text{Beta}(j+1, n-j) \leq p)$, and the Beta–Binomial identity $\Pr(\text{Beta}(j+1, n-j) \leq p) = \Pr(\text{Binom}(n, p) \geq j+1)$ gives the display. The normalization follows from $\sum_{j=0}^{n-1} \Pr(\text{Binom}(n, p) \geq j+1) = \mathbb{E}[\text{Binom}(n, p)] = np$; monotonicity in j is monotonicity of the Binomial upper tail; q_0 follows from $\Pr(\text{Binom}(n, p) \geq 1) = 1 - (1-p)^n$. \square

(All stage quantities below are therefore finite rational expressions in the primitives when p is rational; the instance margins in Step 5 are evaluated in exact arithmetic.)

Lemma B.2 (q_0 monotonicity — for part (ii)’s threshold form). $q_0(n)$ is strictly decreasing in n .

Proof. With $x := 1-p \in (0, 1)$,

$$q_0(n) = \frac{1-x^n}{np} = \frac{1-x}{p} \cdot \frac{1}{n} \sum_{i=0}^{n-1} x^i,$$

an average of the strictly decreasing sequence x^i ; averages of strictly decreasing sequences strictly decrease as terms are appended. \square

Lemma B.3 (Cone lower bound: DI \Rightarrow momentum — proves part (i)). Assume A1–A3 and DI. Then for every state (m, h) in the cone ($m < K, g(m) \leq h \leq T$),

$$P_c[m, h] \geq \widehat{P}(g(m)).$$

Proof. Induction on h .

Base $h = 1$. The only cone states have $g = 1$, i.e. $m = K - 1$ (diagonal). DI gives $A[K - 1, 1] = 1$, so by §0.1 with $b = p$ and $n = N - K + 1$:

$$P_c[K - 1, 1] = \Pr(\text{Binom}(n, p) \geq 1) = 1 - (1-p)^{N-K+1} = \widehat{P}(1).$$

Step. Let $h \geq 2$ and assume the claim for $h - 1$. Fix a cone state (m, h) , $g := g(m)$, $n := n(m) = N - K + g$. Two cases by the selected action $a := A[m, h] \in \{0, 1\}$:

- $a = 0$. Then $b = 0$, so $j = 0$ a.s. and $P_c[m, h] = P_c[m, h - 1]$. By DI, $a = 0$ is impossible at a diagonal state, so $h > g$, hence $h - 1 \geq g$ and $(m, h - 1)$ is in the cone; the inductive hypothesis gives $P_c[m, h] \geq \widehat{P}(g)$.
- $a = 1$. Then $j \sim \text{Binom}(n, p)$ and

$$P_c[m, h] = \sum_j \text{Binom}(n, p)(j) \cdot (\mathbf{1}\{j \geq g\} + \mathbf{1}\{j < g\} \cdot P_c[m + j, h - 1]).$$

For each $j \geq 1$ with $j < g$, the successor $(m + j, h - 1)$ has gap $g - j \leq g - 1 \leq h - 1$, so it is in the cone and $P_c[m + j, h - 1] \geq \widehat{P}(g - j) \geq \widehat{P}(g - 1)$ (monotonicity of \widehat{P}); for $j \geq g$ the summand is $1 \geq \widehat{P}(g - 1)$. Discarding the $j = 0$ term (bounded below by 0),

$$P_c[m, h] \geq \Pr(\text{Binom}(n, p) \geq 1) \cdot \widehat{P}(g - 1) = (1 - (1-p)^{N-K+g}) \cdot \widehat{P}(g - 1) = \widehat{P}(g).$$

Since $T \geq K$ (A1), $(0, T)$ is in the cone and $\pi_m = P_c[0, T] \geq \widehat{P}(K) > 0$. \square

Lemma B.4 (Deadline ignition — proves part (ii)). *Assume A1–A3 and A7. Then $A[K - 1, 1] = 1$.*

Proof. At $(K - 1, 1)$, $n = N - K + 1$, $h - 1 = 0$. For every $j \geq 0$, $m + 1 + j = K + j \geq K$, so $U_c(j) = V - \kappa$ identically. For $j \geq 1$, $m + j \geq K$, so $U_w(j) = \xi V$; for $j = 0$, $m + j = K - 1 < K$ and $h - 1 = 0$, so $U_w(0) = 0$. Hence (the $-\rho$ flows cancel)

$$v_c - v_w = (V - \kappa) - (1 - q_0(n)) \cdot \xi V,$$

which is strictly positive iff A7 holds; the rule then selects commit. The threshold form follows from Lemma B.2: the left side of A7 is decreasing in K , so A7 propagates from K to all $K' \geq K$, i.e. down the θ -grid. \square

Lemma B.5 (Rung-2 ignition for $K = 2$ — with Lemma B.4, proves part (iii)). *Assume A1–A3, A5, A7, A8 and $K = 2$. Then $A[0, 1] = 0$, $V_u[0, 1] = -\rho$, $A[1, 1] = 1$, $P_c[1, 1] = \widehat{P}(1) = 1 - (1 - p)^{N-1}$, $V_u[1, 1] = V - \kappa - \rho$, and $A[0, 2] = 1$.*

Proof. All five intermediate equalities are exact computations of the recursion at $h = 1$, in dependency order.

State (1, 1) (diagonal, $g = 1$, $n = N - 1$). Lemma B.4 applies with $K = 2$ (its proof used only the structure at $(K - 1, 1)$): $A[1, 1] = 1$ by A7. Propagation with $a = 1$, $b = p$: every commit branch clears ($m + 1 + j = 2 + j \geq K$), so $v_c = -\rho + (V - \kappa)$ and $V_u[1, 1] = V - \kappa - \rho$; and $P_c[1, 1] = \Pr(\text{Binom}(N - 1, p) \geq 1) = 1 - (1 - p)^{N-1} = \widehat{P}(1)$.

State (0, 1) (below the diagonal: $g = 2 > h = 1$; $n = N$). Stage values: $U_c(j)$: for $j \geq 1$, $m + 1 + j \geq 2$: $V - \kappa$; for $j = 0$, successor $(1, 0)$ has $P_c[1, 0] = 0$, so $U_c(0) = -L_p$. $U_w(j)$: for $j \geq 2$, $m + j \geq 2$: ξV ; for $j \in \{0, 1\}$, $m + j < K$ and $h - 1 = 0$: 0. Hence

$$v_c - v_w = -L_p \cdot q_0 + (V - \kappa)(1 - q_0) - \xi V \cdot (1 - q_0 - q_1),$$

with $q_j = q_j(N)$. By (C3) this is ≤ 0 , so the rule selects wait: $A[0, 1] = 0$. Propagation with $a = 0$: $b = 0$, $j = 0$ a.s., the wait branch hits $h - 1 = 0$ with $m < K$, so $V_u[0, 1] = -\rho + 0 = -\rho$ (and $P_c[0, 1] = P_c[0, 0] = 0$).

State (0, 2) (diagonal, $g = 2$, $n = N$). Stage values with $h - 1 = 1$, using the entries just computed:

- $j = 0$: $U_c(0) = (V - \kappa) \cdot P_c[1, 1] - L_p \cdot (1 - P_c[1, 1]) = (V - \kappa + L_p) \cdot \widehat{P}(1) - L_p$;
 $U_w(0) = V_u[0, 1] = -\rho$.
- $j = 1$: $m + 1 + j = 2 \geq K$: $U_c(1) = V - \kappa$; $U_w(1) = V_u[1, 1] = V - \kappa - \rho$.
- $j \geq 2$: both branches clear: $U_c - U_w = (V - \kappa) - \xi V = -\Delta_{\text{FR}}$.

Hence

$$v_c - v_w = q_0 \cdot [(V - \kappa + L_p) \cdot \widehat{P}(1) - L_p + \rho] + q_1 \cdot \rho - \Delta_{\text{FR}} \cdot (1 - q_0 - q_1),$$

which is strictly positive by (C2); the rule selects commit: $A[0, 2] = 1$.

For $K = 2$ the diagonal is exactly $\{(1, 1), (0, 2)\}$, so DI holds, and Lemma B.3 gives $\pi_m \geq \widehat{P}(2)$. \square

Step 5 (non-vacuousness of part (iii)). Take $(N, K, T, p) = (10, 2, 10, 1/10)$ and $(V, \kappa, L_p, \rho, \xi) = (3, 1, 3, 1/20, 4/5)$. Then, in exact rational arithmetic (all quantities are the closed forms above; Lemma B.1 reduces every q_j to a Binomial tail):

- A1: $2 \leq 2 \leq 10, T = 10 \geq K = 2$. ✓
- A5: $\xi V = 12/5 > V - \kappa = 2$. ✓
- A6: $K - 1 = 1 \geq \max(1, pN) = \max(1, 1) = 1$. ✓
- A7 at $n = N - K + 1 = 9$: $q_0(9) = 612579511/900000000 \approx 0.68064$; margin

$$v_c - v_w = (V - \kappa) - (1 - q_0) \xi V = 462579511/375000000 \approx +1.2335 > 0. \checkmark$$

- (C3) at $n = 10$: $q_0 \approx 0.65132, q_1 \approx 0.26390, 1 - q_0 - q_1 \approx 0.08478$; margin

$$v_c - v_w = -73003674279/50000000000 \approx -1.4601 \leq 0. \checkmark$$

- (C2): $\widehat{P}(1) = 1 - (9/10)^9 = 612579511/10^9 \approx 0.61258$; $(V - \kappa + L_p) \cdot \widehat{P}(1) - L_p + \rho = 12579511/200000000 + 1/20 \approx 0.11290$; margin

$$v_c - v_w = 105633434992992089/(2 \cdot 10^{18}) \approx +0.05282 > 0. \checkmark$$

So DI holds and $\pi_m \geq \widehat{P}(2) = (1 - 0.9^9)(1 - 0.9^{10}) = 0.39898 \dots$, while A5–A6 put the instance in the contested regime, so by part (iv) below $\pi_r = 0$ and $\text{Adv} \geq \widehat{P}(2) > 0$ *unconditionally* for this instance class. \square

Step 6 (revealed collapse — proves part (iv)). Assume A1, A2, A4, A5, A6. At every state $(0, h)$, $h = 1, \dots, T$: the gap is K , $n = N$, and the pivotality weight is

$$w(0, h) = \text{clip}(1 - (K - 1)/\max(1, pN), [0, 1]) = 0,$$

since $K - 1 \geq \max(1, pN)$ by A6. By A5, $a_{\text{inf}} = 0$. Hence the selected propensity is $a(0, h) = w \cdot a_{\text{marg}} + (1 - w) \cdot a_{\text{inf}} = a_{\text{inf}} = 0$, *regardless of a_{marg}* . Propagation with $b = p \cdot 0 = 0$ gives $j = 0$ a.s., so $P_c[0, h] = P_c[0, h - 1]$ for every $h \geq 1$, and $P_c[0, 0] = 0$ (since $K \geq 1$). By induction $\pi_r = P_c[0, T] = 0$ exactly. (States with $m \geq 1$ never enter the computation of $P_c[0, \cdot]$ because the transition from $m = 0$ is degenerate at $j = 0$.) \square

Interpretation, exactly as claimed: under A6 the initial gap exceeds the expected mover count, the pivotality weight classifies every mover as inframarginal, revelation of that inframarginal status triggers the free-ride action ($a_{\text{inf}} = 0$, optimal because $\xi V > V - \kappa$ by A5), and clearing collapses to zero from the start — a volunteer’s dilemma detonated by pivotality information.

Step 7 (advantage — proves part (v)). On the contested regime (A5, A6), part (iv) gives $\pi_r = 0$, so $\text{Adv} = \pi_m$. Under DI (or under A7–A8 when $K = 2$), Lemma B.3 (resp. Lemma B.5) gives $\pi_m \geq \widehat{P}(K) > 0$, hence $\text{Adv} \geq \widehat{P}(K) > 0$. Globally, $\pi_r \leq 1$ always, so under DI $\text{Adv} \geq \widehat{P}(K) - 1$. λ -freeness and no-inversion are definitional: no object in §0.1–§0.3 contains a rationality parameter, and for fixed primitives Adv is one number, so there is no parameter along which it could invert (contrast v7 S1’s logit, which inverted at high λ). \square

Remarks

Remark B.1 (R1: what was restricted relative to the numerics — masked side). The proposition idealizes the script’s Laplacian quadrature (equal-weight 51-point grid on $[0, 1]$, `laplacian_vals`) to the exact integral; Lemma B.1 makes every stage quantity closed-form. The script’s rule is an equal-weight, endpoint-inclusive 51-point average, whose generic error is $O(1/n_a)$ (a first-order average, not a higher-order rule). Against the $O(10^{-1})$ stage margins at the lemma-relevant states this is harmless, but it is not action-invariant everywhere: at the calibration the idealization flips the selected action at four above-diagonal cone cells where the stage margin is small on at least one side — $\theta = 0.20$, $(m, h) = (4, 9)$ (quadrature $+0.01259$ commit vs exact -0.00036 wait) and $(5, 8)$ (-0.00120 wait vs $+0.00024$ commit); $\theta = 0.30$, $(3, 5)$ (-0.00118 wait vs $+0.00010$ commit); $\theta = 0.50$, $(2, 9)$ (-0.00179 wait vs $+0.00106$ commit). The only outcome-level effect is at $\theta = 0.20$: $\pi_m = 0.9894$ (script) vs 0.9890 (exact rule), a gap $\leq 4 \cdot 10^{-4}$. All lemma-verified actions — every diagonal rung, every below-diagonal wait, and the $K = 2$ instance of Step 5 — and DI itself are identical under both rules, so no proved claim is affected. The masked *policy* is genuinely not “all-commit”: the ground-truth tables show commit on and above the just-in-time diagonal $h = g$ at large gaps, free-riding above the diagonal at small gaps, and wait below the diagonal. The proof is engineered around this: Lemma B.3 needs *only* diagonal ignition (DI) and is agnostic about every off-diagonal action — that is what makes the conditional momentum theorem clean.

Remark B.2 (R2: what was restricted — revealed side). Part (iv) is a theorem about the v7/v9 *reduced-form* revealed game: pivotality revelation is operationalized by the deterministic weight $w(m, h)$ blending the pivotal and inframarginal actions, not derived as the Bayesian equilibrium of an extensive form in which movers observe their decisiveness. The collapse proof is exact for that reduced form (and matches the script to all printed digits: $\pi_r = 0.0000$ at $\theta \in \{0.10, 0.20, 0.30\}$). Deriving the same collapse from a fully specified revealed extensive form is open (it is the same gap v7 §9 flagged; S10 inherits it deliberately, changing only the selection rule).

Remark B.3 (R3: the remaining gap — DI at general K is a conjecture). DI is *proven* at the last rung for every K satisfying A7 (Lemma B.4: at the v9 calibration the exact margins are $+0.5744$, $+0.8480$, $+1.2800$ for $K = 7, 8, 9$, and positive down to $K = 3$), and at both rungs for $K = 2$ (Lemma B.5). At deeper rungs ($g \geq 2$, $K \geq 3$) DI is the diagonal-ignition conjecture (Conjecture 2): verified pointwise at all seven grid θ ’s (the ground-truth solver returns $A[K - g, g] = 1$ for every rung at every grid point, and all-wait below the diagonal), but not proven. The precise obstruction: at a diagonal state the wait branch with $j \geq 1$ co-movers lands on-cone, where the only easy bound on the uncommitted value is $V_u \leq \xi V$; commit beats that crude bound only via the $j = 0$ (pivotal) Laplacian mass $q_0 \approx 1/(np)$, and the resulting sufficient condition fails at the calibration even though the exact inequality holds with ample margin: at the deepest rung ($(m, h) = (0, 7)$, $\theta = 0.30$) the diagonal stage margin $v_c - v_w$ is $+0.417$ under the exact-integral rule of A3 ($+0.432$ under the script’s quadrature), and the shallower $\theta = 0.30$ rungs are larger still (up to $+0.574$ at $g = 1$). Closing it needs $V_u(\text{on-cone}) \leq \xi V \cdot P_c + o(1 - P_c)$ together with gap-indexed momentum bounds — i.e. a simultaneous induction on (DI, below-diagonal freeze, on-cone value bounds). A theorem plus this labeled conjecture is the deliverable; we do not pretend the induction closes.

Remark B.4 (R4: mapping to the numerical anchors; correction to the “ ≥ 0 everywhere” claim). Ground truth (v9_selection_lambdafree.py, §1 table), against the proposition’s quantities,

with $K(\theta) = \lceil(1 - \theta)N\rceil$:

θ	K	A6 (contested)?	$\widehat{P}(K)$ (proved bound under DI)	script π_m	script π_r	script Adv
0.10	9	yes	0.7532	0.9906	0.0000	+0.9906
0.20	8	yes	0.8966	0.9894	0.0000	+0.9894
0.30	7	yes	0.9579	0.9924	0.0000	+0.9924
0.40	6	no	0.9831	0.9978	0.9986	-0.0008
0.50	5	no	0.9933	0.9995	1.0000	-0.0005
0.60	4	no	0.9973	0.9999	1.0000	-0.0001
0.75	3	no	0.9990	0.9999	1.0000	-0.0001

The anchors map as: “ $\pi_m \approx 0.99$ at every θ ” = DI + Lemma B.3 (the bound $\widehat{P}(K)$ is a true but conservative floor; the slack is the $j = 0$ branch discarded in the induction and the above-diagonal commits the bound ignores); “advantage +0.99 at $\theta \in \{0.10, 0.20, 0.30\}$ ” = part (v) with part (iv)’s exact zero; “|Adv| < 0.01 elsewhere” = the global bound $\text{Adv} \geq \widehat{P}(K) - 1 \in [-0.017, -0.001]$ plus $\pi_r \leq 1$. **Correction:** the task-statement form “advantage ≥ 0 everywhere” is *falsified* by the ground truth — the easy-regime advantage is slightly negative ($-8 \cdot 10^{-4}$ at $\theta = 0.40$). The correct λ -free claim, which this proposition states and the script verifies, is: *strictly positive on the contested regime; bounded below by $-(1 - \widehat{P}(K))$ (numerically $\leq 10^{-3}$ in magnitude, a tie) elsewhere; no inversion anywhere because there is no parameter to invert in.* The table quotes the script’s 51-point-quadrature numbers, while the propositions are stated for the exact-integral rule of A3; the only entry that differs between the two rules is π_m at $\theta = 0.20$ (0.9894 quadrature vs 0.9890 exact), a $\leq 4 \cdot 10^{-4}$ discrepancy with no effect on any proved claim (Remark R1).

Remark B.5 (R5: FMP license, scope). What FMP §6.4 supplies is the *formula*: in symmetric two-action many-player games the noise-independent global-game selection is the action preferred under a uniform belief over the opponents’ aggregate — and (per the verified note `notes/frankelmorrispauzner2003equilibrium.md`) own-action quasiconcavity is vacuous here (two actions) while the LP-maximizer characterization rests on strategic complementarities (FMP Thm 3 shows noise-*dependence* can arise outside the licensed class). Our stage games violate global complementarities at exactly the states that drive the result (diagonal states are volunteer’s dilemmas — $U_c - U_w$ decreasing in co-movers). So the citation discipline is: FMP §6.4 licenses the Laplacian criterion as the noise-independent selection *where complements hold*; at substitutes states the same criterion is adopted as the definition of stage-wise risk dominance (the CvD vanishing-noise belief routed through the p -consistent kernel, v9.S10.2), and its robustness there is evidenced — not proven — by the agreement of the independent potential-maximizer refinement (`v9_verify_s10.py`, Tables A–C) on what the two refinements share: the contested-regime split (masking advantage $\approx +0.99$ at $\theta \leq 0.30$) and the exact revealed collapse ($\pi_r = 0$ there). The easy regime is not a tie under the potential-maximizer: Table A shows $\text{adv}_{\text{POT}} = +0.033$ at $\theta = 0.40$ and -0.064 (revealed-favoring) at $\theta = 0.75$. Anyone citing P2 should cite FMP for the criterion’s pedigree, not as a theorem covering every stage game here.

Remark B.6 (R6: other deliberate restrictions). (a) $T \geq K$ (A1) is needed for $(0, T)$ to be in the cone; the v9 calibration has $T = 10 \geq K \leq 9$. (b) Ties select wait (script convention). Under bare A8 this convention is load-bearing: (C3) is a weak inequality, and Lemma B.5’s

$A[0, 1] = 0$ invokes ties \rightarrow wait in the equality case of (C3). At the Step-5 instance the inequality is strict (margin -1.4601), so the convention does no work there; every other proved selection is strict. (c) The per-period cost ρ enters both stage values and cancels in every comparison; it matters only through continuation values, where Lemma B.5 tracks it exactly. (d) The finite- N qualifier of v7 §2 carries verbatim: nothing here is a large- N statement, and $\widehat{P}(K)$ does not survive $N \rightarrow \infty$ at fixed K/N in any uniform way — the masking advantage remains a finite- N , contested-regime object made λ -free, exactly as v9 §2 prices it.

C Proof of Proposition 3

Setup

This appendix formalizes v9 S11 (2026-06-09-ifwishlist-toy-model-v9.md, §3). Ground truth for the model: `runs/numcheck/v9_dynamic_two_channel.py` (functions `solve_unified`, `type_perceived`, `G_and_s`, `G_and_s_decoupled`); independent re-implementation `v9_verify_s11.py`. Notation follows the anchor paper §3–§6 (2026-06-10-ifwishlist-anchor-paper-v3.md): disclosure dial k , threshold share θ with $K = \lceil (1 - \theta)N \rceil$, excludability ξ , bad-coalition prior π_b .

The activation game $\Gamma_T(k_{\text{type}}, k_{\text{piv}})$. The model below is a direct formalization of `solve_unified + type_perceived` in `v9_dynamic_two_channel.py` — exact up to the idealized stage-equilibrium set S (the script approximates S on a grid with bisection; see R1.4) — written with two separate dials; the **bundled** game of S11 is the diagonal $k_{\text{type}} = k_{\text{piv}} = k$, the **decoupled** game is the full square $(k_{\text{type}}, k_{\text{piv}}) \in [0, 1]^2$ (the script’s `G_and_s_decoupled`).

M1 (population, coalition, clearing). N agents are attached to one coalition of unknown type $\psi \in \{G, B\}$ with prior $\Pr(\psi = B) = \pi_b$. Time runs in discrete periods indexed by the remaining horizon $h = T, T - 1, \dots, 1$; the state is (m, h) where m = number of agents committed so far. The coalition **clears** the first time $m \geq K$.

M2 (payoffs and their perceived, type-blended versions). True terminal payoffs: a committer pays $\kappa > 0$; if the coalition clears she receives v_ψ ($v_G = V > 0$, $v_B = -L < 0$); if she commits and the coalition never clears she receives $-L_p < 0$ (exposure loss, not type-blended). A never-committer receives ξv_ψ if the coalition clears ($\xi \in (0, 1]$ = excludability leak) and 0 otherwise. Each period an uncommitted agent is active she pays waiting friction $\rho \geq 0$. The type dial k_{type} enters as a reduced-form **payoff blend** (the script’s `type_perceived`): members of a type- ψ coalition act on the *perceived* payoffs

$$\begin{aligned} P_c(k_{\text{type}}; \psi) &= (1 - k_{\text{type}}) \bar{P}_c + k_{\text{type}} (v_\psi - \kappa), \\ P_n(k_{\text{type}}; \psi) &= (1 - k_{\text{type}}) \bar{P}_n + k_{\text{type}} \xi v_\psi, \end{aligned}$$

where the pooled values are $\bar{P}_c = \mathbb{E} - \kappa$, $\bar{P}_n = \xi \mathbb{E}$, with $\mathbb{E} := (1 - \pi_b)V - \pi_b L$ (prior-mean clearing value). P_c is the perceived payoff to having committed when the coalition clears; P_n the perceived payoff to not having committed when it clears.

M3 (stage decision problem). At state (m, h) with $m < K$ there are $n = N - m$ uncommitted agents. Given a candidate symmetric commit propensity $q \in [0, 1]$, each *other* uncommitted agent commits this period with probability $a = p_{\text{opp}}q$ ($p_{\text{opp}} \in (0, 1)$ = availability). Let $Y \sim \text{Binom}(n - 1, a)$ be the others’ commitments. The deciding agent compares (continuation objects Π, U defined in M6; conventions $\Pi(m', h) = 1$ for $m' \geq K$, $\Pi(m', 0) = 0$ and $U(m', 0) = 0$ for $m' < K$):

$$v^w(q) = -\rho + \mathbb{E} \left[\mathbf{1}\{m + Y \geq K\} P_n + \mathbf{1}\{m + Y < K\} U(m + Y, h - 1) \right],$$

$$v^c(q) = -\rho + \mathbb{E}\left[\mathbf{1}\{m+1+Y \geq K\} P_c + \mathbf{1}\{m+1+Y < K\} (P_c \Pi(m+1+Y, h-1) - L_p (1 - \Pi(m+1+Y, h-1)))\right],$$

and the commit–wait differential is $D(q) := v^c(q) - v^w(q)$ (a polynomial in q , hence continuous). The **stage-equilibrium set** is

$$S = \{q \in (0, 1) : D(q) = 0\} \cup \{0 \text{ if } D(0) \leq 0\} \cup \{1 \text{ if } D(1) \geq 0\};$$

$S \neq \emptyset$ by continuity of D .

M4 (momentum selection — a model primitive, not a derived object). The played propensity is the *largest* stage equilibrium, $a_{\text{mom}}(m, h) = \max S$ (the v6/v7 “optimistic/momentum” selection; the script’s `max(eqs)`). Its λ -free justification is the separate S10 strand; here it is taken as given.

M5 (pivotality leakage / free-ride damping). Define the pivotality weight

$$\text{piv}(m) = \text{clip}\left(1 - \frac{K - m - 1}{\max(1, p_{\text{opp}}(N - m))}, 0, 1\right)$$

and the free-ride gate $\chi = \mathbf{1}\{\xi V > V - \kappa\}$ (free-riding profitable at true good payoffs; the script computes the gate once from true values and applies it to both types). A fraction $f = k_{\text{piv}}(1 - \text{piv}(m))\chi$ of would-be movers free-ride, so the effective propensity is $a_{\text{eff}}(m, h) = a_{\text{mom}}(m, h)(1 - f)$.

M6 (transition, values, outputs). Given a_{eff} , the number of new commitments is $J \sim \text{Binom}(n, p_{\text{opp}}a_{\text{eff}})$, and

$$\begin{aligned} \Pi(m, h) &= \mathbb{E}[\mathbf{1}\{m + J \geq K\} + \mathbf{1}\{m + J < K\} \Pi(m + J, h - 1)], \\ U(m, h) &= a_{\text{eff}} v^c(a_{\text{eff}}) + (1 - a_{\text{eff}}) v^w(a_{\text{eff}}). \end{aligned}$$

The game is solved once with $\psi = G$ perceived payoffs and once with $\psi = B$; the outputs are

$$\begin{aligned} G(k_{\text{type}}, k_{\text{piv}}) &= \Pi_G(0, T), & s(k_{\text{type}}, k_{\text{piv}}) &= 1 - \Pi_B(0, T), \\ W(k_{\text{type}}, k_{\text{piv}}) &= (1 - \pi_b) G V - \pi_b (1 - s) L. \end{aligned}$$

Bundled-dial objects are written $G(k) := G(k, k)$ etc.

Assumptions

Assumption 3 (A0 — regularity). $N \geq 2$; $K \in \{2, \dots, N\}$; $p_{\text{opp}} \in (0, 1)$; $0 < \kappa < V$; $L > 0$; $L_p > 0$; $\rho \geq 0$; $\pi_b \in (0, 1)$; $\xi \in (0, 1]$.

Assumption 4 (A1 — pessimistic pooling). $0 \leq \mathbb{E} < \kappa$, where $\mathbb{E} = (1 - \pi_b)V - \pi_b L$. (Pooled with the bad prior, clearing is not worth the commitment cost; the pooled free-ride payoff $\xi\mathbb{E}$ is still nonnegative.)

Assumption 5 (A2 — free-ride regime). $\varepsilon := \xi V - (V - \kappa) > 0$, i.e. $\xi > 1 - \kappa/V$, so the gate $\chi = 1$. (The anchor’s “channel has bite only when ξ exceeds $\xi^* = 1 - \kappa/V$ ”.)

Assumption 6 (A3 — thin root field). $K - 1 \geq p_{\text{opp}}N$: at the empty state the expected number of available movers does not cover the gap, so $\text{piv}(0) = 0$ — no early mover is perceived pivotal, and disclosure at the root reveals pure inframarginality.

Assumption 7 (A4 — pivotal ignition at full type-news). With $X \sim \text{Binom}(N - 1, p_{\text{opp}})$ and $\beta_1 = \Pr(X = K - 1)$, $\beta_+ = \Pr(X \geq K)$, $\beta_- = \Pr(X \leq K - 2)$:

$$(V - \kappa) \beta_1 > \varepsilon \beta_+ + L_p \beta_-.$$

(At true good payoffs, the pivotal-event gain outweighs the free-ride temptation and the exposure risk. Since $\beta_1 > 0$ under A0, A4 is satisfiable whenever ε and L_p are small enough.)

Assumption 8 (A5 — small waiting friction). $\rho(T - 1) < \min\{L_p, \kappa - \mathbb{E}\}$. (Vacuous at $T = 1$.)

Non-vacuity. The assumption set is non-empty and contains an anchor-shaped instance: $(N, K, p_{\text{opp}}, V, \kappa, \xi, \pi_b, L, L_p) = (10, 7, 0.6, 3, 1, 0.667, 0.5, 3, 0.5)$, any $\rho \in [0, 1)$ at $T = 1$. Check: $\mathbb{E} = 0 \in [0, 1)$ (A1); $\varepsilon = 0.001 > 0$ (A2); $K - 1 = 6 \geq 6 = p_{\text{opp}}N$ (A3); $\beta_1 = 0.2508$, $\beta_+ = 0.2318$, $\beta_- = 0.5174$, so A4 reads $0.5016 > 0.001 \cdot 0.2318 + 0.5 \cdot 0.5174 = 0.2589 \checkmark$.

Stated restrictions relative to the v9 numerics (full ledger in the Remarks below): the interior-existence and decoupled-corner parts are proved at horizon $T = 1$ (one activation round), where the continuation objects $\Pi(\cdot, h - 1), U(\cdot, h - 1)$ vanish identically; the endpoint results $G(0) = G(1) = 0$ are proved for **every** $T \geq 1$. The stage-equilibrium set is the idealized S above (the script approximates it on a grid with bisection).

The proposition

Restatement (Proposition 3, restated: bundled-dial interior optimum; decoupled-dial corner).

- (i) **Bundled dial: both corners dead, a live interior, every optimum interior.**
Endpoints (any horizon). Under A0, A1, A2, A3, A5, for every $T \geq 1$, the bundled game satisfies

$$G(0) = 0, \quad s(0) = 1, \quad G(1) = 0, \quad s(1) = 1, \quad \text{hence} \quad W(0) = W(1) = 0.$$

($k = 0$: pooled news cannot ignite momentum — the unique stage equilibrium is $q = 0$ at every state; $k = 1$: full pivotality leakage makes the empty state absorbing — every root-state mover free-rides.) *Interior (one round). If in addition A4 holds and $T = 1$, there is a threshold $\bar{k} \in (0, 1)$ — the unique root of the affine root-state differential D_1 of Lemma C.4 — such that for every $k \in [\bar{k}, 1)$:*

$$G(k) = \Pr(\text{Binom}(N, p_{\text{opp}}(1 - k)) \geq K) > 0, \quad s(k) = 1, \\ W(k) = (1 - \pi_b) V G(k) > 0.$$

Moreover $W : [0, 1] \rightarrow \mathbb{R}$ is upper semicontinuous, attains its maximum, and **every maximizer lies in the open interval** $(0, 1)$.

(ii) **Decoupled dials: the optimum is the corner.** Under A0–A5 and $T = 1$, the decoupled welfare satisfies, for all $(k_{\text{type}}, k_{\text{piv}}) \in [0, 1]^2$,

$$W(k_{\text{type}}, k_{\text{piv}}) \leq W(1, 0) = (1 - \pi_b) V \Pr(\text{Binom}(N, p_{\text{opp}}) \geq K) > 0,$$

and the argmax set is **exactly** $[\bar{k}, 1] \times \{0\}$ (with \bar{k} from part (i)): every optimum has zero pivotality leakage and full ignition, and $(k_{\text{type}}, k_{\text{piv}}) = (1, 0)$ attains the optimum. In particular every point with $k_{\text{piv}} > 0$ is strictly suboptimal.

(iii) **The interior is a constrained optimum created by bundling.** Under A0–A5 and $T = 1$: $\sup_{k \in [0, 1]} W(k, k) \leq W(1, 0)$, and the bundled diagonal never attains $W(1, 0)$ (every diagonal point has $a_{\text{eff}} < 1$). The bundled optimum is interior (part (i)) while the unconstrained two-dial optimum is the corner (part (ii)); the interior k^* is the constrained optimum of one dial forced to move both channels together — not a type-versus-pivotality balance.

Proof

Throughout, fix a state (m, h) , write $n = N - m$, $g = K - m$ (the gap), and let $Y \sim \text{Binom}(n-1, p_{\text{opp}}q)$ denote the others' current-period commitments in the stage problem. Note $\Pr(Y = 0) = (1 - p_{\text{opp}}q)^{n-1} \geq (1 - p_{\text{opp}})^{n-1} > 0$ for all $q \in [0, 1]$, by $p_{\text{opp}} < 1$ (A0). (★)

Lemma C.1 (strict monotonicity of the binomial tail). *For integers $n \geq 1$, $1 \leq r \leq n$, the map $F(b) = \Pr(\text{Binom}(n, b) \geq r)$ is continuous on $[0, 1]$, nondecreasing, strictly increasing on $(0, 1)$, and $F(b) > 0$ for $b > 0$; indeed $F'(b) = n \binom{n-1}{r-1} b^{r-1} (1-b)^{n-r}$.*

Proof. $F(b) = \sum_{j \geq r} \binom{n}{j} b^j (1-b)^{n-j}$ is a polynomial. Term-by-term,

$$\frac{d}{db} \binom{n}{j} b^j (1-b)^{n-j} = \underbrace{n \binom{n-1}{j-1} b^{j-1} (1-b)^{n-j}}_{=: t_j} - \underbrace{n \binom{n-1}{j} b^j (1-b)^{n-j-1}}_{=: u_j},$$

using $j \binom{n}{j} = n \binom{n-1}{j-1}$ and $(n-j) \binom{n}{j} = n \binom{n-1}{j}$. Since $u_j = t_{j+1}$ and $u_n = 0$, the sum over $j = r, \dots, n$ telescopes to $t_r = n \binom{n-1}{r-1} b^{r-1} (1-b)^{n-r}$, which is > 0 on $(0, 1)$. Positivity for $b > 0$: $F(b) \geq \Pr(\text{Binom}(n, b) = n) = b^n > 0$ when $r \leq n$. \square

Lemma C.2 (no ignition under pooled news; any horizon, any k_{piv} , both types). *Assume A0, A5, the bound $\mathbb{E} < \kappa$ from A1 (the sign restriction $\mathbb{E} \geq 0$ in A1 is not used here), and $k_{\text{type}} = 0$ (with $k_{\text{piv}} \in [0, 1]$ arbitrary, type $\psi \in \{G, B\}$ arbitrary). Then for every state (m, h) with $m < K$, $h \geq 1$: the unique stage equilibrium is $q = 0$, hence $a_{\text{mom}} = a_{\text{eff}} = 0$; and $\Pi(m, h) = 0$, $U(m, h) = -\rho h$. Consequently $G(0, k_{\text{piv}}) = 0$ and $s(0, k_{\text{piv}}) = 1$ for every $T \geq 1$.*

Proof. At $k_{\text{type}} = 0$ the perceived payoffs are the pooled values for **both** types: $P_c = \bar{P}_c = \mathbb{E} - \kappa$, $P_n = \bar{P}_n = \xi \mathbb{E}$. By $\mathbb{E} < \kappa$: $\bar{P}_c < 0$, and $\bar{P}_c - \bar{P}_n = (1 - \xi) \mathbb{E} - \kappa < 0$ (if $\mathbb{E} \geq 0$ this is $\leq \mathbb{E} - \kappa < 0$ using $\xi \leq 1$; if $\mathbb{E} < 0$ then $(1 - \xi) \mathbb{E} \leq 0$ and the expression is $\leq -\kappa < 0$). (†)

Induct on h . *Base $h = 0$* : $\Pi(m, 0) = 0$, $U(m, 0) = 0 = -\rho \cdot 0$ for $m < K$ by convention. *Step*: assume $\Pi(m', h-1) = 0$ and $U(m', h-1) = -\rho(h-1)$ for all $m' < K$. Substituting the

hypothesis into M3 (failure branch of v^c becomes $-L_p$; wait branch on $\{m + Y < K\}$ becomes $-\rho(h - 1)$; the two leading $-\rho$ cancel):

$$D(q) = \left[\bar{P}_c \Pr(Y \geq g - 1) - L_p \Pr(Y \leq g - 2) \right] - \left[\bar{P}_n \Pr(Y \geq g) - \rho(h - 1) \Pr(Y \leq g - 1) \right].$$

Splitting $\Pr(Y \geq g - 1) = \Pr(Y = g - 1) + \Pr(Y \geq g)$ and $\Pr(Y \leq g - 1) = \Pr(Y = g - 1) + \Pr(Y \leq g - 2)$:

$$D(q) = (\bar{P}_c + \rho(h - 1)) \Pr(Y = g - 1) + (\bar{P}_c - \bar{P}_n) \Pr(Y \geq g) - (L_p - \rho(h - 1)) \Pr(Y \leq g - 2). \quad (\text{L2.1})$$

By A5, $\rho(h - 1) \leq \rho(T - 1) < \min\{L_p, \kappa - \mathbb{E}\}$, so all three coefficients are negative: $\bar{P}_c + \rho(h - 1) = \mathbb{E} - \kappa + \rho(h - 1) < 0$; $\bar{P}_c - \bar{P}_n < 0$ by (\dagger); $-(L_p - \rho(h - 1)) < 0$. Hence $D(q) \leq 0$ for all q . Strictness: if $m \leq K - 2$ then $g - 2 \geq 0$ so $\Pr(Y \leq g - 2) \geq \Pr(Y = 0) > 0$ by (\star) and the third term is strictly negative; if $m = K - 1$ then $g - 1 = 0$ so $\Pr(Y = g - 1) \geq \Pr(Y = 0) > 0$ by (\star) and the first term is strictly negative. So $D(q) < 0$ for all $q \in [0, 1]$: $q = 0 \in S$ (since $D(0) < 0$), no $q \in (0, 1)$ is in S (needs $D = 0$), and $1 \notin S$ (needs $D(1) \geq 0$). Thus $a_{\text{mom}} = \max S = 0$, so $a_{\text{eff}} = 0$ regardless of k_{piv} (the damping multiplies zero). The transition is then degenerate at $J = 0$: $\Pi(m, h) = \Pi(m, h - 1) = 0$, and $U(m, h) = v^w(0) = -\rho + U(m, h - 1) = -\rho h$ (at $h = 1$ the wait-no-clear branch is 0, consistent with the convention). The induction closes. Finally $G(0, k_{\text{piv}}) = \Pi_G(0, T) = 0$ and $\Pi_B(0, T) = 0$ gives $s = 1$. \square

Lemma C.3 (full pivotality leakage freezes the root; any horizon, any k_{type} , both types). *Assume A0, A2, A3 and $k_{\text{piv}} = 1$ (with $k_{\text{type}} \in [0, 1]$, ψ arbitrary). Then $a_{\text{eff}}(0, h) = 0$ for every h , the state $m = 0$ is absorbing, and $\Pi(0, T) = 0$ for every $T \geq 1$. Consequently $G(k_{\text{type}}, 1) = 0$ and $s(k_{\text{type}}, 1) = 1$.*

Proof. At $m = 0$: the gap is K and $n = N$, so $\text{piv}(0) = \text{clip}\left(1 - \frac{K-1}{\max(1, p_{\text{opp}}N)}\right)$. If $p_{\text{opp}}N \geq 1$ the ratio is ≥ 1 by A3 ($K - 1 \geq p_{\text{opp}}N$); if $p_{\text{opp}}N < 1$ the denominator is 1 and the ratio is $K - 1 \geq 1$ by $K \geq 2$ (A0). Either way $\text{piv}(0) = 0$. By A2 the gate $\chi = 1$. Hence $f = 1 \cdot (1 - 0) \cdot 1 = 1$ and $a_{\text{eff}}(0, h) = a_{\text{mom}}(0, h) \cdot 0 = 0$ — whatever a_{mom} , k_{type} , and ψ are. The transition from $(0, h)$ is degenerate at $J = 0$, so by induction on h , $\Pi(0, T) = \Pi(0, 0) = 0$ (using $K \geq 2 > 0$). Both types are frozen at the root, so $G(\cdot, 1) = 0$ and $1 - s(\cdot, 1) = \Pi_B(0, T) = 0$. \square

Lemma C.4 (interior ignition at $T = 1$). *Assume A0–A5 and $T = 1$, type G, bundled dial $k_{\text{type}} = k_{\text{piv}} = k$. Define*

$$D_1(k) := P_c(k; G) \beta_1 - (P_n(k; G) - P_c(k; G)) \beta_+ - L_p \beta_-,$$

with $\beta_1, \beta_+, \beta_-$ as in A4. Then: (a) D_1 is the commit–wait differential at the root state $(0, 1)$ evaluated at $q = 1$; (b) D_1 is affine and strictly increasing in k , with $D_1(0) < 0 < D_1(1)$, so there is a unique $\bar{k} \in (0, 1)$ with $D_1(k) \geq 0$ exactly on $[\bar{k}, 1]$; (c) for every $k \in [\bar{k}, 1]$, $a_{\text{mom}}(0, 1) = 1$ and $a_{\text{eff}}(0, 1) = 1 - k$; (d) for every $k \in [\bar{k}, 1)$, $G(k) = \Pr(\text{Binom}(N, p_{\text{opp}}(1 - k)) \geq K) > 0$.

Proof. (a) At $T = 1$ the continuations vanish ($\Pi(\cdot, 0) = 0$, $U(\cdot, 0) = 0$), so at $(m, h) = (0, 1)$, $g = K$, the formula-(L2.1)-style computation gives, for general q ,

$$D(q; k) = P_c(k) \Pr(Y = K - 1) - (P_n(k) - P_c(k)) \Pr(Y \geq K) - L_p \Pr(Y \leq K - 2), \quad (\text{L4.1})$$

where now $P_c(k) = P_c(k; G)$, $P_n(k) = P_n(k; G)$ and $Y \sim \text{Binom}(N - 1, p_{\text{opp}}q)$. At $q = 1$ the three probabilities are $\beta_1, \beta_+, \beta_-$, giving $D(1; k) = D_1(k)$.

(b) $P_c(\cdot)$ and $P_n(\cdot)$ are affine in k (M2), so D_1 is affine. At $k = 1$: $P_c = V - \kappa$, $P_n - P_c = \xi V - (V - \kappa) = \varepsilon$, so $D_1(1) = (V - \kappa)\beta_1 - \varepsilon\beta_+ - L_p\beta_- > 0$ by A4. At $k = 0$ the payoffs are the pooled ones, and each term of (L4.1) is ≤ 0 by (\dagger) of Lemma C.2 ($\bar{P}_c < 0$, $\bar{P}_c - \bar{P}_n < 0$), with the third strictly negative because $\beta_- \geq (1 - p_{\text{opp}})^{N-1} > 0$ (\star), using $K \geq 2$); so $D_1(0) < 0$. An affine function with $D_1(0) < 0 < D_1(1)$ has strictly positive slope, hence a unique root $\bar{k} \in (0, 1)$ and $\{k : D_1(k) \geq 0\} = [\bar{k}, 1]$.

(c) For $k \in [\bar{k}, 1]$, $D(1; k) = D_1(k) \geq 0$, so $q = 1 \in S$ and $a_{\text{mom}}(0, 1) = \max S = 1$. By A3 and A2 (as in Lemma C.3) $\text{piv}(0) = 0$, $\chi = 1$, so $f = k$ and $a_{\text{eff}} = 1 - k$.

(d) For $k \in [\bar{k}, 1)$, $a_{\text{eff}} = 1 - k \in (0, 1]$, and at $T = 1$ the clearing probability is exactly one binomial tail over all $n = N$ uncommitted agents (M6 with $\Pi(\cdot, 0) = 0$): $G(k) = \Pi_G(0, 1) = \Pr(\text{Binom}(N, p_{\text{opp}}(1 - k)) \geq K)$, which is > 0 by Lemma C.1 since $p_{\text{opp}}(1 - k) > 0$ and $K \leq N$. \square

Lemma C.5 (bad coalitions never ignite at $T = 1$, at any dial setting). *Assume A0, A1 and $T = 1$, type B, any $(k_{\text{type}}, k_{\text{piv}}) \in [0, 1]^2$. (A2 is not needed: the stage equilibrium below is already $q = 0$, so the conclusion holds whatever the value of the gate χ .) Then the unique stage equilibrium at $(0, 1)$ is $q = 0$, so $\Pi_B(0, 1) = 0$ and $s(k_{\text{type}}, k_{\text{piv}}) = 1$.*

Proof. Write $t = k_{\text{type}}$. The bad-type perceived payoffs are $P_c(t; B) = (1 - t)\bar{P}_c + t(-L - \kappa)$ and $P_n(t; B) = (1 - t)\bar{P}_n + t(-\xi L)$. Both blend components of $P_c(\cdot; B)$ are negative ($\bar{P}_c = \mathbb{E} - \kappa < 0$ by A1; $-L - \kappa < 0$), so $P_c(t; B) < 0$ for all t . Also

$$P_c(t; B) - P_n(t; B) = (1 - t)((1 - \xi)\mathbb{E} - \kappa) + t(-(1 - \xi)L - \kappa) < 0,$$

since the first bracket is < 0 by (\dagger) and the second is < 0 ($\xi \leq 1$, $L > 0$, $\kappa > 0$). Substituting into (L4.1) (which holds for either type at $T = 1$): all three terms are ≤ 0 and the third is strictly negative ($\Pr(Y \leq K - 2) \geq \Pr(Y = 0) > 0$ by \star and $K \geq 2$). Hence $D(q) < 0$ for all q , $a_{\text{mom}} = 0$, $a_{\text{eff}} = 0$, $\Pi_B(0, 1) = 0$. \square

Lemma C.6 (upper semicontinuity and attainment, $T = 1$, bundled). *Assume A0–A3, A5 and $T = 1$. Then $k \mapsto a_{\text{mom}}(0, 1)(k)$ is upper semicontinuous on $[0, 1]$; hence $W(k) = (1 - \pi_b)V \cdot \Pr(\text{Binom}(N, p_{\text{opp}} a_{\text{mom}}(k)(1 - k)) \geq K)$ is upper semicontinuous and attains its maximum on $[0, 1]$.*

Proof. By Lemma C.5, $s \equiv 1$, so $W(k) = (1 - \pi_b)V G(k)$ with $G(k) = F(p_{\text{opp}} a_{\text{eff}}(k))$, F the (continuous, nondecreasing) binomial tail of Lemma C.1 and $a_{\text{eff}}(k) = a_{\text{mom}}(k)(1 - k)$ (Lemma C.4(c)'s damping form, valid at the root for all k by A2–A3). $D(q; k)$ in (L4.1) is jointly continuous (polynomial in q , affine in k). Let $k_j \rightarrow k$ and pass to a subsequence with $q_j := a_{\text{mom}}(k_j) \rightarrow \alpha = \limsup_j a_{\text{mom}}(k_j)$. Each $q_j \in S(k_j)$ falls in one of three cases, one of which holds along a further subsequence: ($q_j = 0$, $D(0; k_j) \leq 0$): then $\alpha = 0 \leq a_{\text{mom}}(k)$ since $S(k) \subseteq [0, 1]$ is nonempty. ($q_j = 1$, $D(1; k_j) \geq 0$): then $D(1; k) \geq 0$ by continuity, so $1 \in S(k)$ and $a_{\text{mom}}(k) = 1 \geq \alpha$. ($q_j \in (0, 1)$, $D(q_j; k_j) = 0$): then $D(\alpha; k) = 0$ by joint continuity; if $\alpha \in (0, 1)$ then $\alpha \in S(k)$; if $\alpha \in \{0, 1\}$ then $D(\alpha; k) = 0$ also certifies $\alpha \in S(k)$ (the endpoint inequalities hold with equality). Either way $a_{\text{mom}}(k) \geq \alpha$. So a_{mom} is u.s.c.; then a_{eff} is u.s.c. (product of a nonnegative u.s.c. function and a continuous nonnegative function), and $W = (1 - \pi_b)V \cdot F(p_{\text{opp}} \cdot)$ composed with a_{eff} is u.s.c. (F continuous and nondecreasing). An u.s.c. function on a compact set attains its maximum. \square

Proof of Proposition 3(i). Endpoints. $G(0) = 0$ and $s(0) = 1$: Lemma C.2 with $k_{\text{piv}} = 0$ (any T). $G(1) = 0$ and $s(1) = 1$: Lemma C.3 with $k_{\text{type}} = 1$ (any T). Then by M6, $W(0) = W(1) = (1 - \pi_b) \cdot 0 \cdot V - \pi_b \cdot 0 \cdot L = 0$.

Interior, $T = 1$. Lemma C.4 gives $\bar{k} \in (0, 1)$ and, for $k \in [\bar{k}, 1)$, $G(k) = \Pr(\text{Binom}(N, p_{\text{opp}}(1 - k)) \geq K) > 0$; Lemma C.5 gives $s(k) = 1$, so $W(k) = (1 - \pi_b)V G(k) > 0$ (using $\pi_b < 1$, $V > 0$).

Every maximizer interior. By Lemma C.6 the maximum is attained; its value is $\geq W(\bar{k}) > 0 = W(0) = W(1)$, so neither endpoint is a maximizer; every maximizer lies in $(0, 1)$. \square

Proof of Proposition 3(ii). At $T = 1$, by Lemma C.5, $s \equiv 1$ on $[0, 1]^2$, so $W(k_{\text{type}}, k_{\text{piv}}) = (1 - \pi_b)V F(p_{\text{opp}} a_{\text{eff}}(k_{\text{type}}, k_{\text{piv}}))$ with $F(b) = \Pr(\text{Binom}(N, b) \geq K)$ and, at the root (A2, A3, as in Lemma C.3),

$$a_{\text{eff}}(k_{\text{type}}, k_{\text{piv}}) = a_{\text{mom}}(k_{\text{type}})(1 - k_{\text{piv}}),$$

where $a_{\text{mom}}(k_{\text{type}})$ depends only on k_{type} (the stage problem M3 involves the perceived payoffs only; the damping M5 is applied after).

Upper bound and attainment. $a_{\text{eff}} \leq 1$ always, and F is nondecreasing (Lemma C.1), so $W \leq (1 - \pi_b)V F(p_{\text{opp}})$. By Lemma C.4(b)–(c) (whose ignition part uses only the type dial), $a_{\text{mom}}(k_{\text{type}}) = 1$ exactly for $k_{\text{type}} \in [\bar{k}, 1]$; in particular $a_{\text{mom}}(1) = 1$ (A4). Hence $a_{\text{eff}}(1, 0) = 1$ and $W(1, 0) = (1 - \pi_b)V F(p_{\text{opp}})$, which is > 0 by Lemma C.1 ($p_{\text{opp}} > 0$, $K \leq N$). So the upper bound is attained at $(1, 0)$.

Exact argmax set. A point $(k_{\text{type}}, k_{\text{piv}})$ attains the bound iff $F(p_{\text{opp}} a_{\text{eff}}) = F(p_{\text{opp}})$. Since F is strictly increasing on $(0, 1)$ (Lemma C.1) and $0 < p_{\text{opp}} < 1$, this holds iff $a_{\text{eff}} = 1$, i.e. iff $a_{\text{mom}}(k_{\text{type}}) = 1$ **and** $k_{\text{piv}} = 0$, i.e. iff $(k_{\text{type}}, k_{\text{piv}}) \in [\bar{k}, 1] \times \{0\}$. (If $a_{\text{eff}} = 0$ the value is $F(0) = 0 < W(1, 0)$; if $a_{\text{eff}} \in (0, 1)$ the value is strictly below the bound.) In particular any $k_{\text{piv}} > 0$ forces $a_{\text{eff}} \leq 1 - k_{\text{piv}} < 1$, hence strict suboptimality. \square

Proof of Proposition 3(iii). The diagonal is a subset of the square, so $\sup_k W(k, k) \leq \max W = W(1, 0)$. No diagonal point attains it: attainment requires $a_{\text{eff}} = 1$, i.e. $a_{\text{mom}}(k) = 1$ and $k = 0$ simultaneously; but $a_{\text{mom}}(0) = 0$ by Lemma C.2 (so $k = 0$ gives $a_{\text{eff}} = 0$), and $k > 0$ gives $a_{\text{eff}} \leq 1 - k < 1$. Combined with part (i) (every bundled maximizer interior) and part (ii) (every decoupled maximizer at $k_{\text{piv}} = 0$, $k_{\text{type}} \in [\bar{k}, 1]$), the bundled interior optimum is strictly below the decoupled corner optimum: it is the constrained optimum of the yoked dial, not a channel balance. \square

Remarks

Remark C.1 (R1 — what was restricted relative to the v9 numerics).

1. *Horizon.* The interior-existence step (Lemma C.4) and the decoupled-corner result (part (ii)) are proved at $T = 1$, where the continuation objects vanish. The endpoint results $G(0) = G(1) = 0$ (Lemmas C.2–C.3) hold for **every** $T \geq 1$, including the script's $T = 10$. The script's run is the $T = 10$ game; see the conjecture block below (C1 open; C2 falsified).
2. *Assumption regime vs. the anchor calibration.* The anchor calibration ($\pi_b = 0.2$, $V = L = 3$, $\kappa = 1$) has $\mathbb{E} = 1.8 > \kappa$, violating A1. It also violates A4: with the anchor's $L_p = 3$, the right side of A4 is $\varepsilon\beta_+ + L_p\beta_- = 0.001 \cdot 0.2318 + 3 \cdot 0.5174 = 1.5524$, far above the left side $(V - \kappa)\beta_1 = 2 \cdot 0.2508 = 0.5016$. And A5 fails automatically once

A1 does ($\kappa - \mathbb{E} = -0.8 < 0$, so no $\rho \geq 0$ satisfies it at $T = 10$). Of A0–A5 the anchor thus satisfies only A0, A2, A3: no interior result (Lemma C.4, part (i)’s interior clause, part (ii)) is instantiated at the anchor — see R3 for what this does to the mapping. Its $G(0) \approx 0$ is produced by a different sufficient mechanism — large exposure risk $L_p = 3$ plus a thin field $p_{\text{opp}} = 0.6$ (even at $q = 1$ the root failure probability is large, so the $-L_p$ term dominates and the high stage equilibrium fails) — which we have **not** formalized. A1 (pessimistic pooling, $\mathbb{E} < \kappa$) is the clean sufficient route to the same endpoint, and is the economically natural one (“masking pools you with the bad prior, and the pool is not worth the commitment cost”). The conditions here are sufficient, not necessary, and not the anchor calibration’s own mechanism at $k = 0$. The $k = 1$ endpoint mechanism (Lemma C.3), by contrast, **is** exactly the anchor’s: in the $T = 10$ run, $G(1) = 0$ arises because $f = 1$ at $m = 0$ ($K - 1 = 6 \geq 6 = p_{\text{opp}}N$, A3 holds with equality at the anchor) — the same root-freezing as our proof.

3. *Equilibrium selection.* The optimistic selection $a_{\text{mom}} = \max S$ is a model primitive (M4, the v6/v7 convention; its λ -free license is the S10/FMP §6.4 strand, not re-proved here). Lemma C.4’s interior ignition shows $q = 1$ is a stage equilibrium and relies on the max selection to conclude it is played. Under a pessimistic selection ($\min S$) the interior would die at $T = 1$ ($D(0) < 0$ always, so $0 \in S$ always); the inverted-U is a property of the momentum-selected dynamics, exactly as v9 frames it.
4. *Idealized equilibrium set.* The script approximates S on a 201-point grid with bisection and 10^{-12} tolerances; we use the exact S . No result depends on the approximation.
5. *Reduced-form channels.* The type channel is a payoff blend, not Bayesian updating from a signal, and the free-ride fraction is the script’s $k_{\text{piv}}(1 - \text{piv})\chi$ form — both inherited verbatim from the ground-truth script (its own header flags them as “honestly a unified-but-reduced model”). P3 formalizes that reduced model, not the deeper S12 information-theoretic channel.

Remark C.2 (R2 — a fully verified instance). $(N, K, p_{\text{opp}}, V, \kappa, \xi, \pi_b, L, L_p, T) = (10, 7, 0.6, 3, 1, 0.667, 0.5, 3, 0.5, 1)$ satisfies A0–A5 (arithmetic in the Setup above). Here \bar{k} solves $0.98402k = 0.74130$, i.e. $\bar{k} \approx 0.7533$. Cross-check against the ground-truth solver run at these parameters with $T = 1$ (sanity check only; no proof step uses it): $G(k) = 0$ for $k \in \{0, 0.2, \dots, 0.75\}$, $G(0.76) = 1.03 \times 10^{-4} > 0$, $G(0.9) = 2.9 \times 10^{-7} > 0$, $G(1) = 0$, $s \equiv 1$; decoupled 11×11 grid argmax value $0.57342 = (1 - \pi_b)V F(p_{\text{opp}})$ exactly as part (ii) predicts, reported at $(k_{\text{type}}, k_{\text{piv}}) = (0.8, 0.0)$ — the smallest grid point of the predicted flat argmax set $[\bar{k}, 1] \times \{0\} = [0.7533, 1] \times \{0\}$ (the script’s tie-break keeps the first maximizer found).

Remark C.3 (R3 — map to the numerical anchors (v9 S11, $T = 10$, $\theta = 0.30$, $\xi = 0.667$, $\pi_b = 0.2$)). $\theta = 0.30 \mapsto K = 7$ — A3 holds with equality ($K - 1 = 6 = p_{\text{opp}}N$). $\xi = 0.667 \mapsto$ A2 with $\varepsilon = 0.001$ (the anchor sits just inside the free-ride regime, $\xi^* = 2/3$). “ $G(0) = 0$: masking cannot ignite momentum” \mapsto Lemma C.2’s no-ignition conclusion ($a \equiv 0$ at every state; sufficient condition A1+A5, see R1.2). “ $G(1) = 0$: full leakage fully unravels” \mapsto Lemma C.3 (root absorbing; the anchor satisfies the exact hypotheses A2+A3). “Interior $G(0.70) = 0.997$, $W = +2.39$ ” \mapsto Lemma C.4’s live band $[\bar{k}, 1)$ — illustrative only, NOT an instantiation: the anchor violates A1, A4, and A5 (R1.2), so Lemma C.4’s hypotheses do not hold there and the anchor’s interior is not an instance of the proposition. Among the anchor

facts, only the $G(1) = 0$ endpoint (Lemma C.3, hypotheses A2+A3, which the anchor meets exactly) is a genuine instance; the $G(0) = 0$ endpoint and the interior are qualitative matches produced by mechanisms we have not formalized. The genuine A0–A5 instance with a live interior is R2’s ($L_p = 0.5$). At $T = 1$ magnitudes are tiny ($G \sim 10^{-4}$) because a single round has no momentum amplification, while the anchor’s $T = 10$ chain amplifies the same ignition into $G \approx 1$. “Decoupled corner $(k_{\text{type}}, k_{\text{piv}}) = (1, 0)$ at every ξ ” \mapsto part (ii): the corner $(1, 0)$ attains the optimum for every ξ satisfying A2 (the proof nowhere fixes ξ beyond A2/A4); for $\xi \leq \xi^*$ the gate is off, $f \equiv 0$, k_{piv} is payoff-irrelevant, and $k_{\text{piv}} = 0$ is (weakly) optimal trivially. “Interior $W = +2.39$ beats both corners by the whole range” \mapsto Proposition 3(i) final clause: $\max W > 0 = W(0) = W(1)$.

Conjectures, a falsification record, and an observation (R4)

The following records the precise open statements and what would close them, together with the falsification record of the original general-horizon decoupled-corner conjecture.

Restatement (Conjecture 3 (C1), restated: general-horizon interior). *Under A0–A3, A5, and a strengthened ignition condition, for every $T \geq 1$ there exists $k \in (0, 1)$ with $G(k) > 0$ in the bundled game.*

Status: proved at $T = 1$ (Lemma C.4); numerically corroborated at $T = 10$ at the anchor (live interior with $G(0.70) = 0.997$ — corroboration only, since the anchor lies outside A0–A5, R1.2). *Obstruction:* for $T \geq 2$ the wait branch of $D(1)$ contains the endogenous option value $U(m + Y, T - 1)$ of staying uncommitted; the crude bound $U \leq P_n$ is too loose (it makes the wait branch dominate for any $\varepsilon > 0$). Closing C1 needs a sharper upper bound on the waiting option value at momentum-selected play — e.g., $U(m', h) \leq P_n \Pi(m', h) - c(1 - \Pi(m', h))$ for a positive constant c — or a direct monotone-induction argument that ignition at the root only improves at deeper states (where the gap is smaller and $\text{piv} > 0$ weakens the damping).

Remark C.4 (A falsified conjecture: C2 — general-horizon decoupled corner). As originally conjectured, C2 read: under A0–A5, for every $T \geq 1$ the decoupled argmax has $k_{\text{piv}} = 0$ and is attained at $(1, 0)$. This is false. The counterexample hunter exhibited, inside A0–A5,

$$\begin{aligned} & (N, K, p_{\text{opp}}, V, \kappa, \xi, \pi_b, L, L_p, \rho) \\ & = (8, 4, 0.29, 4.84, 2.00, 0.958, 0.32, 9.625, 0.408, 0.0126) \end{aligned}$$

— all of A0–A5 verified, A4: $0.6160 > 0.4770$ — where the ground-truth solver gives $W(k_{\text{type}}=1, k_{\text{piv}}=0.2) = 0.9903 > W(1, 0) = 0.7629$, identically at $T = 2, 5, 10$, with $s \equiv 1$ — a pure good-coalition effect: at the full-disclosure corner, leakage $k_{\text{piv}}: 0 \rightarrow 0.2$ strictly RAISES the activation gain (the hunter reports the gain rising from 0.232 to 0.296; our re-extraction from `v9_dynamic_two_channel.py` gives $W(1, 0.2) = 0.99025$, $W(1, 0) = 0.76286$, with $G: 0.2318 \rightarrow 0.3009$). Confirmed by `v9_verify_s11.W_general` to 5 decimals. Per the hunter, the failure is GENERIC: 6/8 fresh random A0–A5 draws violate C2 at $T=10$ (k_{piv} in $\{0.1-0.3\}$ strictly beats every $k_{\text{piv}}=0$ point), and $(1, 0)$ is not even the argmax of the $k_{\text{piv}} = 0$ line at $T \geq 2$ ($k_{\text{type}} = 0.9$ beats $k_{\text{type}} = 1$ in the instance above) — we re-verified the line claim: $W(0.9, 0) = 0.7974 > 0.7629 = W(1, 0)$ at $T = 10$, and $W(0.8, 0) = 0.8284 > W(1, 0)$ already at $T = 2$.

Replication note (recorded for completeness, not as a hedge — the falsification stands on the re-verified numbers above): one hunter figure did not reproduce, the “max over the entire

$k_{\text{piv}} = 0$ line = 0.9378”. On an 11-point k_{type} grid the line max is 0.8284 (so $W(1, 0.2)$ does beat every coarse-grid $k_{\text{piv}} = 0$ point), but refining the line yields $W(0.85, 0) = 1.0313$ and $W(0.8625, 0) = 1.0788$, both above $W(1, 0.2)$, and a 21×21 scan with local refinement places this instance’s global argmax at $(0.8625, 0)$ — at $k_{\text{piv}} = 0$ after all. What is decisively falsified is therefore C2’s attainment at $(1, 0)$ and the part-(ii) comparative static “every $k_{\text{piv}} > 0$ point is strictly suboptimal”; whether the weaker “some $k_{\text{piv}} = 0$ point is optimal” survives is C2’ below (Conjecture 4; the hunter’s 6/8 line-dominance draws carry the same coarse-grid caveat).

What survives: (a) the decoupled corner is a THEOREM at $T = 1$ (Proposition 3(ii), proved above, untouched); (b) at the v9 anchor calibration at $T = 10$ it is a numerical fact — re-verified against `v9_dynamic_two_channel_best_2d` (11×11): argmax $(1.00, 0.00)$ with $W^* = +2.4000, +2.3995, +2.3855$ at $\xi = 0.667, 0.70, 0.75$. (The anchor lies outside A0–A5, R1.2, so this corroborates rather than instantiates.)

Mechanism (per the hunter; consistent with our re-runs): at $T \geq 2$ a leakier future cuts the waiting option value and re-ignites mid-states — exactly the continuation channel named in C1’s obstruction as what breaks the $T = 1$ factorization $a_{\text{eff}} = a_{\text{mom}}(k_{\text{type}})(1 - k_{\text{piv}})$.

Restatement (Observation 1 (O1), restated: positive leakage can strictly help at $T \geq 2$). *There exist instances satisfying A0–A5, horizons $T \geq 2$, and dials $k_{\text{type}} \in [0, 1]$, $k_{\text{piv}} > 0$ such that $W(k_{\text{type}}, k_{\text{piv}}) > W(k_{\text{type}}, 0)$ with $s \equiv 1$ at both points: positive pivotality leakage is strictly welfare-improving, through the good-coalition continuation/waiting-option channel. Verified instance: the C2 counterexample (Remark C.4) at $k_{\text{type}} = 1$, $k_{\text{piv}} = 0.2$, where W rises $0.7629 \rightarrow 0.9902$ for every $T \in \{2, 5, 10\}$. At $T = 1$ this is impossible (Proposition 3(ii): a_{eff} is nonincreasing in k_{piv} and F is nondecreasing), so the phenomenon is strictly dynamic.*

Restatement (Conjecture 4 (C2’), restated: narrowed residual conjecture). *Under A0–A5, for every $T \geq 1$ the decoupled welfare attains its maximum at some point with $k_{\text{piv}} = 0$ — with no claim that $(1, 0)$ attains it, and no pointwise dominance claim in k_{piv} .*

Status: theorem at $T = 1$ (Proposition 3(ii)); at the C2-falsifying instance the refined scan is consistent with it (argmax $(0.8625, 0)$); at risk from the hunter’s 6/8 genericity report if those violations survive k_{type} -grid refinement — open. *Falsifier:* any A0–A5 instance and T where a fine k_{type} scan certifies $\sup_{k_{\text{type}}} W(k_{\text{type}}, 0) < W(k'_{\text{type}}, k'_{\text{piv}})$ for some $k'_{\text{piv}} > 0$. *Obstruction (unchanged from the original C2):* for $T \geq 2$, k_{piv} feeds back into a_{mom} through the continuation values Π, U , so the clean factorization used in part (ii) fails state-by-state — and O1 shows the feedback can be strictly beneficial, so no monotone (Topkis-style) comparative static in k_{piv} can hold pointwise; any proof of C2’ must compare optima, not paths.

Remark C.5 (R5 — reading). Part (i)’s endpoints are the two channels’ pure failure modes derived from primitives: A1 makes the *pooled* commit option dominated (masking destroys the type news needed to ignite), and A2+A3 make the *root* mover provably inframarginal under full disclosure (leakage destroys the perceived pivotality needed to move first). Part (ii) shows neither failure is about *balancing* the channels: given two instruments, the planner sets each at its own best corner. The interior k^* exists only because one physical k-anonymity dial yokes the two — exactly the v9 S11 claim, here as a theorem at $T = 1$. Beyond one round the two extensions now part ways: the bundled-interior extension C1 remains an open conjecture (Conjecture 3), while the decoupled-corner extension C2 is falsified (Remark C.4) — at $T \geq 2$ positive leakage can strictly raise welfare through the continuation channel (Observation 1), so the corner result is a one-round theorem plus an anchor-calibration numerical fact, not a general-horizon truth.