

Research Article

Multiplicative Amplification of Position in Social Hierarchies: Identifiable Dynamics and an Unidentifiable Seed

Kristian Sestak¹

1. Independent researcher

Trajectories of positions are widely used to argue that a hierarchy was built by amplification: the early ordering predicts the final one better than diffusion allows, and the excess is read as compounding. We show that this inference fails. Entities that merely differ in a persistent trait produce the same excess at zero gain, and that null carries a free parameter tunable to any lock-in speed: magnitude identifies nothing, only a within-system contrast does. Inside a falsifiable *normal form*, a multiplicative cascade amplifying *position*, we separate what a trajectory settles from what it cannot. The *dynamics* are identifiable: a necessity theorem, a non-monotone dispersion-maximising gain, a Hill-verified closed-form Kesten tail index tying tail heaviness to contestability, and a $\sqrt{\tau}$ early-lead law. The *seed* is not: the persistent input is recovered, but its decomposition into capability and position is not a function of the trajectory. The obstruction is the amplified form of one panel econometrics knows, state dependence against unobserved heterogeneity, exact here. This is observational equivalence, not imprecision, so no quantity of data bears on it; the split must come from another observable, and what an exogenous handle returns is a proportion, the *manufactured share* $R = 1 - \text{corr}^2$, not a verdict. Both points bite: a within-Wikipedia revocation contrast supplies the capability-free test ($\Delta\rho = 0.11$, 95% CI [0.09, 0.13]; a third of it heterogeneity by placebo), and an online-chess cohort with every appearance of manufactured order is at least $\sim 62\%$ revealed skill ($R \approx 0.38$, an upper bound).

Corresponding author: Kristian Sestak, kristian.sestak@gmail.com

Highlights

- Early-lead excess over the diffusive law is matched by heterogeneity at zero gain.
- The split of the amplified input into skill and position is unidentifiable.
- Observational equivalence, not imprecision: no quantity of data recovers it.
- What identifies instead is a revocation contrast internal to one system.
- Chess: an order that looks manufactured is at least 62% revealed skill ($R \approx 0.38$).

1. Introduction

That an early advantage compounds into a large final one, whether or not it was deserved, is by now well established, and established causally: randomised interventions have planted an arbitrary head start and watched it grow, in artificial cultural markets ^[1], in comment scoring ^[2], and across four domains at once ^[3]. A large observational literature has followed, and it commonly argues from the shape of a trajectory: positions are tracked over time, the early ordering is found to predict the final one far better than diffusion would allow, and that excess is read as evidence that something is amplifying small differences.

This paper is about why that inference does not go through, and what does. The excess is real and easy to reproduce, but it is not diagnostic. A population whose members merely differ in some persistent trait produces the same excess with *no* feedback compounding anything, and we give the closed forms (Eqs. (14)–(15)) and show that such a population has no characteristic lock-in speed at all: sweeping its one free parameter moves the lock-in across the entire range an amplifier occupies, and past it. Magnitude therefore cannot separate amplification from ordinary heterogeneity, let alone identify which mechanism is doing the amplifying. What survives is a *contrast* built inside one system, and we give one that works.

Behind the measurement problem lies a sharper one. Suppose the amplification is granted. The question everyone actually wants answered is then whether the order that formed reflects who was better or merely who got ahead, and that question, we show, has no answer in the trajectory: the persistent input is recovered, but its decomposition into capability and position is not a function of the observed motion. This is observational equivalence rather than imprecision, so no quantity of data bears on it. The split must be imported from a different observable, and what an exogenous channel returns is a proportion.

The obstruction is not new in kind. Panel econometrics has long known that persistent outcomes cannot be separated into true state dependence and persistent unobserved heterogeneity without restrictions, and that the initial condition is the sharpest form of the difficulty ^{[4][5]}. What the multiplicative cascade adds is that the confound becomes *exact* rather than a matter of restrictions one might weaken, two parameter families inducing the same law over trajectories, and that what replaces the lost split is a measurable share rather than an unidentified coefficient. Section 3 locates the correspondence, including what our statement does *not* add to that literature.

The two problems compound, and one dataset shows them doing so. An online-chess cohort (Section 12) wears every appearance the amplifier reading predicts, and a trajectory-only analysis would classify its order as manufactured from noise. Admitting a convergent, non-amplified estimate of skill turns that verdict over: the order is mostly earned. The trajectory could have raised a suspicion, since the cohort's own dispersion betrays that its near-equal entry was an artefact of measurement, but suspicion is as far as it reaches.

Drawing these boundaries requires a model to draw them in, and the model matters less than the boundaries. Status hierarchies form even among people of near-equal ability ^[6], and their inequality is usually read as merit sorted by ability, as a coordination device that a group needs ^[7], or as cumulative advantage in which small leads compound ^{[8][9]}. We work inside a fourth reading, not because we claim it is the operative one, but because it is the reading under which the identification question is sharpest: if a cascade inflates whatever it is given, what survives in the output about what went in? It helps to fix where it sits, as a modelling choice:

$$\text{hierarchy} \xrightarrow{\text{read as}} \text{symmetry-breaking amplifier of position,} \\ \text{as against meritocracy or coordination.}$$

The arrow is deliberate. This is a *reading* of what generates the order, not an identity. Half of it, that the process amplifies, a trajectory can settle; the other half, that what was amplified was position rather than capability, it cannot, which is the non-identifiability theorem of Section 3. The reading becomes a measurable quantity, the manufactured share R of Eq. (7), only through the exogenous channel already described. Which hierarchies fall under it, and to what degree, is then an empirical question rather than a premise, and one that the trajectory signatures answer only in the restricted circumstances Section 8 sets out. With that caveat fixed, meritocracy sorts on a difference that is already there, and coordination answers a functional need for someone to decide. Both take the *direction* of the resulting order as meaningful. The reading we follow is one in which the order has no direction to begin with.

A near-homogeneous population sits at an unstable symmetric state, and the hierarchy is what breaks that symmetry, manufacturing a gap out of noise and calling it “up” only afterwards. This reading is not new. That a hierarchy can form among near-equals from amplified random differences is established [10] [11]. What we add is not the mechanism but a falsifiable *normal form* for it and, principally, a line drawn through what such a hierarchy can and cannot recover about its own origin.

The line has two sides. On one, the *dynamics* (A, whether the process amplifies or merely diffuses) are identifiable from a trajectory. We fix them with four results, each stated so it can fail. A *necessity theorem* settles when amplification, not mere multiplicativity, is required for structure. A *non-monotone dispersion-maximising gain* is one that plain cumulative advantage does not produce. A closed-form *Kesten tail index* (Section 6), checked against a Hill estimate, ties the heaviness of the stationary tail to how contestable the status is. And a $\sqrt{\tau}$ *early-lead law* separates amplification from diffusion without measuring capability, on a homogeneous population directly and, where entities differ, through a contrast the heterogeneity cancels out of.

On the other side, the *seed* (B, whether what got inflated was capability or position, real skill or amplified noise) is *not* identifiable, and this is the boundary the rest of the paper is organised around. A *non-identifiability theorem* (Section 3) has two parts of unequal weight. That the map carries capability and position only through their sum is a structural consequence of an amplifier being blind to the provenance of what it receives; it is close to definitional given that premise, and its interest lies in the premise being forced by the normal form and refutable, since a capability-dependent gain or noise would break it and would leave trajectory signatures. The sharper part is that an inherited head start and the earliest random shock are carried into every later observation through one identical term, which makes them indistinguishable in law and not just in practice. Together they turn the familiar informal statement, that the amplifier keeps no signature of what it amplified, from a metaphor into a property of the likelihood. The remedy is the one already named, an exogenous handle on capability that does not pass through the amplifier.

With the boundaries stated, here is the model that makes them precise; it locates the mechanism inside an existing class. Model an entity’s outcome as governed by two inputs: its own capability k (how many possibilities it can explore) and ρ_{eff} , the density of favourable possibilities the environment makes accessible to it: its position. Suppose the dispersion of the second dominates the first, $Var(\ln \rho_{eff}) \gg Var(\ln k)$, so that where one stands matters more than how able one is. Where does that

large positional dispersion come from? We take it to be *manufactured*: hierarchy is a multiplicative amplifier of position,

$$S_i(t+1) = S_i(t) (1 + gf(S_i(t))) + \text{noise}, \quad (1)$$

with gain g and a saturating feedback f we take, following Gould's model of deference ^[6], to be

$$f(S) = \frac{\theta S}{1 + S/S^*}, \quad (2)$$

linear at the bottom (the heavy-tail engine) and approaching a soft ceiling θS^* at the top (we derive S^* from first principles in Proposition 1). Capability k enters once, at the input, and the cascade then inflates whatever it finds, real merit or noise indifferently, so the grown quantity decouples from capability.

Read as a pathology this cascade is a machine for erasing merit. It reads another way too, and the point is worth making here because Sections 5 and 9 return to it. What makes the cascade a pathology, that it inflates *whatever it finds*, correctness notwithstanding, is also what lets a system move at all when it has no way to know what “correct” is. Consider a homogeneous population, $Var(\ln k) \approx 0$: every member is approximately equal, consistent with the Sah–Stiglitz equal-evaluator assumption and with Gould's finding that status ranks form even among near-equals ^{[6][7]}. Such a population occupies a symmetric configuration with no gradient. There is nothing to select on and no preferred direction. Without amplification it would sit there. The amplifier's function, read positively, is to make that symmetric state untenable: to take an infinitesimal fluctuation and lift it into a macroscopic difference, so the population moves *some* way, not necessarily a good one.

The claim that the direction need not be good is essential and is not a defect of the mechanism but its condition of operation. A system that acts before it can evaluate (a market, an unplanned population) has no oracle for correctness *ex ante*. It has, at each step, only an action and its consequence. Amplification supplies motion in the absence of an evaluator; selection, if it comes, comes afterwards and only on the branches that motion has already generated. Dead ends are therefore not a malfunction but the price of moving without knowing where.

Relation to existing principles

The idea sits at the confluence of four established ones, and our contribution is that a single calibrated cascade realises all four at once: *spontaneous symmetry breaking* (a ball on a symmetric peak falls in a direction fixed by fluctuation, not by any “correct” axis, which Eq. (1) projects onto a macrostate); *Fisher's*

generation of variance ^[12] (adaptation rate scales with the variance to select on, so manufacturing variance is logically prior to selection); *neutral drift* ^[13] (most fixed differences are fixed regardless of fitness, our “not necessarily the right direction”); and *ergodicity economics* ^[14] (a multiplicative process is non-ergodic, the formal face of “the system advances while its median member does not”). None of the four, alone, identifies the amplifier of position as what does this in social and biological hierarchies; that identification is the contribution.

What is new relative to cumulative advantage

That multiplicative growth produces heavy tails is old and well understood, and we claim none of it. It is the Yule–Simon birth process ^{[15][16]}, the Matthew effect ^[8] and preferential attachment ^[9] in the network literature; and, in the statistical-physics literature this paper is addressed to, it is the family of multiplicative stochastic processes with a lower repulsion, exchange or reset term, whose stationary law is a power law of index fixed by a Cramer condition: Levy and Solomon’s logarithmic Boltzmann law ^[17], Sornette and Cont’s convergent processes repelled from zero ^[18], Gabaix’s derivation of Zipf’s law from shocks with a reflecting barrier ^[19], and, closest to the reset dynamics used here, Bouchaud and Mézard’s wealth-condensation model ^[20]. Our Eq. (1) with reverse-dominance renewal sits squarely inside that class, and we use its machinery rather than rederive it: Section 6 computes the resulting tail index in closed form.

Closer still, that a hierarchy forms among near-equals from *amplified random differences* rather than merit is itself established: in networked endorsement dynamics ^[10], where a hierarchy emerges only once winner effects overcome noise (our diffusion-to-amplification threshold), and in group-hierarchy models ^[11], where initially random advantages self-reinforce independently of the agents’ intrinsic quality (our decoupling). We take that qualitative phenomenon as given, and we are not adding to it. The mechanism is not ours, the multiplicative machinery is standard, and the propositions below are elementary by design. What separates this paper from that literature is the direction of the question. It asks what mechanism generates the order; we ask what an observer of the resulting order can recover about the mechanism. That is an identification question, its answer is largely negative, and the negative part is the contribution: several signatures in current use do not identify what they are taken to identify, and saying so precisely is worth more here than another mechanism would be. Four differences from that literature follow, cutting across the results listed above. First, a *necessity* theorem (Proposition 3): without the feedback the process is a structureless martingale, so within the model amplification is not one route

to inequality among many but the precondition for any structured development. Second, a *dispersion-maximising* gain: the saturating feedback makes manufactured dispersion non-monotone in strength, a downturn plain preferential attachment (no per-unit ceiling) does not produce, and it is horizon-dependent rather than an intrinsic optimum (the robust claim is the non-monotonicity itself). Third, a *scope* condition: the effect is confined to similar entities and dissolves into ordinary competence as $Var(\ln k)$ grows (Section 5). Fourth, the *decoupling from correctness* is turned from a qualitative “the rich get richer” into a quantity with a sign and a scale, the manufactured share R , though only once an exogenous handle on capability is admitted, since the non-identifiability theorem of Section 3 shows a trajectory alone cannot supply it. Cumulative advantage says leads compound; it does not say when compounding is required, when it is self-limiting, and when it erases merit.

Alternative mechanisms, and how to tell them apart

The amplifier is one route to hierarchy, not the only one, and we do not claim it is always the operative one. Hierarchy can also arise from (i) genuine ability differences sorted by a meritocratic process; (ii) a functional need for coordination, where someone must decide, as in the architecture of Sah and Stiglitz [2]; (iii) rank-order tournaments and contests, where prizes are assigned by relative performance; (iv) inherited or seized resources, where position is transferred rather than grown; and (v) network position from homophily or chance connectivity. These are not rivals to be dismissed but alternatives to be distinguished, and the model offers three signatures that do the distinguishing. Merit sorting predicts outcome tracks capability ($\rho_{eff} \parallel k$); the amplifier predicts the opposite, increasingly so with gain (Section 4). Coordination or tournaments predict a bounded spread set by roles or prizes; the amplifier predicts a heavy tail from reset-plus-drift (Section 5). Inheritance predicts the order is set at the input and preserved; the amplifier predicts it is *manufactured* from near-nothing when $Var(\ln k)$ is small, inherited only when it is large (the scope condition). Which mechanism dominates is an empirical question these signatures make answerable *in principle*, and the qualifier is doing real work: the first and third of them are stated in terms of capability, which Section 3 shows a trajectory does not supply, so they are usable only where an exogenous handle is. Our claim is only that where the amplifier operates, it has these consequences.

What the model represents

A fair objection is that no single multiplicative equation can stand for every hierarchy. It is not meant to: it is a *normal form* for the class of hierarchies generated by multiplicative amplification of position, its three ingredients each forced (multiplicativity is “position begets position”; saturation is finiteness, derived as a carrying capacity in Proposition 1; reset and noise are a stochastic, revocable environment). What we claim is pitched at the level of the class, not the parametrisation: the robust consequences (a log-normal bulk with a power-law tail, a Kesten class ^[21] insensitive to the exact f ; non-ergodicity; gain-growing decoupling; the ceiling) depend only on this structure, as Table 1 shows, while the specific magnitudes depend on calibration and are not claimed. Hierarchies dominated instead by ability, coordination, tournaments, or inheritance lie outside the class, and the model’s signatures are how one tests whether a given hierarchy lies inside it, subject throughout to the caution of Section 8: read as magnitudes they test nothing, and only within-system contrasts carry the inference. The interpretive statements of Section 9 are readings of these model properties, not theorems about them; a reader who declines the interpretation keeps every result. The mathematics establishes properties of a model; the philosophy is offered separately, and as such.

2. The symmetric state is unstable, and hierarchy manufactures the dispersion

We simulate Eq. (1) in log-status $x = \ln S$ (the schematic additive “noise” of Eq. (1) is realised, here and throughout, as Gaussian noise in the log, i.e. multiplicatively in S),

$$x_i(t+1) = x_i(t) + \ln(1 + gf(S_i)) + \eta_i, \quad \eta_i \sim \mathcal{N}(0, \sigma^2), \quad (3)$$

with the Gould feedback (2), starting from a homogeneous state $x_i(0) = 0$ perturbed only by infinitesimal noise (10^{-6}). We use an illustrative calibration ($\sigma = 0.05$, $\theta = 0.02$, $S^* = 50$; the qualitative claims are shown robust to it in Table 1) and vary the gain g .

The first lesson is that dispersion grows *even at* $g = 0$, but only diffusively: with no feedback, Eq. (3) is a random walk and $\text{Var}(x) \approx \sigma^2 t$ accumulates linearly from noise alone. So “symmetry breaks” in the trivial sense at any temperature; the question with content is whether the growth is diffusive (additive, slow) or amplified (multiplicative, fast). Hierarchy makes it the latter. The clean, gain-controlled quantity is therefore not an escape time, which is set mostly by how fast noise seeds the initial spread and only weakly by g , but the *amplification factor*

$$A(g) \equiv \frac{Var_g(T)}{Var_0(T)}, \quad (4)$$

a “manufactured by hierarchy” term: how much dispersion the structure adds beyond the diffusion a flat population would produce anyway. By construction $A(0) = 1$.

Figure 1 shows the result. In the left panel the $g = 0$ curve grows sub-linearly on the log axis (diffusion), while the larger gains turn over into a multiplicative phase and reach a far higher plateau. The turnover is not instantaneous and the smallest gains have not reached it by $T = 300$: at $g = 0.05$ the curve is still essentially diffusive ($A = 1.9$), which is what the blow-up time $t^* \approx 1/(g\theta S_0)$ of the next paragraph requires, since it exceeds the horizon there. In the right panel $A(g)$ rises steeply, to $\sim 400\times$ near $g \approx 0.4$, and is then *non-monotone*: at higher gain it falls back. This is not numerical noise. It is the saturating ceiling of Eq. (2): at very high gain every entity races to the ceiling θS^* and clusters there, so the dispersion the amplifier was manufacturing is destroyed by the same insurance that stabilises the amplifier. At any given horizon there is thus a *dispersion-maximising gain*. Too little and the population only diffuses; too much and Gould’s saturation collapses the spread. The manufactured dispersion is maximal in between. As the next paragraph shows, that maximising gain is a property of the observation window, not an intrinsic constant.

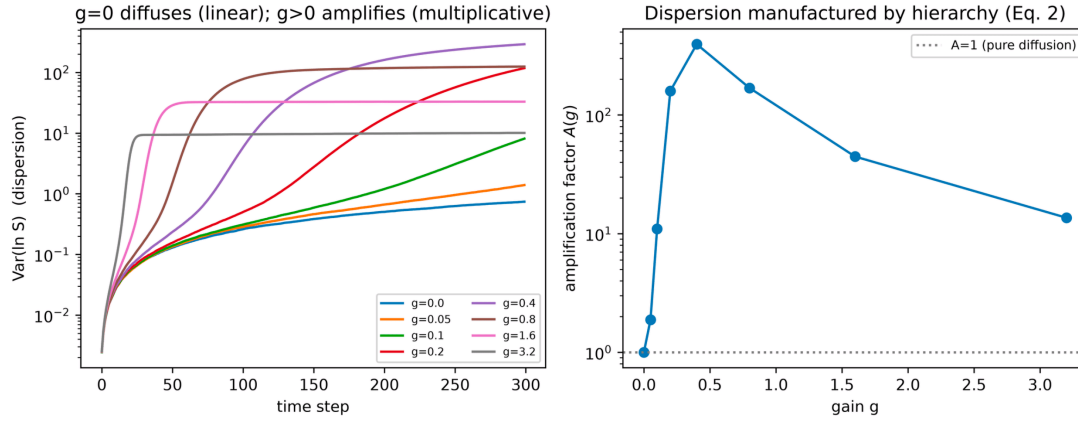


Figure 1. The homogeneous state is dynamically unstable, and hierarchy manufactures dispersion multiplicatively. *Left:* from a homogeneous start, $g = 0$ (blue) grows dispersion only diffusively ($\text{Var} x$ linear in t , hence concave on the log axis), while sufficient gain turns the curve over into a multiplicative phase and up to a plateau; the smallest gains do not reach their turnover within this horizon. Note that the highest gains ($g = 1.6, 3.2$) saturate lower than $g = 0.4$, which is the non-monotonicity seen directly. *Right:* the amplification factor $A(g) = \text{Var}_g / \text{Var}_0$ (Eq. (4)), the dispersion manufactured by hierarchy, rises to $\sim 400\times$ and is non-monotone: the saturating feedback (2) that stabilises the amplifier also caps and then destroys the dispersion at high gain. The dispersion-maximising gain is horizon-dependent (Eq. (5)), not an intrinsic constant.

The gain that maximises dispersion is horizon-dependent, not intrinsic

The peak in Figure 1 is a crossover of two timescales, and a property of the observation window, not a universal constant. In the linear-bottom regime $f(S) \approx \theta S$ the mean status blows up in finite time, $\dot{S} \approx g\theta S^2$, reaching the ceiling at $t^*(g) \approx 1/(g\theta S_0)$. Dispersion is manufactured only while the map is stretching ($\varphi' > 1$); once the trajectory hits the ceiling the remaining horizon adds only diffusion. The manufactured dispersion is therefore maximal where the blow-up time meets the horizon, $t^*(g^*) \approx T$, i.e.

$$g^*(T) \approx \frac{1}{\theta S_0 T}. \quad (5)$$

So there is no single “optimal gain”: at each horizon T a different g maximises $A(g)$, and as $T \rightarrow \infty$ it tends to zero.

That much is right, but the exponent is not. Locating the peak on a dense log-spaced gain grid, with the maximum refined to sub-grid resolution and the horizon extended to $T = 2400$ (Figure 2; `optimal_gain_fit.py`), gives $g^* = 0.62, 0.37, 0.24, 0.17, 0.14, 0.13$ at $T = 150, 300, 600, 1200, 1800, 2400$. A log-log regression returns $g^* \sim T^{-\alpha}$ with $\alpha = 0.56 \pm 0.02$ over seeds, and restricting to the asymptotic half of the range ($T \geq 600$), where the noise-seeding transient occupies a smaller share of the horizon, moves it *further* from unity, to 0.44 ± 0.04 , not closer. Correspondingly the crossover constant $g^* \theta S_0 T$ drifts monotonically from 1.9 to 6.4 across the range (computed from the unrounded g^*) instead of settling. Equation (5) therefore captures the *direction* (the maximising gain falls with the horizon and vanishes asymptotically) but not the exponent, which is closer to $T^{-1/2}$; the linear-regime blow-up argument evidently omits the noise-seeding stage that sets the effective S_0 , and that stage does not become negligible at the horizons reachable here. An earlier four-point scan on a coarse gain grid appeared to confirm $1/T$; that appearance was an artefact of the grid spacing, and the dense-grid fit supersedes it. What the paper needs from this section is the non-monotonicity, and it is unaffected by the exponent.

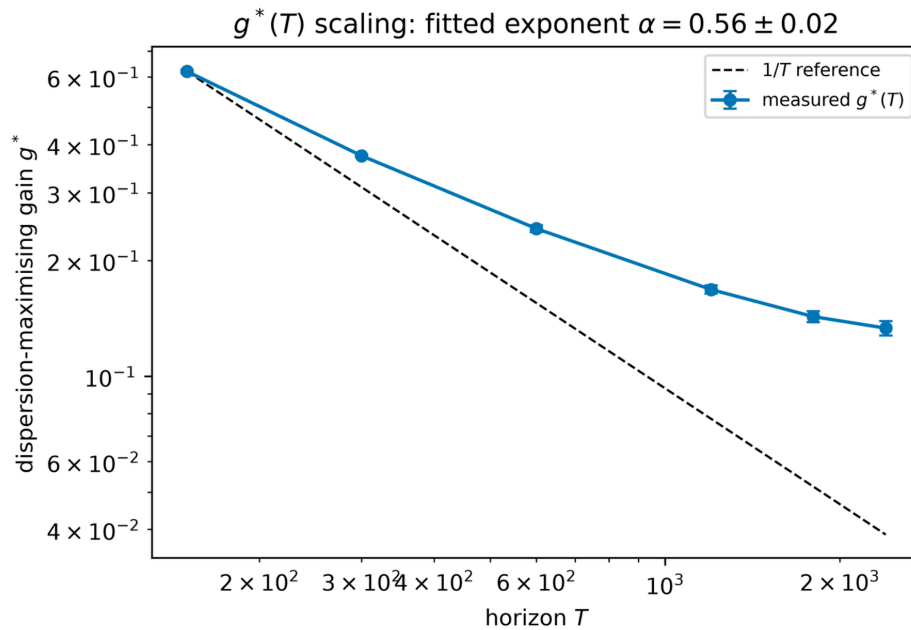


Figure 2. The dispersion-maximising gain falls with the horizon, but not as $1/T$. Peak of $A(g)$ located on a dense log-spaced gain grid with sub-grid refinement, over six horizons and three seeds (error bars are the spread over seeds); the dashed line is a $1/T$ reference anchored at $T = 150$. The fitted exponent is $\alpha = 0.56 \pm 0.02$, and restricting to $T \geq 600$ gives 0.44 ± 0.04 , moving away from unity rather than towards it. The mean-field crossover (5) therefore gives the direction but not the exponent (`optimal_gain_fit.py`).

Where the ceiling comes from

The self-limitation need not be assumed, and its origin is more general than Gould's deference. It is a maintenance carrying capacity.

Proposition 1. (The ceiling is a maintenance carrying capacity). *Let a self-reinforcing quantity S be acquired preferentially (at a rate increasing in S) while its upkeep draws a bounded flow Φ shared across its S units, so per-unit service is Φ/S and a unit is retained only while $\Phi/S \geq \varrho$. Then S is capped at $S^* = \Phi/\varrho$, independent of the acquisition rate. The cap is finite iff Φ is; a quantity needing no upkeep ($\varrho = 0$) has none.*

Proof. Retention fails once $\Phi/S < \varrho$, i.e. $S > \Phi/\varrho$; acquisition, however aggressive, cannot hold S past the point where per-unit service drops below viability. $\square$

Gould's $S^* = D/\kappa$ is the instance with Φ the reciprocation budget D and $\varrho = \kappa$; the smooth feedback follows as $f(S) = \theta S \cdot \Phi/(\Phi + \varrho S) = \theta S/(1 + S/S^*)$. The same structure governs a metabolic

throughput [22], a cognitive bandwidth [23], or a monitoring capacity [24]: whatever the flow, S^* is upkeep budget over per-unit demand. The prediction is falsifiable and non-obvious, the ceiling being set by upkeep, not ambition: in simulation with a preferential giving rule and no imposed cap, $S^*/D = 1$ across a factor-sixteen range of the acquisition gain (`regulation.py`). Hence a criterion for *which* systems self-regulate: a currency that must be actively maintained (reciprocated, monitored, defended) has an S^* ; a freely self-replicating token (download counts, followers, a ledger number) has $\varrho = 0$, no ceiling, and runs away until an external limit binds, which is why the optimal-gain downturn is absent in the download market of Section 10.

3. The seed is unidentifiable: what the amplified output cannot recover

The previous section concerned the *map* question, which we call A. Whether the dynamics amplify or merely diffuse is a property of the trajectory and is recoverable from it. We now isolate the *seed* question, B, and prove it is not. Write the per-entity latent input to the cascade as a sum of a capability part and a positional part,

$$a_i = k_i + p_i, \quad (6)$$

where $k_i = \ln k$ is (log-)capability and $p_i = \ln \rho_{eff}$ is (log-)position, the density of favourable possibilities the environment makes accessible. The amplifier acts on the entity through a_i alone. Capability and position enter the same multiplicative channel (Eq. (1)), and this is the sense in which the cascade “inflates whatever it finds.”

Proposition 2 (Non-identifiability of the seed). *Consider the linearised amplifier in log-status, $x_i(t+1) = \lambda x_i(t) + a_i + \eta_i(t)$, with gain-induced slope $\lambda = \varphi' > 1$ (Proposition 3), $\eta_i(t) \sim \mathcal{N}(0, \sigma^2)$ i.i.d., and latent input (6). Let the cohort start from the common observed position $x_i(0) = 0$ (the shared provisional entry rating of Section 12 is the empirical instance), and let any inherited head start enter as a latent initial impulse w_i carried by the first step, $x_i(1) = a_i + w_i + \eta_i(0)$: what is inherited is precisely what the entry position does not show. Let the observable be the entire trajectory $\mathbf{x}_i = (x_i(0), \dots, x_i(T))$. Then:*

- i. (Static: components are non-separable.) *The law of $\{\mathbf{x}_i\}$ depends on (k_i, p_i) only through the sum $a_i = k_i + p_i$. Hence the family $\{(k_i + \delta_i, p_i - \delta_i) : \delta_i \in \mathbb{R}\}$ is observationally equivalent: no statistic of the trajectory distinguishes a high-capability, low-position entity from a low-capability, high-position one with the same a_i .*

ii. (Dynamic: a lucky early draw is inseparable from an inherited seed.) In the solution $x_i(T) = \lambda^{T-1}w_i + \sum_{s=0}^{T-1} \lambda^{T-1-s} (a_i + \eta_i(s))$, the inherited seed w_i and the earliest noise increment $\eta_i(0)$ enter every coordinate $x_i(t)$, $t \geq 1$, through the single term $\lambda^{t-1} (w_i + \eta_i(0))$: only their sum is carried, with the identical weight. The noise variance $\text{Var} [\sum_s \lambda^{T-1-s} \eta_i(s)] = \sigma^2 \frac{\lambda^{2T}-1}{\lambda^2-1}$ is dominated by the smallest s . An entity that received a favourable early fluctuation is therefore observationally identical to one that inherited a head start: “manufactured by early noise” and “inherited from the seed” induce the same law on $\{\mathbf{x}_i\}$.

Consequently the persistent amplified component a_i is identified, but its decomposition is not. This holds both for the split into capability and position and, within position, for the split into a recurrent environmental advantage and an amplified transient fluctuation.

Proof. Solving the linear recursion from the common start $x_i(0) = 0$ gives, for every i ,

$$x_i(T) = \lambda^{T-1}w_i + \left(\sum_{s=0}^{T-1} \lambda^{T-1-s} \right) a_i + \sum_{s=0}^{T-1} \lambda^{T-1-s} \eta_i(s),$$

and analogously for each intermediate $x_i(t)$ with T replaced by t . Every coordinate of $\mathbf{x}_i$ is thus an affine function of (w_i, a_i) plus a Gaussian term whose covariance depends only on λ, σ, T . The map $(k_i, p_i) \mapsto \text{law}(\mathbf{x}_i)$ factors through $a_i = k_i + p_i$; replacing (k_i, p_i) by $(k_i + \delta_i, p_i - \delta_i)$ leaves a_i , hence the entire law, unchanged, proving (i). For (ii), note that w_i and $\eta_i(0)$ enter each coordinate $x_i(t)$, $t \geq 1$, only through the single term $\lambda^{t-1} (w_i + \eta_i(0))$, while $x_i(0) = 0$ carries neither; the geometric sum $\sum_s \lambda^{2(T-1-s)} = (\lambda^{2T} - 1)/(\lambda^2 - 1)$ is maximised term-wise at $s = 0$. A shift $w_i \mapsto w_i + \varepsilon$ and a shift in the realised $\eta_i(0)$ of ε produce identical values of every coordinate of $\mathbf{x}_i$; the two data-generating stories are therefore indistinguishable at the level of the observed trajectory. Identifiability of a_i itself follows because the drift is the only recurrent (per-step) term and is recovered from the trajectory slope; only its composition is lost. $\square$

Which half of this is substantive

The two parts of Proposition 2 carry unequal content. Part (i), taken by itself, is close to a tautology: if capability and position enter the dynamics only through the sum $a_i = k_i + p_i$, then no observable separates them, and the algebra is bookkeeping that records the modelling decision made in Eq. (6). The work is done elsewhere.

Single-channel entry is *forced* rather than chosen, and that converts (i) from bookkeeping into a restriction with empirical content. A model in which capability entered a second time, by scaling the gain ($\lambda_i = \lambda(k_i)$) or the noise ($\sigma_i = \sigma(k_i)$), would break (i) outright: an entity's capability would then show up in the curvature or the heteroskedasticity of its own trajectory, and k would be identified from the trajectory alone without any exogenous channel. What Eq. (6) claims, and what a critic should attack, is therefore not the algebra but the physical premise: that an amplifier of position is blind to the provenance of what it receives, so capability can act only by changing where an entity stands and never by changing how the cascade treats it. Where that premise fails the non-identifiability fails with it, and the failure is testable, since capability-dependent gain and capability-dependent noise both leave trajectory signatures. Read this way, (i) is a statement about a class of mechanisms, and it is the class the paper is about.

Part (ii) does the rest, and it is not tautological under any reading. It says that an inherited head start w_i and the earliest noise increment $\eta_i(0)$ are carried into *every* later coordinate by the one term $\lambda^{t-1}(w_i + \eta_i(0))$, with the identical weight. Nothing in the setup arranges this; it follows from where in the recursion the two enter, and it is what makes “manufactured by early luck” and “inherited from the seed” observationally identical rather than merely hard to tell apart. This is the amplified form of the initial-conditions problem, and it is the part that uses the linearisation, which serves only to make the equivalence exact rather than local. Part (i) is structural and survives the nonlinearity: in the full dynamics (1), with or without the reset, the entity's law depends on (k_i, p_i) only through the channel by which they enter, whatever the feedback f and however far from linear the map. The proofs are elementary by design; what the proposition contributes is the identification boundary it draws, not the machinery that draws it.

What is not recoverable, and in what sense

The result is easy to overstate in either direction, so here it is exactly: *the decomposition of the amplified input is not a function of the trajectory.*

What is lost is the decomposition, not the input. The persistent component a_i is recovered from the drift, as are the dynamics of Section 2; it is the split that survives nowhere, whether into capability and position or, within position, into a recurrent advantage and an amplified transient. Nor is that split badly estimated, which would be a different and milder complaint. Proposition 2 says that two families of parameters induce the *same* distribution over observables, so the difficulty is not a noisy estimator but

the absence of any consistent one, and the usual remedies do not apply: longer horizons, larger cohorts and denser sampling leave a flat likelihood flat. The signature is not hard to read; it is absent.

That is a claim about an observable, though, not about the world. Nothing here says the split is unknowable, only that it is unknowable from a trajectory. A different kind of observation settles it immediately, and the discriminator below names three that do, with Section 12 a case where one of them overturns the trajectory-only reading outright. What such a channel returns, finally, is a proportion: the manufactured share R of Eq. (7) lands somewhere between pure symmetry-breaking and pure revelation. Asking whether a hierarchy was earned or manufactured is therefore the wrong question, and would remain so even with the exogenous channel in hand.

This relocates the philosophical reading of Section 9. The statement that “forward” is a label applied afterwards is no longer an interpretation laid over the model but a corollary of (6), since the map carries the sum and not the split.

Relation to state dependence versus unobserved heterogeneity

The sibling of this result named in Section 1 can now be located exactly. In panel data, persistent outcomes can reflect true state dependence (the past outcome causally raising the future one) or persistent unobserved heterogeneity, the two are observationally confounded without restrictions, and the initial condition is the sharpest obstruction ^{[4][5]}. Two things follow, a difference and a debt. Question A is the analogue of state dependence ($\lambda > 1$ against $\lambda \leq 1$), and, consistent with that literature, we treat it as identifiable, though by different means (variance growth, the non-monotone gain, and within-system contrasts rather than serial-correlation restrictions). Proposition 2 is posed one level deeper: it grants the state dependence and shows that the composition of the persistent individual effect a_i , capability against position, remains unidentified even then, with part (ii) as the amplified form of the initial-conditions problem (w_i against $\eta_i(0)$). The remedy also differs. Where the panel literature buys identification with functional-form and error-structure assumptions, the discriminator below buys it with an exogenous channel to one component, and returns a share rather than a coefficient.

The debt should be stated as plainly as the difference. That a persistent individual effect resists decomposition is not our discovery, and a reader who knows that literature will find the obstruction familiar; what Proposition 2 adds is that in a multiplicative amplifier the confound is exact rather than a restriction one might relax, that the single channel carrying it is forced by the normal form and refutable, and that the quantity replacing the lost split is the share R of Eq. (7) rather than a coefficient one declines

to interpret. The elementary proof is the point: no machinery is needed once the channel structure is granted.

The discriminator, and the decomposition it returns

Identifying B requires a second observation of one component through a channel that does *not* pass through the amplifier's early lock-in, that is, an exogenous handle on k . Three such channels exist. Capability can be held fixed *by construction* (parallel worlds, $Var k = 0$, so $a = p$). The seed can be *randomised* ($k \perp$ assignment, so the persistent component is p by design). Or a *convergent independent estimate* $\hat{k}$ can be obtained by a different sampling that averages the early fluctuation away (a later equilibrium). Given any such $\hat{k}$, the split becomes a regression and the identified quantity is a share,

$$R = \frac{Var(p)}{Var(a)} \leq 1 - corr^2(\text{final order}, \hat{k}), \quad (7)$$

the manufactured fraction of the dispersion. The bound assumes capability and position are uncorrelated ($k \perp p$) and that $\hat{k}$ is an unbiased estimate of k , and it is a bound rather than an equality for two reasons, both in the same direction. First, the final order is $\Lambda a_i + N_i$, with $\Lambda = \sum_{s=0}^{T-1} \lambda^s$ the amplification of the persistent input and N_i the accumulated trajectory noise, so $1 - corr^2 = (\Lambda^2 Var p + Var N) / (\Lambda^2 Var a + Var N) \geq Var(p) / Var(a)$, with equality only in the noiseless limit $Var N = 0$: the right-hand side charges the amplified noise to manufacture, which is what one wants of it, but that makes it exceed the seed-level share $Var(p) / Var(a)$. Second, measurement error in $\hat{k}$ attenuates the correlation and inflates the right-hand side again. Both give an *upper* bound on manufacture, so a small measured value is the informative case; the empirical sections report the measured right-hand side and call it R throughout, read in that direction. Because the share in (7) is a decomposition of *variance*, the correlation in it must be a Pearson correlation on the scale the model is written in, namely log-status; a rank correlation does not decompose variance, and $1 - \rho_S^2$ is not a variance share. We therefore compute R from Pearson correlations of log-ratings throughout, reporting rank correlations alongside as the monotone-invariant summary of the ordering. Here $R \rightarrow 1$ is the pure symmetry-breaking limit and $R \rightarrow 0$ pure revelation. It is the same manufactured share that the scale scan of Section 5 reports, falling from 1.00 towards zero as $Var(\ln k)$ grows. This is why the free-token domains of Section 10 (downloads, stars, citations) settle nothing about B on their own. Lacking any such channel, they leave R a free parameter. Where a channel exists the number is finite and informative. The online-chess worked example of Section 12 measures $R \approx 0.38$ for a cohort of near-equal entry ratings,

so an order that looked manufactured, in that it was uncorrelated with the entry position, is roughly 62% revealed skill once the convergent $\hat{k}$ is admitted.

4. An illustration of the non-separability: direction versus correctness

Proposition 2(i) is best seen in a simulation built as its special case. We present it as an *illustration of the construction* rather than as an independent finding. When capability enters through the same channel as position, the direction the amplifier selects decouples from correctness by construction, and it does so increasingly as amplification grows, so the positional density ρ_{eff} ends up orthogonal to capability k .

We test this in the most favourable case for merit. Give every entity a true, fixed quality $q_i \sim \mathcal{N}(0, 1)$ (the capability k) and let it enter as an *honest, deterministic* per-step drift: quality helps, every step, with no noise on its own contribution. If any design should keep the winner correlated with quality, this one should. We then measure the rank correlation between the final ordering and q as the gain g rises.

Figure 3 shows the correlation falling monotonically from 0.94 at $g = 0$ (no amplification: quality plus diffusion) to 0.66 at $g = 0.8$. The amplifier does not reverse merit, since the correlation stays positive, but it progressively *decouples* the winner from the quality, because an early environmental draw, once amplified, outruns a steady honest advantage. This is $\rho_{eff} \perp k$ shown as *generated* by the gain, not assumed: the more hierarchy amplifies, the less the direction of motion tracks who was actually right. The decoupling is *partial*, not total, and it is *monotone in gain*, so the same knob that manufactures dispersion (Section 2) is the knob that severs it from correctness. Amplification and decoupling are one process.

What is construction here, and what is not

A multiplicative amplifier that inflates *whatever it finds* will decouple outcome from quality more or less by definition, and a simulation of it cannot “discover” otherwise. The qualitative direction, that gain erodes the outcome–quality correlation, is therefore a consequence of the model’s construction, and we do not present it as a finding. What the simulation adds is quantitative: that the erosion is *monotone* in gain rather than a threshold (monotone net of sampling noise: the market model of Figure 8 is flat within noise over its lowest gains before falling), that it is *partial* (a decline to 0.66, not a collapse to zero) under the most merit-favourable additive design, and that its rate is set by the same g that fixes the dispersion

of Section 2. The empirical weight of the claim rests not on this figure but on whether the decoupling appears where quality is held fixed by *construction of the data*: the Music Lab experiment of Section 11, where identical songs across worlds make between-world differences causally attributable to the amplifier and nothing else. The simulation states the prediction; the experiment is what can refute it.

The additive design is the conservative one

A natural objection is that we let quality enter *additively*, on its own noise-free channel, which may understate merit's staying power. The opposite is true. If quality instead enters *multiplicatively*, scaling each entity's gain, so that better entities amplify faster, the correlation with quality is even *lower* ($\lesssim 0.10$ at every gain, versus 0.66–0.94 additively; the comparison is informative only at $g > 0$, since at $g = 0$ a quality that acts by scaling the gain has no channel at all and the correlation is trivially zero, whereas at $g = 0.8$ it is 0.10 against 0.66), because the amplifier cannot tell a quality-driven increment from a noise-driven one and magnifies both identically: an early environmental draw into a high position outruns a real but proportionally amplified advantage. Merit that must flow through position is thus *more* decoupled, not less; the additive design of Figure 3 is the case most favourable to competence, and it already decouples. (Consistent with the scope condition of Section 5, a large enough quality spread would still dominate, since the decoupling is a claim about similar entities.) One caveat bounds the reading: the decoupling holds for competence expressed as an *ongoing* capacity; a fixed initial endowment, injected once and then amplified like any other early input, is instead preserved by the gain rather than eroded by it.

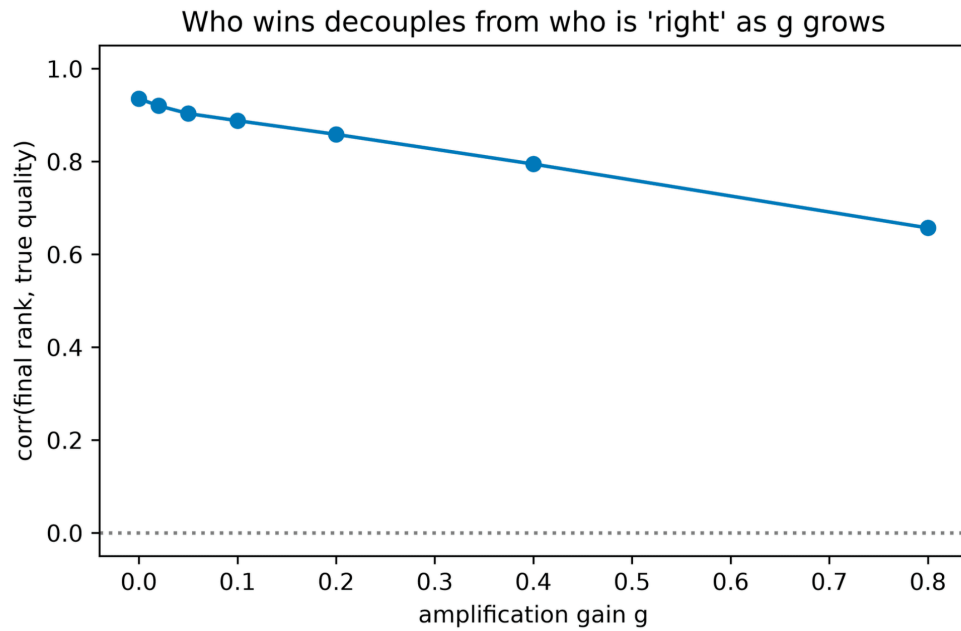


Figure 3. The direction of motion decouples from correctness as amplification grows. Even when true quality enters as an honest per-step drift (the best case for merit), the rank correlation between the final winner and true quality falls monotonically with gain, from 0.94 ($g = 0$, no amplification: quality plus diffusion) to 0.66 ($g = 0.8$). The decoupling $\rho_{eff} \perp k$ is generated by the gain; it is partial, not a reversal.

5. Within the model, amplification is necessary for development, not just for its size

Section 2 showed the amplifier controls *how much* dispersion appears. A stronger claim holds within the model: without the multiplicative feedback the system does not *develop* at all, in the sense of forming persistent, structured differentiation; it at most diffuses. Unlike a functionalist appeal this is a statement about the dynamics, not a value judgement. It also corrects a natural but wrong intuition: absent hierarchy the entities do not literally freeze (noise still moves them); what they lose is not motion but *development*.

Proposition 3 (Amplification is necessary for development in the model). *Take the dynamics (3) with distinct initial attributes $x_i(0)$. If $g = 0$, then*

- i. each entity's expected state is invariant, $\mathbb{E}[x_i(t)] = x_i(0)$ for all t (the population stays in its original conditions in expectation);
- ii. the process is a martingale, so no lead compounds: the cross-sectional variance grows only diffusively, $Varx(t) = Varx(0) + \sigma^2 t$ (linear, not exponential), and the ordering decays, $corr(x_0, x_t) = [1 + \sigma^2 t / Varx_0]^{-1/2} \rightarrow 0$;
- iii. the cross-sectional distribution converges to Gaussian once rescaled by $\sigma\sqrt{t}$ (the increments being Gaussian already, no limit theorem is needed: the law is the initial distribution convolved with $\mathcal{N}(0, \sigma^2 t)$, which the growing noise variance dominates): no heavy tail, hence no new level of structure forms.

Conversely, any $g > 0$ makes the per-step map $\varphi(x) = x + \ln(1 + gf(S))$ have slope $\varphi'(x) = 1 + gf'(S)S/(1 + gf(S)) > 1$ wherever $f' > 0$, a positive Lyapunov exponent that breaks (i)–(iii): leads compound, variance grows super-linearly, and, with the reset of Section 7, a heavy tail forms. Amplification is therefore necessary for directed, structure-forming development; without it the system only diffuses.

Proof sketch. At $g = 0$, (3) reduces to $x_i(t+1) = x_i(t) + \eta_i$ with $\mathbb{E}\eta = 0$, so $\mathbb{E}[x_i(t+1) | x_i(t)] = x_i(t)$; iterating gives (i), and the martingale property is immediate. Writing $x_i(t) = x_i(0) + \sum_{s \leq t} \eta_{is}$ with independent increments, the cross-sectional variance is $Varx(0) + \sigma^2 t$ and $corr(x_0, x_t) = Varx_0 / \sqrt{Varx_0(Varx_0 + \sigma^2 t)}$, which is (ii); the accumulated noise is exactly Gaussian and its variance $\sigma^2 t$ swamps $Varx_0$, giving (iii). For $g > 0$ the stated slope $\varphi' > 1$ makes infinitesimal differences grow geometrically, so the martingale and Gaussian conclusions both fail. $\square$

The proposition is stated for the reset-free dynamics. Adding the reverse-dominance reset of Section 7 at $g = 0$ replaces the Gaussian in (iii) with a non-Gaussian but still light-tailed, mean-reverting stationary law (a killed-and-restarted random walk): it changes the shape of the bulk but adds no heavy tail and no persistence, so the operative conclusion, no structured development without amplification, is unchanged. A heavy tail needs the reset *and* the multiplicative drift together, the Kesten mechanism [21], which requires $g > 0$.

The three signatures are visible in simulation (Figure 4), where entities are given genuinely distinct initial attributes and either develop them or wash them out. At $g = 0$ the variance grows linearly, the final distribution is Gaussian, and the initial ordering decays (measured $corr = 0.35$ at $T = 300$, matching the predicted $[1 + \sigma^2 T / Varx_0]^{-1/2} = 0.33$); at sufficient gain the variance runs away, a heavy right tail appears, and the ordering locks in ($corr$ rising to 0.70 at $g = 0.8$). As in Section 2, the separation from diffusion is not immediate in g : the variance ratio at $g = 0.05$ is only twice the diffusive one over this

horizon. The histogram here shows only the qualitative tail signature, a heavy right shoulder and positive excess kurtosis at moderate gain; Section 6 makes the tail quantitative and explains why the figure cannot show a fitted power law at this calibration. The reading is not that hierarchy preserves merit: by Section 4 what locks in may be amplified noise as easily as a real attribute. The reading is that *some* structure forms only under amplification, and which structure forms is exactly what the amplifier then makes unrecoverable.

Scope: this is a claim about similar entities

Proposition 3 concerns the regime of interest, $Var(\ln k)$ small. When the initial attributes instead differ by orders of magnitude, the amplifier no longer *manufactures* the hierarchy; it only *magnifies* one that is already there. Sweeping the initial-attribute spread s_0 at fixed gain (`scale_scan()`) makes this quantitative: the share of the final variance the cascade manufactures, rather than inherits from the initial attributes, falls from 1.00 at $s_0 = 0.01$ (near-identical entities, final order set entirely by amplified noise, $corr(x_0, x_T) \approx 0$) to 0.20 at $s_0 = 6$ (order-of-magnitude differences, $corr(x_0, x_T) \approx 0.90$, the order reflects real capability). This manufactured share is the R of Eq. (7), here measured against the initial attributes, which in the simulation are the injected true capability and so serve as an exact $\hat{k}$ (unlike a provisional entry position such as the chess entry rating of Section 12), and it falls towards zero as $Var(\ln k)$ grows, so the amplifier’s manufactured contribution vanishes and the ordering comes to express real capability. A hierarchy still forms among very different entities, even more extreme in magnitude, but it now expresses competence rather than being decoupled from it. “Position, not competence” is thus a statement about *similar* entities specifically; make them different enough and the amplifier becomes an amplifier of competence.

The scope condition is not directly testable

A caution follows from Proposition 2 and bears on every empirical use of the condition just stated. The condition is predicated on $Var(\ln k)$, which is not observable at entry: an entry measurement is produced by the same channel that has not yet had the chance to learn anything, so what a narrow entry band demonstrates is a narrow dispersion of the *estimate*, and this is consistent both with entities that are alike and with entities the measurement has merely not yet told apart. Establishing that a population really sits in the small- $Var(\ln k)$ regime therefore requires the same exogenous handle on k that Section 3 requires for R , and is unavailable without one. The online chess of Section 12 is the cautionary instance: an entry band of 40 Elo points, an apparently symmetric start, and a month-end dispersion that

turns out to equal the whole population's, because the seeded rating recorded ignorance rather than equality. The practical rule is that “these entities started out alike” is a claim about the evidence unless a channel to k says otherwise, and the amplifier reading may not rest on it until one does. This does not weaken the scope condition as a statement about the model, where $Var(\ln k)$ is a parameter one sets; it bounds what may be concluded from a symmetric-looking start in data, and it is the same boundary the non-identifiability theorem draws elsewhere, met here from the other side.

Beyond making development possible, the same instability makes it *fast*: while the initial *escape* from the symmetric state is seeded by noise and depends only weakly on g (Section 2), amplification reaches a *locked* dominant configuration exponentially faster than diffusion and, unlike diffusion, holds it, so a system with no oracle settles on a path quickly rather than never (in simulation, ≈ 130 steps to a locked dominant share at $g \approx 0.2$, against a weak, unlocked pseudo-leader only at ≈ 1500 steps under diffusion). Whether that path is a dead end does not bear on the speed; the system develops either way.

The theorem, and what would refute it

Proposition 3 is a deductive fact about the *model*: at $g = 0$ the dynamics are a martingale, so by (ii)–(iii) the order washes out and no heavy tail forms. Within the equation this is certain. The empirical claim, that real ranking-forming systems fall under the equation, is separate and falsifiable, and it takes the form of a taxonomy: a persistent, self-generated, heavy-tailed order requires the loop; without it a system yields at most a *transient* order (diffusion washes it out) or an *imposed* one (set at the input or by an external gradient and merely conserved, the inheritance case of Section 4). The conjunction *self-generated, persistent, and loop-free* is the empty cell the model predicts no system occupies. A refutation must live there: a near-homogeneous population that develops a persistent heavy-tailed order with no self-reinforcing feedback and no imposed gradient. We report none; the absence of a counterexample is of course not a proof, and the theorem about the model holds regardless.

One apparent counterexample is worth disposing of at once, because it is our own and a reader will meet it two sections later. Equation (15) describes a population at $g = 0$ whose entities differ in per-step drift, and such a population plainly does develop a persistent order: the drift predicts the final ranking, with no feedback compounding anything. That is not the empty cell, and the taxonomy above already places it. A heterogeneous drift is an *imposed* gradient in the sense used here, the persistent trait a_i of Eq. (6) being fixed at the input and merely integrated by the dynamics, so the order it produces is conserved rather than self-generated; and being a sum of independent increments it stays light-tailed, for any light-tailed

distribution of the drifts, by a convolution argument akin to Proposition 3(iii), so it is not heavy-tailed either. It fails the empty cell's conjunction on two of its three terms. The distinction matters beyond bookkeeping: precisely because heterogeneous drift mimics amplification on the persistence signature while differing from it on tail shape and on self-generation, it is the null that Sections 8 and 11 must and do control for, by contrast rather than by magnitude.

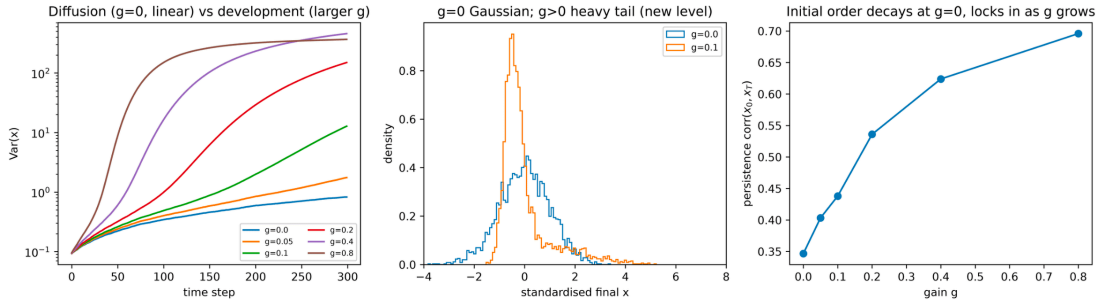


Figure 4. In the model, amplification is necessary for development. *Left:* with distinct initial attributes, $g = 0$ grows the cross-sectional variance only linearly (diffusion), while sufficient gain grows it super-linearly; as in Figure 1, the smallest gains have not separated from diffusion within this horizon. *Middle:* the final distribution at $g = 0$ (excess kurtosis 0.03: Gaussian, no new level) against $g = 0.1$ (excess kurtosis 4.2: a heavy right shoulder, structure). The comparison is drawn at moderate gain deliberately, since excess kurtosis is not monotone in g and turns negative at $g = 0.2$ (-1.1) as entities pile up at the ceiling, the same saturation that produces the downturn in Figure 1. *Right:* the persistence of the initial ordering $\text{corr}(x_0, x_T)$ decays at $g = 0$, to 0.35 by $T = 300$ and to zero only as $t \rightarrow \infty$ (Proposition 3ii), and locks in as g grows. Without the amplifier the system diffuses; only with it does it develop.

6. The tail index in closed form: contestability sets the exponent

The previous section leaves the heavy tail qualitative. It need not be. In the saturated regime the feedback is flat, $f(S) \rightarrow \theta S^*$, so a surviving entity's status is multiplied each step by $M = \exp(c + \eta)$ with

$$c = \ln(1 + g\theta S^*), \eta \sim \mathcal{N}(0, \sigma^2), \quad (8)$$

while the reverse-dominance reset returns the entity to the floor. That reset is *dominance-triggered* rather than constant-rate, and since the distinction carries the argument below we fix it here: an entity at status S is toppled in a given step with probability

$$h(S) = \frac{p_{hi}}{1 + (S_{top}/S)^{1/w}}, \quad (9)$$

a logistic in $\ln S$ about a dominance threshold S_{top} of width w : negligible for $S \ll S_{top}$, and saturating at the over-dominant toppling rate p_{hi} for $S \gg S_{top}$. Throughout we use $S_{top} = 200$, $w = 0.5$ and, in the illustrative calibration, $p_{hi} = 0.01$; of these it is p_{hi} that is independently measurable in the field, as a rate at which incumbents are actually deposed, which is what makes the prediction below falsifiable. Only that asymptote enters the derivation below. With it, the dynamics are a multiplicative random walk with geometric killing and renewal, the mechanism of [21] and, in the form used here, of [20][17][18][19]. Its stationary law has a Pareto tail $\Pr(S > s) \sim s^{-\mu}$ with μ the positive root of the Cramér condition $(1 - p_{hi})\mathbb{E}[M^\mu] = 1$, that is

$$\mu = \frac{-c + \sqrt{c^2 - 2\sigma^2 \ln(1 - p_{hi})}}{\sigma^2} \xrightarrow{\mu\sigma^2 \ll c} \frac{-\ln(1 - p_{hi})}{c}. \quad (10)$$

The limiting form is the content: *the tail index is the ratio of the toppling rate to the per-step amplification c of Eq. (8).* The smallness condition is on $\mu\sigma^2$ rather than on σ^2 alone, since expanding the root leaves a correction of order $\sigma^2 \ln(1 - p_{hi})/c^2$; equivalently the limit needs $2\sigma^2 |\ln(1 - p_{hi})| \ll c^2$, so it degrades in the thin-tailed corner where toppling is frequent and amplification weak, and it is the exact root that is used throughout below. A status that is contested often relative to how fast it compounds has a thin tail; one that compounds faster than it is revoked has a heavy one (Figure 5, right).

One boundary of this result should be stated, because it limits how far it travels. The derivation lives in the *saturated* regime and therefore presupposes a per-unit ceiling S^* : it describes statuses that must be maintained, in the sense of Proposition 1. Within that class the toppling rate is a genuine control parameter, and the free-token/toppling contrast of Section 8 acquires a distributional face alongside the early-lead one: as $p_{hi} \rightarrow 0$ the index $\mu \rightarrow 0$ and the stationary law disappears, which is the runaway a revocation-free status exhibits. But a truly ceiling-free token (a download or follower count, $\varrho = 0$ in Proposition 1) is not the $p_{hi} \rightarrow 0$ limit of (10); it lies outside the derivation altogether, since with $f(S) = \theta S$ unbounded the per-step multiplier is not the constant of Eq. (8) and the growth is super-exponential rather than geometric. We therefore claim the index for contestable, maintained statuses and not for free tokens, the same scope restriction the non-monotone gain carries (Section 10).

Equation (10) is a prediction about the *full* nonlinear dynamics, so we test it on those (`kesten_tail.py`). Running Eq. (3) with the Gould feedback and the logistic reverse-dominance hazard to stationarity and applying a Hill estimator [25] to the simulated tail, over a grid of three gains and six toppling rates, (10) is confirmed wherever the tail is heavy enough to be resolved (Figure 5, left): across the twelve grid cells with $\mu \gtrsim 2$ the mean absolute relative error is 3.1% and the worst is 8.9% (for instance

$\mu = 0.576$ predicted against 0.572 measured at $g = 0.2, p_{hi} = 0.1$, and $\mu = 1.510$ against 1.511 at $g = 0.4, p_{hi} = 0.4$). At larger μ the Hill estimate falls increasingly short of the prediction, by up to 60% at $\mu \approx 6.7$. We read this as estimator bias rather than model failure, and the diagnostic is internal: moving the Hill threshold further into the tail, from the top 2% to the top 0.5%, moves the estimate *towards* the prediction in every such cell, which is the signature of a threshold not yet in the asymptotic regime. We therefore claim (10) for $\mu \lesssim 2$, the heavy-tailed regime the paper is about, and leave the thin-tail corner untested.

The illustrative calibration has no fitted power law, and we do not report one

Substituting the calibration used throughout ($\sigma = 0.05$, $\theta = 0.02$, $S^* = 50$, $p_{hi} = 0.01$) into (10) gives $\mu = 0.105, 0.055, 0.030$ at $g = 0.1, 0.2, 0.4$. An index this far below 1 means the stationary law has no mean, and, more practically, that no simulation of any feasible length comes near it: the heavy right shoulder of Figure 4 is a pre-asymptotic transient of a distribution still spreading, not a realised power law. This is why we report the shoulder and the excess kurtosis rather than a fitted exponent there, and it sharpens the claim of Section 5: at the calibration where the amplifier manufactures the most dispersion, the tail it is heading towards is one without a mean.

Why this does not contradict the necessity theorem

A reader may substitute $g = 0$, hence $c = 0$, into (10) and obtain a finite $\mu = \sqrt{-2\ln(1 - p_{hi})}/\sigma \approx 2.84$, apparently a power-law tail with no amplification at all, against Section 5's claim that a heavy tail requires the reset *and* the drift together. The two are consistent, and the reason locates a real modelling assumption. Equation (10) is derived in the *saturated* regime, where the hazard has reached its asymptote p_{hi} ; substituting $c = 0$ into it silently keeps that asymptotic hazard while removing the drift that was the only thing carrying entities up to it. Our reset is not a constant-rate killing but the *dominance-triggered* hazard (9): a coalition topples the over-dominant, not the typical. At $g = 0$ the population never becomes over-dominant. Over $T = 300$ steps at $\sigma = 0.05$ the most extreme of 4000 walkers reaches only $S \approx 18$ against $S_{top} = 200$, where the hazard is $\sim 8 \times 10^{-5}$, two orders below p_{hi} , so renewals essentially never fire and the process is the plain martingale of Proposition 3. The distinction is the modelling content here: a Kesten process with *state-independent* killing does produce a power law without any drift, as (10) at $c = 0$ correctly reports for that different model. What the necessity theorem needs, and what reverse dominance supplies, is that revocation be triggered by dominance rather than applied uniformly.

Where a system topples at a constant rate regardless of standing, the necessity claim does not hold, and that is a testable difference between the two kinds of renewal.

A falsifiable threshold

Setting $\mu = 1$ in (10) gives the contestability a status must have for its stationary distribution to possess a mean at all,

$$p_{hi}^* = 1 - \exp(-(c + \sigma^2/2)), \quad (11)$$

which at the paper's calibration is $p_{hi}^* = 0.092, 0.168, 0.287$ at $g = 0.1, 0.2, 0.4$. The prediction is sharp and refutable in either direction: a domain in which positions are revoked at a measurable rate should show a finite mean above this line and a divergent one below it, with the crossing located by the independently measurable c and p_{hi} rather than fitted.

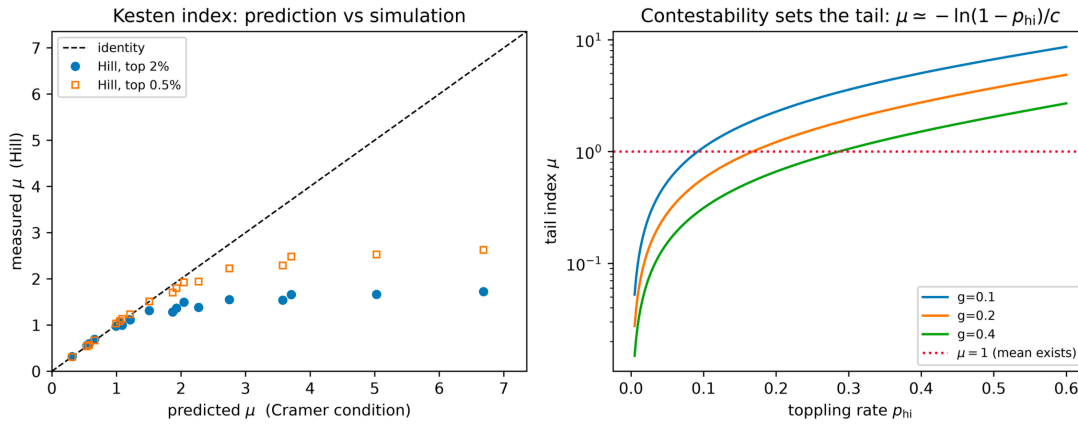


Figure 5. The tail index is available in closed form. *Left:* the Cramér root (10) against a Hill estimate from the simulated stationary law of the full nonlinear dynamics, over three gains and six toppling rates. Agreement is close for $\mu \lesssim 2$; at larger μ the Hill estimate is biased low, and moving the threshold from the top 2% to the top 0.5% (open squares) moves it back towards the prediction. *Right:* μ against the toppling rate, the limiting form $\mu \simeq -\ln(1 - p_{hi})/c$ making contestability the control parameter. The line $\mu = 1$ is Eq. (11); the paper's illustrative calibration ($p_{hi} = 0.01$) sits far below it.

7. The ensemble moves; the typical entity does not

The multiplicative dynamics (3) are non-ergodic, and this is where “the system advances” and “dead ends occur” become the same statement rather than opposed ones. We run $M = 5000$ independent single-

entity trajectories with a reverse-dominance reset (a coalition toppling the over-dominant [26], a Kesten renewal [21]) at $g = 0.2$, with the hazard (9), and compare two averages: the ensemble mean $\langle x \rangle$ (the population as a whole) and the median trajectory (the typical member).

Figure 6 shows the decomposition. The ensemble mean climbs to ≈ 6 and holds there; the median trajectory rises to ≈ 1.5 and then *declines* back towards ≈ 0.85 . The population moves, but its movement is carried by a thin heavy tail of winners, visible in the right panel, not by the typical entity, which peaks and falls back. Per step the ensemble advances about $7\times$ faster than the median. And $\sim 20\%$ of trajectories end at or below where they started: dead ends, in the literal sense that amplification took them nowhere or backward.

This is the quantitative form of the intuition that motivated the paper. The advantage of not stalling, that the population does not sit frozen at the symmetric state but explores, is real at the level of the ensemble. But it is bought with the dead ends: the same multiplicative structure that lets a few trajectories run into the tail sends a fifth of them into arrest or decline. The ensemble's motion and the typical member's dead end are not two phenomena; they are the ergodic and time averages of one non-ergodic process. These are individually catastrophic, high-variance events whose net cost cannot be settled: the gains (ensemble motion) are continuous and visible, the costs (individual dead ends) are distributed and hidden in the bulk of the histogram.

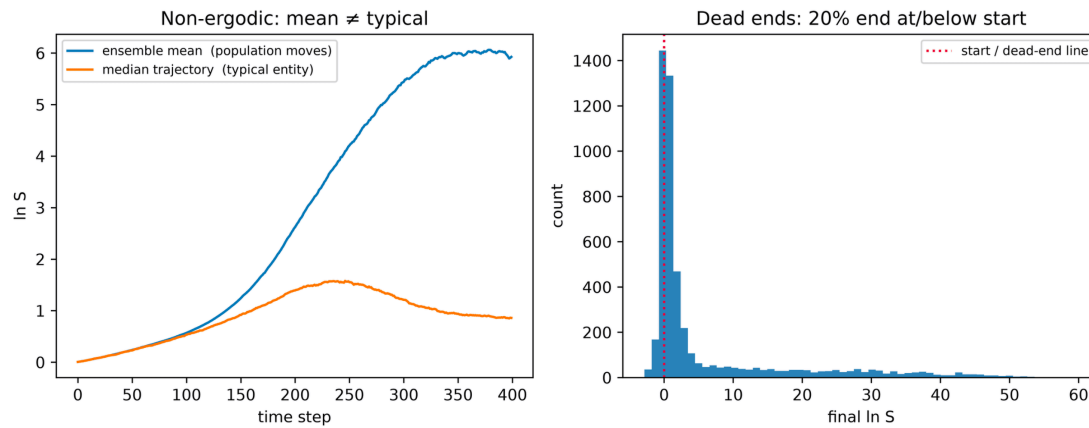


Figure 6. The motion is non-ergodic: the population advances, the typical entity does not. *Left:* the ensemble mean (blue) climbs steadily while the median trajectory (orange) peaks near 1.5 and declines; the advance is carried by a heavy tail of winners, not the typical member. *Right:* distribution of final log-status; $\sim 20\%$ of trajectories end at or below their start (dead ends). The population’s movement is bought with those dead ends; they are the ensemble and time averages of one process.

Robustness across the calibration

The point estimates above ($A \approx 400\times$, the $0.94 \rightarrow 0.66$ decoupling, the $\sim 20\%$ dead-end fraction) are one point in a family; only their *signs* are claimed robust. Table 1 sweeps the noise σ and feedback θ over a factor of five each. Across the grid the three directions never fail: the amplifier always manufactures dispersion ($A_{peak} \gg 1$, range 78–378), the decoupling is always present (E2 correlation drop > 0 , range 0.11–0.47), and dead ends always occur. But the magnitudes move, and one moves a lot: the dead-end fraction ranges from under 1% to nearly 40%, rising with noise and falling with feedback strength. So “a fifth of trajectories” is not a universal constant but the value at the illustrative calibration (the $\sim 20\%$ of Section 7); the robust statement is that the fraction is strictly positive and grows with the noise-to-feedback ratio. The scan reports a slightly lower 17% at the same (σ, θ) because it runs a common reduced protocol, described in the caption, not because the magnitude is unstable at fixed settings.

σ	θ	A_{peak}	E2 corr. drop	dead-end %
0.02	0.02	318	0.23	3.6
0.05	0.02	378	0.21	17.0
0.10	0.02	247	0.20	35.5
0.05	0.01	156	0.12	24.9
0.05	0.04	378	0.44	08.4

Table 1. Robustness of the three signatures across the calibration (`robustness_scan()`; representative rows of a 3×3 grid). The directions are invariant ($A_{peak} \gg 1$, correlation drop > 0 , dead ends > 0 throughout) while the magnitudes vary, the dead-end fraction most of all ($< 1\%$ to $\sim 40\%$ over the full grid). To keep the nine-cell sweep fast, the scan uses a common reduced protocol ($N = 3000$, $T = 300$ for the dispersion and dead-end columns, $N = 1500$, $T = 250$ for the E2 column, and the E2 drop measured to $g = 0.6$), so its magnitudes sit slightly below the headline point estimates of Sections 2, 4 and 7 (which use $N = 4000 - 5000$, $T = 300 - 400$, and the E2 drop to $g = 0.8$); only the signs are compared across the grid, not the exact values.

8. A capability-free signature, and the null that defeats it

Every test so far needs a measure of correctness the amplifier has not already contaminated: the true quality q of Section 4, the independent-condition quality of Design 1 (Section 10). That measure is a latent variable we can rarely observe cleanly, and it is the single hardest requirement in taking the model to data. This section isolates a signature that does not need it, and then shows why, used as it usually is, it settles nothing. The amplifier makes a prediction about the *internal geometry of a trajectory* alone, about how early the final order is fixed, that references capability nowhere and is therefore measurable wherever positions are tracked over time. That much is right. What does not follow, and is the substance of this section, is that observing the predicted geometry licenses the conclusion.

Ask: how much of the final ordering is already determined by the first τ -fraction of a trajectory? Define

$$\rho(\tau) \equiv \text{corr}(\text{rank}_x(\tau T), \text{rank}_x(T)), \quad (12)$$

the rank correlation between the ordering at an early time τT and the final ordering, on a population that starts near-homogeneous and into which *no* quality is injected. The shape of $\rho(\tau)$ separates diffusion from amplification, and the diffusive case has a clean closed form.

Proposition 4 (Early-lead persistence and the $\sqrt{\tau}$ law). *Under the dynamics (3) from a near-homogeneous start:*

(i) At $g = 0$ the increments are i.i.d. and independent of the past, so $x(T) = x(\tau T) + \sum_{s > \tau T} \eta_s$ with the tail sum independent of $x(\tau T)$, giving

$$\text{corr}(x(\tau T), x(T)) = \sqrt{\text{Var}(x(\tau T)) / \text{Var}(x(T))} = \sqrt{\tau}.$$

The first τ -fraction of history correlates only $\sqrt{\tau}$ with the final order, and explains only the fraction τ of its variance: the ordering after the first 1% correlates $\sqrt{0.01} = 0.1$ with the final one, and early leads are uninformative.

(ii) At $g > 0$ the per-step slope $\lambda = \varphi' > 1$ (Proposition 3) makes the final state dominated by its earliest amplified fluctuation: linearising, $x(T) \approx \sum_s \lambda^{T-1-s} \eta_s$, whose variance is dominated by the smallest s : the earliest amplified inputs win. Hence $\rho(\tau) \rightarrow 1$ for $\tau \ll 1$, and the excess $\rho(\tau) - \sqrt{\tau}$ is strictly positive. It is also non-decreasing in g while the feedback is unsaturated, in the sense that $\lambda = \varphi' > 1$ over the trajectory; once saturation binds, $\varphi' \rightarrow 1$ (with the feedback of Eq. (2), $f'(S) = \theta / (1 + S/S^*)^2 \rightarrow 0$) and ρ is in any case bounded by one, so no global monotonicity in g is claimed here, consistently with the non-monotone $A(g)$ of Section 2.

The statement mentions no quality q : it is a property of the map, not of any latent capability, and is measurable from the trajectory of positions alone.

The rank null is slightly below $\sqrt{\tau}$, and that is the whole of the residual

One technical point must be made, because without it the verification below looks marginally off. Part (i) derives $\sqrt{\tau}$ for the *Pearson* correlation of x , whereas the quantity we measure and report throughout, Eq. (12), is a *rank* correlation, which is preferable in data because it is invariant to how a position is scaled. At $g = 0$ the pair $(x(\tau T), x(T))$ is bivariate normal, so the two are related exactly by $\rho_S = (6/\pi) \arcsin(\tau/2)$ and the rank-scale null is

$$\rho_S(\tau) = \frac{6}{\pi} \arcsin\left(\frac{\sqrt{\tau}}{2}\right) < \sqrt{\tau}, \quad (13)$$

lying a little under the Pearson law, by about 0.013 at $\tau \approx 0.1$ and by at most about 0.018 near $\tau \approx 0.35$. This matters twice. It accounts entirely for the small, systematically negative residual of the simulated $g = 0$ curve against $\sqrt{\tau}$: that residual averages -0.0088 across the grid and is negative at almost every point, whereas against (13) it averages $+0.0007$ and changes sign, i.e. the bias vanishes and only sampling scatter remains. The diffusive baseline is thus verified more exactly than a comparison with $\sqrt{\tau}$ suggests. And where we use $\sqrt{\tau}$ as the null on real data, the substitution is conservative, since the true rank null is lower and clearing $\sqrt{\tau}$ is therefore the harder test. We keep $\sqrt{\tau}$ in the figures for readability and note the correction here.

Figure 7 confirms both regimes. The $g = 0$ curve sits very close to the diffusive law, and on it once the rank correction (13) is applied: measured 0.306, 0.618 at $\tau = 0.1, 0.4$ against a rank null of 0.303, 0.614 (and against $\sqrt{\tau} = 0.316, 0.632$). Amplification bows the curve sharply upward: at $g = 0.4$ the ordering after the first 1% of the horizon already correlates $\rho = 0.35$ with the final one (against 0.10 under diffusion), and after the first 10% it correlates $\rho \approx 0.90$. Writing $\tau_{90}(g)$ for the smallest early fraction whose ordering correlates ≥ 0.9 with the final one, diffusion needs the whole horizon ($\tau_{90} = 1$) while amplification collapses it to a sliver ($\tau_{90} \approx 0.2$ at both $g = 0.2$ and $g = 0.4$; the value of ρ at the first tenth, 0.898 at $g = 0.4$, sits just under the 0.9 mark, so τ_{90} falls at the next grid point). The order freezes early, and how early is set by the gain.

The reading is the same anti-meritocratic one, sharpened. Because the population is homogeneous, the early configuration $\rho(\tau)$ locks in is pure amplified noise: there is no capability to track. In that homogeneous setting, and there only, a large $\rho(\tau) - \sqrt{\tau}$ is positive evidence of multiplicative amplification (over pure diffusion) precisely *without* implying the frozen order is deserved; it measures how completely history is decided by its opening, not whether the opening was earned. The homogeneity is a premise, and where it fails the same excess arises at $g = 0$, as the next paragraph shows. (What it does *not* on its own establish is that the amplifier is *this* position-mechanism rather than plain cumulative advantage; that separation needs the revocation contrast below.) This is the cleanest form of the paper's claim: it needs neither a quality measure nor an assumption that the frozen order is meritless, only the trajectory itself.

The premise is stronger than “a homogeneous start”

The $\sqrt{\tau}$ null is weaker than it looks, and the point governs every empirical use of it below. Proposition 4(i) derives $\rho(\tau) = \sqrt{\tau}$ at $g = 0$ from a near-homogeneous start, but the premise it needs is

not merely that entities begin at a common position: the increments must also be identically distributed across them, so the *population* must be homogeneous. Either kind of persistent difference breaks it. A dispersed start gives, at $g = 0$,

$$\rho(\tau) = \sqrt{\frac{Varx_0 + \sigma^2\tau T}{Varx_0 + \sigma^2T}} > \sqrt{\tau} \quad \text{whenever } Varx_0 > 0, \quad (14)$$

and a common start with a heterogeneous per-step drift a_i , the case of any stock accumulated from zero by entities of differing quality, gives the related but distinct

$$\rho(\tau) = \frac{\tau T^2 Vara + \sigma^2 \tau T}{\sqrt{(\tau^2 T^2 Vara + \sigma^2 \tau T)(T^2 Vara + \sigma^2 T)}}, \quad (15)$$

which likewise exceeds $\sqrt{\tau}$ for any $Vara > 0$ and tends to 1 once the drift dominates the noise. Both say the same thing: a persistent trait predicts the final order on its own, with no feedback compounding anything. The simulation makes the point more sharply than a single number would: sweeping $Vara$ over four decades at $g = 0$ (`data_collapse.py`) drives τ_{90} from 0.79 down to 0.005, straddling the whole range the amplifier family occupies (0.07–0.26) and the preferential-attachment null with it. There is thus no characteristic τ_{90} for heterogeneity at all: it has a free parameter, and any observed lock-in speed can be matched by choosing it. An excess over $\sqrt{\tau}$ therefore fails to separate amplification not only from cumulative advantage but from *no amplification at all*. What survives is the contrast, which is why every empirical claim in this paper is built from one: heterogeneity inflates both arms of a within-domain comparison alike, so it can supply only a minority of a gap between them (the placebo of Section 11 calibrates how much), whereas the mechanism predicts exactly such a gap.

One discriminating caveat: reverse dominance erases the early lead

The lock-in is a property of the *runaway* amplifier. Adding the reverse-dominance reset of Section 7, the coalition that topples the over-dominant, returns τ_{90} towards 1 even at $g = 0.2$: active toppling periodically wipes the early leader, so no opening advantage survives to the horizon. The caveat is a second axis the signature resolves. A freely accumulating position with no per-unit ceiling and no toppling (downloads, followers, citation stock: the free tokens of Section 2) should show strong early-lead persistence; an actively contested status that is reciprocated and can be revoked should show *weaker* persistence at the same gain. This contrast, not the raw magnitude of $\rho(\tau)$, is what carries the identification: plain preferential attachment on a monotone stock produces early-lead persistence as strong as the amplifier's, or stronger, so a high $\rho(\tau)$ alone does not distinguish this mechanism from

cumulative advantage in general. Only the collapse of that persistence under revocation does. Early-lead persistence and its absence together separate the free-token amplifier from the toppling-limited one.

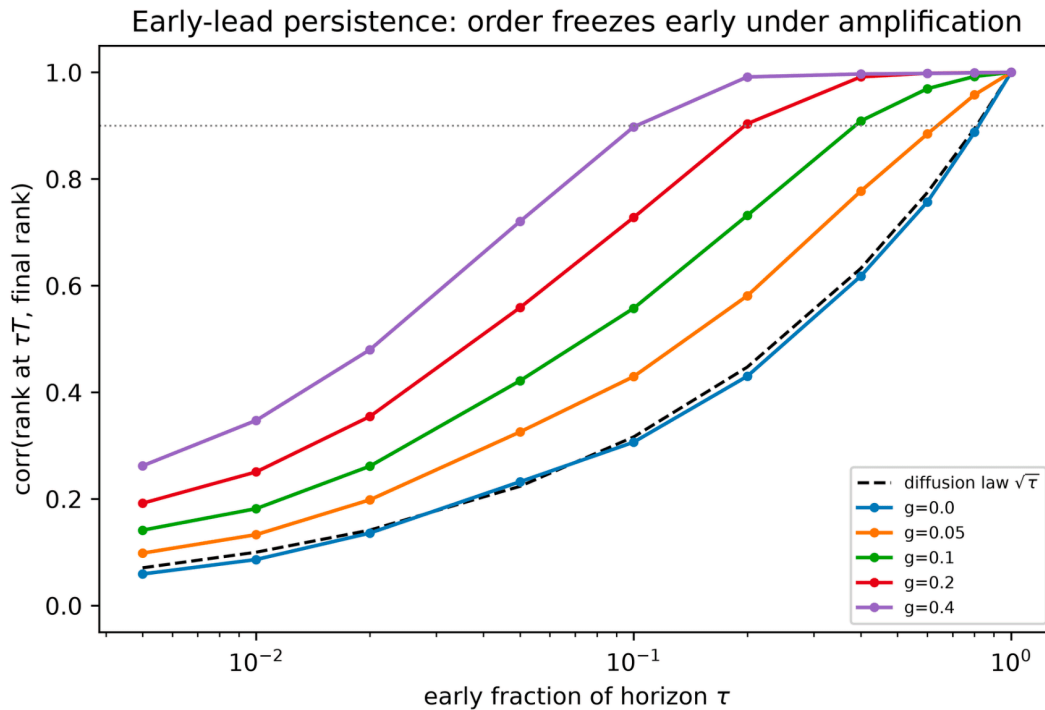


Figure 7. Early-lead persistence, a capability-free signature. The rank correlation $\rho(\tau)$ (Eq. (12)) between the ordering at an early fraction τ of the horizon and the final ordering, on a homogeneous population with no quality injected. Diffusion ($g = 0$, blue) lands on the $\sqrt{\tau}$ law (dashed, Proposition 4i): early leads are uninformative. The blue curve sits a little *under* the dashed line throughout, which is not sampling error but the rank correction (13): $\sqrt{\tau}$ is the Pearson-scale null, and the exact rank-scale null $(6/\pi)\arcsin(\sqrt{\tau}/2)$ lies just below it, where the measured curve in fact sits. Amplification bows the curve up: at $g = 0.4$ the ordering after the first 10% of the horizon already correlates $\rho = 0.898$ with the final one. Here the excess over $\sqrt{\tau}$ tracks amplification without any measurement of capability, but only because this population is homogeneous by construction: on heterogeneous entities the same excess is produced at $g = 0$ (Eqs. (14)–(15)), so in data it is the contrast, not the excess, that identifies the mechanism.

9. Interpretation: instability, not teleology

The preceding sections are properties of the model; this one is interpretation, offered as such. The dispersion, development and ergodicity results (Sections 2, 5 and 7) invite a functionalist reading, that

hierarchy exists because it lets systems move and moving is good. We resist it. The symmetric state is an unstable fixed point ($\sigma > 0$, $f'(0) > 0$), so the population leaves it by dynamics, not purpose; and calling not-stalling a benefit would need a metric on destinations, which is what the amplifier destroys, having eroded the capability signal such an account requires (Section 4). By Proposition 2 the mechanism keeps no signature of whether it amplified a real advantage or pure noise, so “forward” is a label applied afterwards to whatever persisted. The positive content is correspondingly modest: the amplifier manufactures the dispersion on which any later selection can act, in a direction uninformed by correctness, at the price of dead ends in a non-ergodic ensemble. That is a claim about the generation of variance, not about the value of its destinations.

10. Empirical predictions and two identifying designs

The right question is not “which dataset represents a hierarchy?”, since the mechanism is buried under others in every real system and no clean dataset exists, but “which empirical signatures must the mechanism leave behind, and in what kind of system?” Datasets are chosen by *mechanism*, not by field: any system where a visible position signal (stars, views, reputation, followers, capital) reinforces itself multiplicatively from a common starting position is a candidate. We advance two designs, Design 1 identifying the decoupling causally and Design 2 needing no capability measure at all, and report a first real-data pass on both (Section 11).

Design 1: the decoupling is causal (parallel worlds)

The Salganik–Dodds–Watts artificial cultural market ^[1] runs the same songs through many independent “worlds,” so between-world differences cannot come from song quality, which is held fixed, but only from amplified early differences (the social-signal draw). This is Section 4 in a laboratory, with two levers it already contains: an *independent* condition (no visible counts) that measures intrinsic quality Q_i uncontaminated by the amplifier, and a social-signal strength that is the gain g . The prediction is that the rank correlation between market success and independent quality *falls monotonically with g* : per world $\rho_w = \text{corr}(r^{k_w}, Q)$ regressed as $\rho_w = \beta_0 + \beta_1 g_w$ with $\beta_1 < 0$. In simulation the correlation falls from 0.76 to 0.65 over $g \in [0, 4]$ (`power_analysis.py`; the dose–response sweep of Figure 8, a separate seeded run of the same market model, reads $0.75 \rightarrow 0.63$); a Monte–Carlo power analysis gives power ≥ 0.96 already at ten worlds per gain level, so five gain levels at twenty worlds each (~ 100 worlds) is a

comfortable margin. A within-world shuffle placebo must return $\beta_1 \rightarrow 0$; the design fails the model if $\beta_1 \geq 0$.

Design 2: early-lead persistence (no capability measure needed)

Section 8 predicts that on a near-homogeneous entry cohort the rank correlation $\rho(\tau)$ between the ordering at an early fraction τ of the horizon and the final ordering is high for small τ , so a sliver of early history ($\tau_{90} \ll 1$) already fixes most of the final order. Because it references no quality, it is testable wherever a positional signal is tracked over time and entrants share a starting position: do a repository's first weeks of stars, a video's first hours of views, a startup's seed round predict the final ranking? The subtlety is the null, and here a synthetic check disciplines the claim. On a *monotone* accumulating stock the analytic $\sqrt{\tau}$ diffusion law is *not* a sufficient control: simulating plain preferential attachment, which carries none of this model's noise-lock-in or ceiling, gives an early-lead persistence that not only exceeds $\sqrt{\tau}$ but *exceeds the amplifier's own* ($\tau_{90} \approx 0.1$ for preferential attachment against 0.375 for the free-token amplifier at $g = 0.2$, the ceiling-free parameterisation of this synthetic check, slightly slower to lock in than the Gould-saturated simulation of Section 8 at $\tau_{90} \approx 0.2$; both against $\tau_{90} = 1$ for diffusion). So the *magnitude* of $\rho(\tau)$ cannot separate the amplifier from cumulative advantage in general: an excess over $\sqrt{\tau}$, or even over a preferential-attachment null, is not diagnostic of *this* mechanism. What is diagnostic is the free-token/toppling contrast of Section 8: the *same* amplifier run with a reverse-dominance renewal (a revocable, actively contested status) collapses its persistence back to the diffusion regime ($\tau_{90} : 0.375 \rightarrow 1$ in simulation), whereas a freely accumulating position with no per-unit ceiling and no toppling (download, star, follower stock) stays high. Revocation is something plain cumulative advantage does not contain, so its signature, strong persistence in free-token domains and weak in contestable ones, is the cleanest separation of this mechanism from its rivals. The price is that it identifies amplification without, on its own, certifying that the frozen order is undeserved.

The identifying signature is a joint dose–response

What separates the amplifier from its rivals is not any single curve but that *several* signatures move together with the *one* gain knob. Figure 8 sweeps the market model across gain and shows the decoupling $\text{corr}(\text{success}, \text{quality})$ falling while the concentration (Gini) and the early-lead lock-in $\rho(0.1)$ both rise, a joint dose–response no rival reproduces (meritocracy predicts no gain effect at all; plain cumulative advantage has no single tunable gain that moves all three at once). The two solid Music Lab points, weak and strong social signal (0.765 and 0.651), sit on the predicted decoupling curve, subject

to the presentation confound recorded in Section 11: the two arms differ in how songs were displayed as well as in signal strength, so they anchor the curve's *direction* rather than calibrating its abscissa. This is the simulated dose–response with a two-point, partly confounded real anchor. Tracing the full multi-level curve, whether by the pre-registered high-power market above or by an exogenous gain change such as the platform experiments of Section 11, is the test that would confirm the joint response rather than one of its projections.

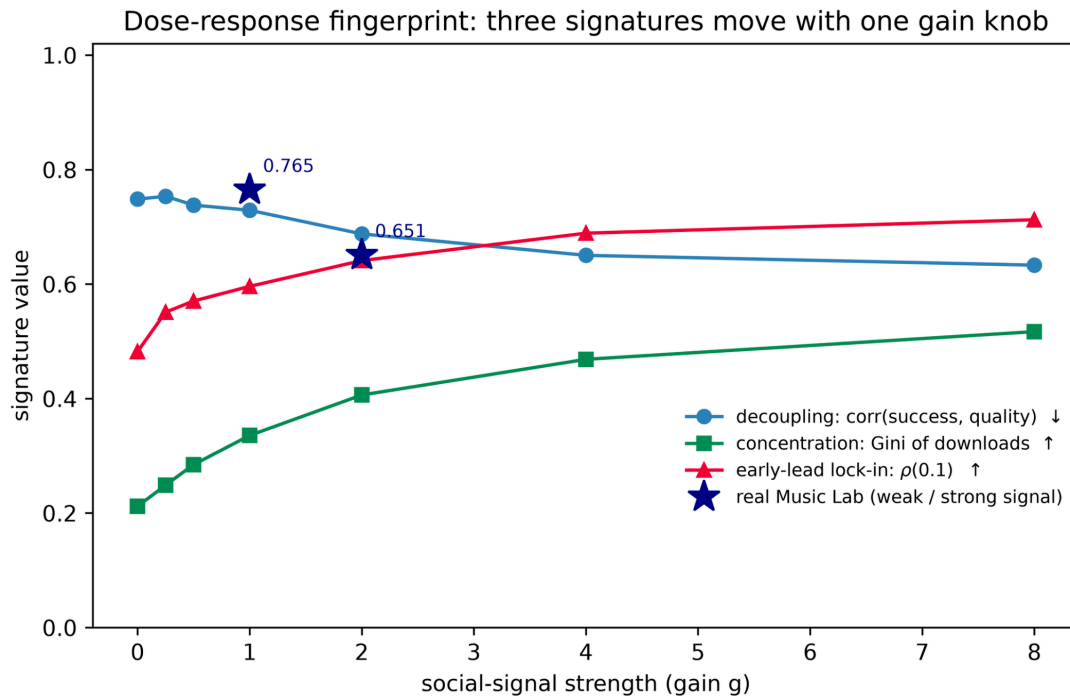


Figure 8. The identifying dose–response: three signatures move with one gain knob. As the social-signal strength (gain) rises in the market model, decoupling from quality falls (blue, 0.75 → 0.63) while concentration (green, 0.21 → 0.52) and early-lead lock-in (red, 0.48 → 0.71) rise together. The two latter are monotone throughout; the decoupling is flat within noise over the lowest gains (0.748 against 0.753 at $g = 0$ and 0.25) before falling steadily. The real Music Lab points (navy stars, weak/strong signal: 0.765 → 0.651) reproduce the predicted *direction*. Their placement on the abscissa, at $g = 1$ and $g = 2$, is nominal and carries no claim: the two experiments differ in presentation as well as in signal strength (Section 11), so they cannot calibrate a gain axis, only bracket a slope. Rivals do not produce this joint response to a single knob; a full multi-level sweep is the test (`dose_response.py`).

A prediction we withdraw for free-token systems

The non-monotone $A(g)$ with a dispersion-maximising gain (Section 2) does *not* transfer to every amplifying setting. The downturn requires a *per-unit* ceiling (Gould's θS^*), so that at high gain every entity converges to the same saturation and dispersion collapses. A free token with no per-unit cap (downloads, stars, followers) has no S^* , so concentration only grows: in a simulated download market the Gini rises monotonically ($0.20 \rightarrow 0.52$ across $g \in [0, 8]$; `power_analysis.py`; the separate sweep of Figure 8 reads $0.21 \rightarrow 0.52$), no downturn. The non-monotonicity is thus a property of systems with individual saturation, and should be tested only there.

A sharper universality test: does $\rho(\tau)$ collapse?

The strongest cross-domain claim, that these systems are one normal form rather than an analogy (Section 1), makes a prediction beyond “each curve bows above $\sqrt{\tau}$.” Rescaling each system's horizon by its own τ_{90} , the early-lead curves $\rho(\tau/\tau_{90})$ should *collapse* onto a single master shape, as correlation-length rescaling collapses curves at a critical point; a different mechanism should fall off it (Figure 9). In simulation this is clean: an amplifier family spanning gains, noise levels, and ceiling-versus-free-token parameterisations (raw τ_{90} ranging 0.07–0.26; a different parameterisation from the Design-2 synthetic check, whose free-token τ_{90} is 0.375) collapses to an internal RMS spread of ≈ 0.01 , while diffusion and plain preferential attachment sit $6\text{--}15\times$ that spread off the master, on every seed (`data_collapse.py`). The same holds against the heterogeneity null that defeats the magnitude test: a $g = 0$ population whose entities merely differ in drift (Eq. (15)) does not collapse at all, its own internal spread reaching ≈ 0.13 across four decades of V_{ara} , and its mean curve sits $10\times$ the amplifier's internal spread off the master, alongside diffusion at $7\times$ and preferential attachment at $12\times$. Persistent heterogeneity, unlike amplification, has no single rescaled shape, since each V_{ara} traces a different one, so shape separates in simulation what magnitude cannot. The rescaling is also what makes the otherwise incommensurable real axes comparable at all (years of cumulative startup capital, a months-long star stock, a days-long vote sequence), a comparison Section 11 declines to make on raw τ_{90} .

We ran the test on *three independent* real free-token domains (`real_data_collapse.py`): startup funding (cumulative capital of a 2010-entry cohort, 2325 companies from the open Crunchbase export), GitHub stars, and Wikipedia support-only counts, all re-extracted on a common fine 18-point grid (`refine_grids.py`). On real data the test is *inconclusive*. The three real curves do cluster after rescaling, with mutual RMS spread ≈ 0.07 , about three times closer to one another than the $\sqrt{\tau}$ diffusion null.

Little evidential force survives, however, once the clustering is examined. It is fragile to extraction: a coarse 10-point grid gave a visibly tighter collapse, and refining to the 18-point grid weakened it, which is the wrong direction for a real shape. The real domains sit $7.7\text{--}9.8\times$ the internal spread off the *simulated* master, *farther* than the $\sqrt{\tau}$ null itself at $4.1\times$; the master is calibration-dependent so this is not decisive, but a family the data sit farther from than they sit from the null cannot be said to describe them. And decisively, the heterogeneity null of Eq. (15) is *not* excluded: the real master sits 0.032 from that null's mean curve, nearer than the three domains sit to one another. The one rival the simulation separates cleanly is thus the rival the real data cannot rule out, so the observed clustering is equally well explained by three domains whose entities simply differ in quality. We therefore claim nothing from this test beyond the simulated half. What would decide it is a free-token and a revocable arm measured *inside one system*, the design Section 11 realises and this cross-domain comparison cannot.

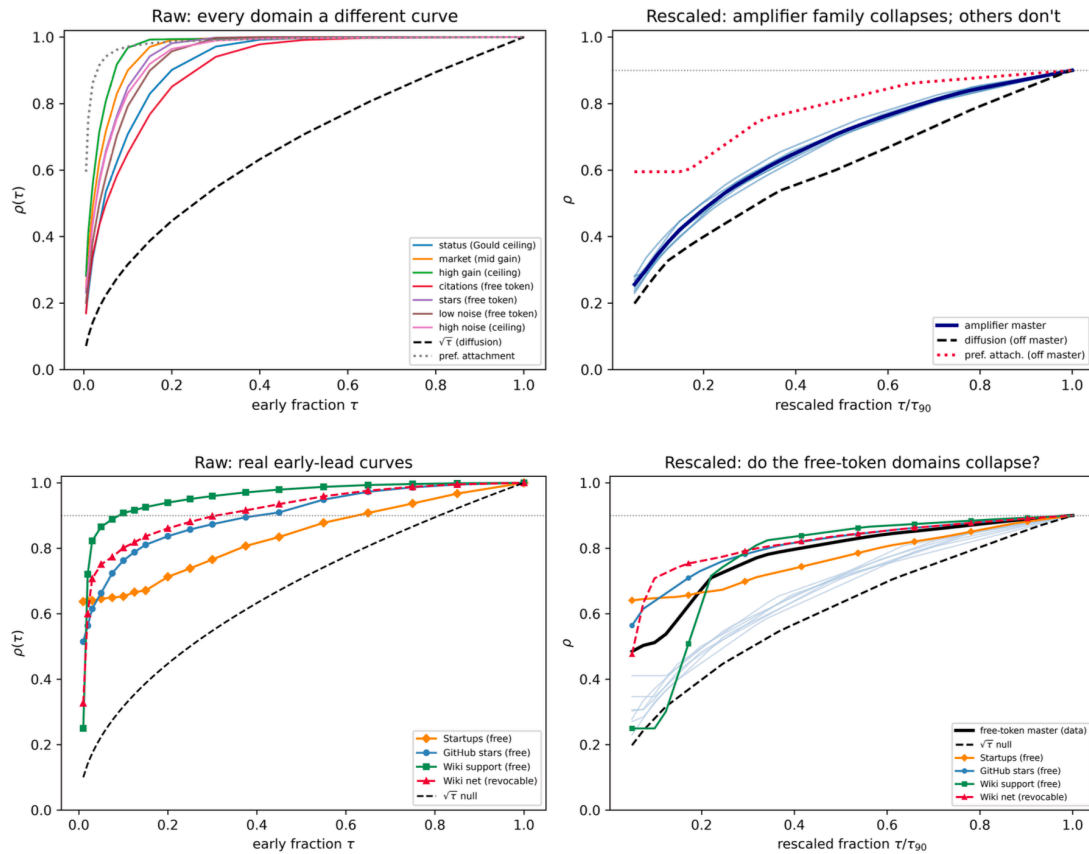


Figure 9. The $\rho(\tau)$ -collapse test, where the evidence is a *shape* rather than a magnitude. *Top, simulation:* rescaling each curve by its own τ_{90} collapses an amplifier family spanning gains, noise levels and ceiling-versus-free-token parameterisations onto one master (internal RMS spread ≈ 0.01), while diffusion (dashed) and preferential attachment (dotted) sit far off it on every seed. The heterogeneity null of Eq. (15) sits off it too and, unlike the amplifier family, does not collapse onto any shape of its own (numerical, not drawn). *Bottom, real data:* the same rescaling on three independent free-token domains (startup capital, GitHub stars, Wikipedia support-only). They cluster, but for the reasons given in the text the test is inconclusive, so we draw no conclusion from this panel; the identifying comparison is the within-system one of Figure 10.

What would falsify each contribution

Because a multiplicative model with saturation and reset is flexible, it is worth as much as its stated failure conditions, so we give one pre-statable observation per contribution. (i) The *necessity theorem* fails if anything is found in the empty cell of Section 5. (ii) The *non-monotone gain* fails if a system with a genuine per-unit ceiling shows manufactured dispersion rising monotonically in gain, with no downturn

at any horizon. (iii) The *tail index* fails if a maintained, revocable status whose toppling rate p_{hi} and per-step amplification c are separately measurable shows a stationary tail index away from the Cramér root (10), and in particular if a status sitting above the line (11) shows a divergent mean or one below it a finite one; the prediction is a number fixed in advance by two independently observable quantities, not a fitted exponent. (iv) The *decoupling* fails if market success tracks independent quality equally well at high and low gain ($\beta_1 \geq 0$ in Design 1); and the identification of the mechanism as *positional* amplification (rather than cumulative advantage in general) fails if early-lead persistence does *not* collapse under revocation where a free-token and a contestable arm are compared inside one system (Design 2). Each is a single observation stated in advance; none is met in the data we report, but that is a failure to refute, not a confirmation.

11. First real-data evidence

We report real-data results for both designs. We lead with Design 2, the capability-free early-lead test, because its identifying comparison is made *within a single dataset* and is therefore the cleanest evidence in the paper: a toppling arm and a free-token arm on Wikipedia adminship, followed by an independent free-token pole on GitHub stars. We then turn to Design 1 on the Music Lab, which is causally identified but has only two gain levels.

The capability-free toppling arm on real data: Wikipedia RfA

The first result tests Design 2 without any quality measure, and exploits the fact that the discriminating signal is the free-token/toppling contrast, not the magnitude of $\rho(\tau)$ (Section 8): on a monotone stock plain cumulative advantage already lifts $\rho(\tau)$ far above $\sqrt{\tau}$, so only the *collapse* of persistence under revocation identifies the mechanism. Wikipedia Requests for Adminship (the SNAP wiki-RfA corpus [27]; 2269 elections with ≥ 25 timestamped, signed votes) lets us measure both arms *inside one dataset*: the running *net* support (support minus oppose) is a revocable, contestable position an early front-runner can be toppled from, while the *support-only* cumulative count is the monotone free token. Ordering votes in time and reading $\rho(\tau) = \text{corr}(\text{rankposition at } \tau, \text{rankfinal})$ across elections, the free token locks in markedly earlier than the revocable position (Figure 10): $\tau_{90} = 0.10$ for support-only against 0.375 for net support, and at the first tenth of the vote sequence $\rho = 0.91$ versus 0.80. The gap carries statistical weight: resampling elections with replacement ($B = 2000$), the free-token-minus-revocable difference at the first tenth is $\Delta\rho(0.1) = 0.11$ (95% CI $[0.09, 0.13]$), it is positive with zero

excluded at every $\tau < 1$, and the τ_{90} difference is 0.275 in the full sample (bootstrap mean 0.245, 95% CI [0.20, 0.28]). The predicted ordering holds, free-token persistence $>$ revocable persistence, measured with no capability variable (`wiki_rfa_toppling.py`). That both arms also sit far above $\sqrt{\tau}$ carries no inference, candidates plainly differing in quality (Section 8); the design earns its conclusion from the *gap*, which heterogeneity can supply only in small part.

The two arms are the same votes of the same candidates read two ways, so a persistent difference between candidates enters both with nearly the same signal-to-noise: at an early fraction the support-only count has cross-candidate signal τn against within-candidate noise $\sqrt{\tau n \pi (1 - \pi)}$, while the net count has both quantities doubled, and the factor cancels. We quantify what survives that cancellation with an explicit placebo (`wiki_rfa_toppling.py`): each election keeps its real number of votes, candidates are given the empirical spread of final support shares, and votes are i.i.d. given the candidate, so heterogeneity is maximal and nothing compounds. The placebo returns a mean early gap of $+0.033$ and τ_{90} of 0.10 against 0.05, versus $+0.115$ and 0.375 against 0.10 in the data: heterogeneity buys just under a third of the observed gap in ρ and about a fifth of the observed τ_{90} separation. The residual is not zero, because ρ is a rank correlation and each arm is referred to its own final ordering, so we report the placebo as a calibration rather than as a proof of zero. What the real contrast has to beat, it beats by a factor of three or more, and what distinguishes support-only from net support is precisely that one is revocable. The effect is real but *milder* than the runaway-versus-toppling gap of the clean simulation ($\tau_{90} : 0.375 \rightarrow 1$), for a substantive reason: in adminship elections the opposition is rarely a majority, so net support seldom falls far and the toppling is only partial. The contrast is in the predicted direction; its size is a property of how contestable the particular status is.

The free-token pole: GitHub stars

The matching free-token arm is a positional signal with fine-grained entry times, a common starting position (every repository begins at zero stars, though not, of course, at a common quality), and no revocation. On a cohort of 295 GitHub repositories created in the first quarter of 2019 (final stars 80–3000) tracked over a 24-month horizon, with star timestamps read through the GitHub GraphQL API and cumulative star stock as the position (`github_earlylead.py`), the early-lead persistence is strong: $\rho(\tau)$ reaches 0.68 by the first tenth of the horizon and 0.89 by $\tau = 0.375$ ($\tau_{90} = 0.5$), well above the $\sqrt{\tau}$ diffusion null (0.32 at $\tau = 0.1$), though clearing that null carries no inference on its own (Section 8). Free tokens do lock in early. But the result also illustrates, on real data, why magnitude is not the test: the

observed persistence sits *below* a preferential-attachment null fitted to the same cohort ($\tau_{90} = 0.02$ for that null), so GitHub star accrual is if anything *less* early-frozen than pure cumulative advantage, and its raw persistence does not on its own single out the amplifier. It is also why we do not read a between-domain GitHub-versus-Wikipedia comparison as the contrast: their positions and time axes are not commensurable (a 24-month star stock against a fraction of a days-long vote sequence), and the raw τ_{90} ordering across the two is not even in the predicted direction. The identification rests on the *controlled* within-Wikipedia contrast above, where revocation is toggled inside one dataset; GitHub supplies the free-token pole and shows strong early-lead lock-in, but a free-versus-revocable comparison *within a single* free-token-plus-toppling system remains the cleanest test to run next.

Design 1 on the Music Lab: the decoupling is causal

This turns the decoupling from a correlation into a *causal* claim: Design 1 on the archived Salganik–Dodds–Watts artificial cultural market [11][28], a direct rerun of a published experiment. It ran the same 48 songs through many independent worlds, with an independent condition (no visible counts) that measures each song’s intrinsic quality Q and two experiments differing in social-signal strength: experiment 1 a weak signal, experiment 2 a stronger one. Because the songs are identical across worlds, between-world differences *cannot* come from quality; they are the amplifier. On the released download counts, the rank correlation between market success and independent quality falls from 0.765 under the weak signal to 0.651 under the strong one (eight worlds each; Welch $t = 4.19$, $df = 7.9$, one-sided $p = 1.5 \times 10^{-3}$), while concentration rises with the signal (Gini $0.34 \rightarrow 0.50$). This is Section 4’s decoupling, measured where quality is held fixed by construction, and it lands almost exactly on the simulated Design 1 figure ($0.76 \rightarrow 0.65$). Two features of this design deserve emphasis, because they are what the chess example of Section 12 lacks. Identical songs across worlds mean the heterogeneity that defeats the trajectory signatures there cannot operate here: with quality held fixed, a between-world difference has nowhere to come from but the amplifier. And the measured drop is, if anything, *conservative*. Each experiment estimates Q from its own independent world, and experiment 2’s is the larger of the two (2192 against 1578 downloads), so its Q is the *less* attenuated and its correlation should on that count be the higher. It is the lower. Resampling experiment 2’s independent world down to experiment 1’s precision drops its correlation further, to 0.573 ± 0.038 , widening the gap from 0.114 to about 0.19. The gain effect is thus large enough to overcome a measurement advantage working against it.

The comparison is bounded in two ways. Power is the lesser one: with only two levels and eight worlds each the test is under-powered, so the pre-registered high-power version remains a multi-level gain sweep. The other is sharper. The two experiments do not differ in social-signal *strength* alone, with everything else held fixed, as a clean dose of gain would require. They differ in how the songs were presented: experiment 1 showed them in a grid in random order with their download counts attached, while experiment 2 showed them in a single column *ranked* by current downloads. Ranking raises the salience of the social signal, which is why it is the stronger-signal arm, but it also introduces an ordinary position effect, items near the top of a list being chosen more often whatever the list means. Between-world differences remain causally attributable to the social signal, since quality is held fixed by construction and the presentation format is common to all worlds within an experiment; what is not clean is the reading of the two arms as two settings of the single knob g . The comparison is therefore good evidence that amplification decouples success from quality, which is the claim Section 4 makes, and weaker evidence for the dose–response profile of Figure 8, whose two real anchors inherit this confound. Tracing the curve within one presentation format, by varying signal strength alone, is the version we flag as needed and do not have.

Corroborating randomised evidence, and the prediction it leaves open

The mechanism’s core claim, that an *arbitrary* early advantage compounds into final position independently of merit, is not ours to establish from scratch: it has been shown causally, by randomised field experiments, more cleanly than any observational trajectory can. Muchnik, Aral and Taylor ^[2] planted a single random up-vote on comments and it lifted the final score by $\sim 25\%$; van de Rijt and colleagues ^[3] randomly conferred initial success (donations, ratings, endorsements) across four separate domains and watched it self-reinforce; and the Music Lab itself, run with the download counts *inverted*, manufactured self-fulfilling winners from the falsified signal ^[29]. These establish the amplifier’s engine (early position begets position, correctness notwithstanding) at population scale and with random assignment. What none of them varies is the *amount* of amplification: they fix the gain at one setting, whereas the model’s distinctive claim is a *dose–response*, every signature growing monotonically with gain and then *breaking* at the ceiling (Section 2), a profile neither meritocracy (no gain effect) nor plain cumulative advantage (no per-unit ceiling) produces. A setting where amplification is exogenously dialled would trace that curve; platform changes to visible counts are the natural instances (Instagram’s hidden like counts, YouTube’s 2021 removal of the public dislike count, which converts a partly

contestable position into a purer free token and so should *raise* early-lead persistence and *lower* outcome–quality coupling). Tracing that curve is the priority we flag.

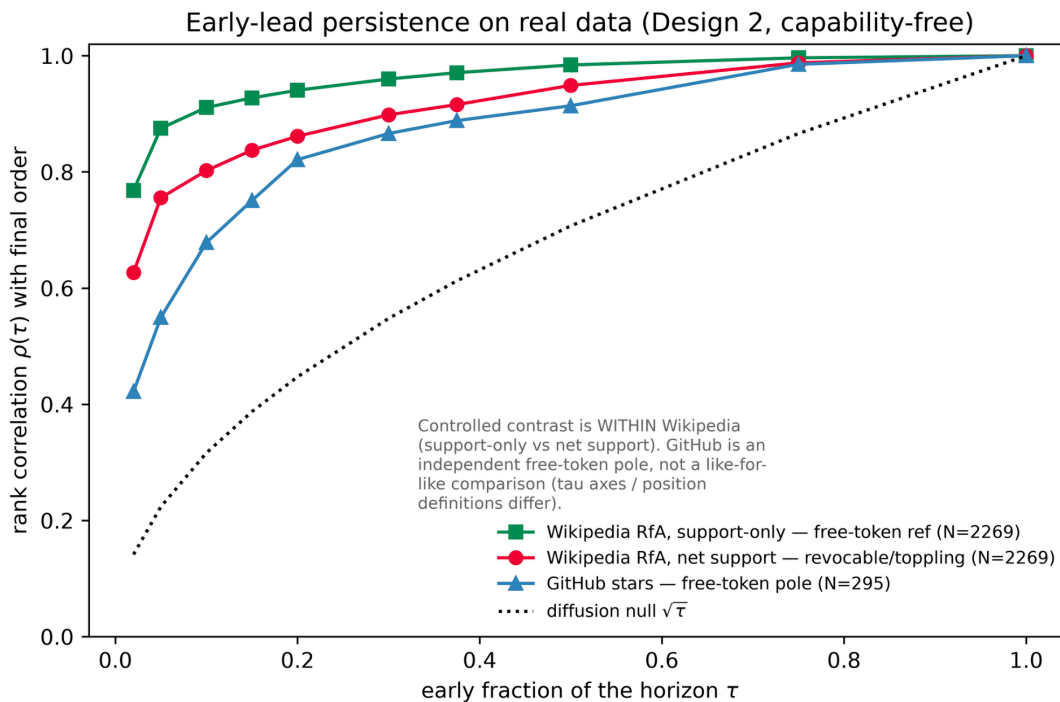


Figure 10. Early-lead persistence on real data (Design 2, capability-free): $\rho(\tau)$ between the ordering at an early fraction τ and the final ordering. All three real curves sit far above the $\sqrt{\tau}$ diffusion null (dotted), so early leads are informative everywhere, but that alone is only cumulative advantage. The identifying signal is the *controlled* within-Wikipedia contrast: support-only (green, the free token) locks in earlier than net support (red, revocable), $\tau_{90} = 0.10$ versus 0.375 . GitHub stars (blue) are an independent free-token pole; they show strong lock-in but are *not* a like-for-like comparison against Wikipedia (incommensurable τ axes and position definitions), and their magnitude even sits below a preferential-attachment null fitted to that cohort, so magnitude is not the test, the revocation contrast is.

12. The boundary at work: an order that looks manufactured and is not

This section is the paper’s central exhibit, and it is one the paper loses. If the identification boundary of Section 3 were merely a technical caveat, one could note it and proceed to read trajectories as before. The case below shows what that costs. A domain that wears *every* appearance the amplifier reading predicts,

and that a trajectory-only analysis would therefore classify as manufactured order, turns out to be mostly the opposite once an exogenous channel to capability is admitted. The trajectory is not silent throughout: as the dispersion diagnostic below shows, it does warn that the near-equal start was an artefact of measurement, so the dynamics reading (A) was never entitled. But between manufacture and revelation (B) it says nothing at all, and that is the half no amount of data repairs. The boundary is not a caveat; it is the difference between a right and a wrong answer.

Online chess supplies all three ingredients in one public dataset: a start that *looks* homogeneous, a plausible amplifier, *and* an exogenous channel to capability. The first two turn out to be less than they appear, which is the lesson; the third is what makes the lesson visible. The Lichess open database ^[30] holds billions of rated games under a CC0 licence. It gives every entrant a provisional rating near a common seed. That rating is adjusted and revocable game by game. It also gives a *convergent* skill estimate months later, once hundreds of further games have averaged the early fluctuation away.

Design

From the 2013-01 database we form an entry cohort of active players (at least 30 games that month) and read each player's log-rating trajectory. We contrast two regimes by entry rating. The *near-equal* cohort has entry Elo 1480–1520 (entry s.d. 5.3); the *heterogeneous* control has entry Elo 1000–2200 (s.d. 172). Within the month we measure the order that forms and its early-lead persistence. For the exogenous $\hat{k}$ of Eq. (7) we take each player's mean rating in a *later* month (+6 months, 2013-07, and +12 months, 2013-12), restricted to players with at least 20 games there so the estimate has converged.

What the band selects, and what it does not

One point of interpretation must be fixed before the results, because the cohort invites a misreading we want to rule out: it is *not* an instance of the small- $Var(\ln k)$ regime of Section 5. What the band makes small is the dispersion of the entry *estimate*, not of capability. Lichess seeds every entrant near a common provisional rating precisely because it does not yet know the player, so a 40-point band selects players the system has not yet told apart. The data say so directly. The near-equal cohort's month-end s.d. is 213 against the control's 211: the two cohorts have essentially the *same* underlying capability dispersion, and differ only in how much of it was visible at entry. Two consequences follow, and both matter for how the numbers below should be read. First, restricting entry to a narrow band leaves mostly measurement noise inside it, so the near-zero entry correlations are in part a property of that selection rather than a

discovery about chess; the informative number is not the -0.06 but the 0.79 . Second, the example accordingly does not trace the dominance-ratio axis of Section 5, since $Var(\ln k)$ is not what varies across the two cohorts. It does something else, which is what we ask of it here: it isolates *small observed dispersion with a live exogenous channel*, the configuration in which A and B come apart most sharply, the entry position being uninformative, the trajectory looking manufactured, and only $\hat{k}$ able to say otherwise.

The order looks manufactured

In the near-equal cohort a persistent order forms out of a near-symmetric start, and it is decoupled from the entry position exactly as symmetry-breaking predicts. The rank correlation between entry Elo and month-end Elo is -0.05 in the near-equal cohort against $+0.60$ in the control. Reading each player's running net score (wins minus losses) at fractions of their own games, the early-lead persistence rises from 0.32 at the first fiftieth to 0.90 by the halfway point, an excess over the $\sqrt{\tau}$ diffusion law that is positive at every τ and peaks at $+0.28$ (Proposition 4; `lichess_worked_example.py`). Taken at face value this is signature A together with an order uncorrelated with where players began: read from the trajectory alone, the manufactured-from-noise story.

Why the A reading is unsafe here too

That reading does not survive the previous paragraph, and the failure is instructive enough to state, because it is not the one the paper has been warning about. Proposition 4 derives $\rho(\tau) = \sqrt{\tau}$ at $g = 0$ from a homogeneous *population*, and a common entry rating does not supply one: the band equalises the starting position, leaving $Var x_0$ small, while capability, and hence the per-step drift a_i , goes on differing exactly as the month-end s.d. of 213 shows. This cohort is therefore the heterogeneous-drift case of Eq. (15) rather than the dispersed-start case of Eq. (14), and that null exceeds $\sqrt{\tau}$ for any $Var a > 0$, because a persistent trait predicts the final order without any feedback compounding anything. Once the drift dominates the per-game noise, Eq. (15) returns an excess well above the $+0.28$ peak measured here, with the gain set to zero. The observed excess is therefore consistent with a stable skill difference and no amplification at all, and the same holds for the entry decoupling, which a 40-point band would produce whatever the dynamics. Neither of the cohort's two trajectory signatures identifies A. This does not touch Proposition 4, which is conditional on the homogeneous start it names; it removes this cohort from the set of cases where the test may be applied, and it leaves the identification

of A to signatures a dispersed start does not mimic, such as super-linear variance growth, which a single month of chess does not resolve.

A population-level check, which also fails to settle B

The dispersion the month produces is large, and it lands in a suggestive place. The entry s.d. of 5.3 grows to a month-end s.d. of 213, a factor of 40, which is essentially the s.d. of the entire active population that month (212 across all players with at least 30 games), rather than some intermediate value set by the horizon. This does *not* settle B, and Proposition 2 says it cannot. A cohort relaxing onto the population width is what the revelation of pre-existing differences would look like; it is equally what an amplifier with a stationary scale would produce, since manufacturing dispersion is exactly what $g > 0$ does. The trajectory carries nothing that separates the two readings. We report the fact because it sharpens the puzzle the next paragraph resolves, not because it resolves it.

The convergent channel refutes it

Admitting $\hat{k}$ overturns that reading (Table 2). In the near-equal cohort the month-end order that was uncorrelated with entry is strongly correlated with converged skill six months on, at $\text{corr}(\text{month-end order}, \hat{k}) = 0.79$ (95% bootstrap CI [0.70, 0.86] over players). The entry position, by contrast, does not predict $\hat{k}$ at all (-0.06). The order was not noise. One month of play *revealed* latent skill that the provisional entry rating had failed to capture. Through Eq. (7) this is a manufactured share $R = 1 - 0.79^2 \approx 0.38$ (95% CI [0.26, 0.51]). For both reasons given with (7), the amplified within-month noise being charged to manufacture and measurement error in $\hat{k}$ attenuating the correlation, 0.38 is an *upper* bound on manufacture, so at least about 62% of the order is revealed skill. The heterogeneous control sits where the model says it should, near the revelation end throughout, with $\text{corr}(\text{entry}, \hat{k}) = 0.74$ already at entry. The jump from entry to $\hat{k}$, from -0.06 to 0.79 in the near-equal cohort against 0.74 to 0.86 in the control, is the sharpest form of the lesson. As the entry estimate's dispersion shrinks, the share the entry position explains collapses, while the share that play reveals does not: capability was equally present in both cohorts (month-end s.d. 213 against 211), and only its visibility at entry differed. A near-symmetric start is therefore not evidence of near-equal capability, and reading it as such is the error the discriminator exists to prevent.

Cohort	Channel to $\hat{k}$	$corr$ (entry, $\hat{k}$)	$corr$ (month-end, $\hat{k}$)	R
Near-equal (1480–1520)	+6 mo	−0.06	0.79 [0.70, 0.86]	0.38 [0.26, 0.51]
Near-equal (1480–1520)	+12 mo	−0.09	0.71 [0.57, 0.81]	0.50 [0.35, 0.67]
Control (1000–2200)	+6 mo	0.74	0.86 [0.83, 0.89]	0.26 [0.21, 0.32]
Control (1000–2200)	+12 mo	0.70	0.81 [0.76, 0.85]	0.35 [0.28, 0.42]

Table 2. The discriminator on online chess (Lichess 2013). In the near-equal-rated cohort the month-end order is decoupled from the entry position (−0.06) yet strongly predicts converged skill (0.79), so the order that looked manufactured is mostly revealed. This gives $R \approx 0.38$ (Eq. (7)), an upper bound on manufacture. The control is skill-dominated at entry. The retained sample after the convergence filter (at least 20 later games) is 111 (+6 months) and 96 (+12 months) of the near-equal cohort, and 480 and 429 of the control. Brackets are 95% bootstrap confidence intervals over players ($B = 2000$); the near-equal cohort’s $corr(\text{month-end}, \hat{k})$ excludes zero by a wide margin at both horizons (`lichess_worked_example.py --bootstrap 2000`). Correlations are Pearson on log-ratings, as (7) requires of a variance share; the corresponding rank correlations are within 0.05 of them in every cell (0.77 against 0.79 in the headline cell), so nothing here turns on the choice. Correlations are displayed to two decimals while R is computed from the unrounded values, so a cell can differ from $1 - corr^2$ of its displayed correlation in the last digit (the control +12-month $R = 0.35$ comes from an unrounded 0.808).

What the example shows, and what it does not

It does not show that chess hierarchies are manufactured. It shows the opposite, that this one is mostly revealed, and that is the point. The value is methodological. A domain that looks like pure symmetry-breaking, with a near-equal start, an order decoupled from entry, and super-diffusive lock-in, everything

signature A is supposed to look like, is exposed as about 62% revelation the moment an exogenous, non-amplified estimate of capability is admitted. This is what Proposition 2 requires and Eq. (7) quantifies. The strong early-lead persistence here is skill-driven, a stable trait revealing itself, and is not read as the free-token signal of Section 8; chess ratings are only mildly revocable, so this example bears on the R discriminator, not on the free-token/toppling contrast. The trajectory-only reading was silent about the seed (B), as it must be, and, in this cohort, not entitled to its verdict on the dynamics (A) either, since a dispersed start mimics both of the signatures it rested on. The example is thus doubly cautionary: what looked like an amplifier acting on noise was neither shown to be an amplifier nor, where it acts, shown to be acting on noise, and one exogenous channel is what corrects both errors. Two caveats of sampling bound the number, not the logic. The order that forms is measured over a single month, a short horizon, and attrition is heavy, since the convergence filter retains about a fifth of the cohort, so $R \approx 0.38$ is one estimate on one cohort rather than a constant. What is robust is the sign of the correction, an entry-decoupled order that a convergent channel reveals to be mostly skill, and the fact that this correction is unavailable to any trajectory-only analysis. That is the whole of the claim.

A third caveat: the convergence filter is also a selection

Requiring at least 20 games in the reference month is what makes $\hat{k}$ converged, but it is not an innocent filter: remaining active for six months is not independent of skill, nor of how the entry month went, and such selection would raise $\text{corr}(\text{month-end}, \hat{k})$ and so lower R , pushing in the opposite direction to the attenuation and capability-drift arguments that make 0.38 an upper bound. The estimate is only usable if it is stable when the filter is relaxed, so we check that directly (`lichess_worked_example.py --survivorship`). Sweeping the threshold over 1, 5, 10, 20, 40 games moves the retained sample from 163 players down to 92 and returns $R = 0.45, 0.35, 0.35, 0.38, 0.33$: a spread of 0.12, comfortably inside the bootstrap interval $[0.26, 0.51]$ and with no monotone trend in the threshold, which is what selection of this kind would produce. The loosest filter, which keeps every player who appears at all six months later and so applies almost no selection, returns the *highest* R of the five, so if anything the headline number is conservative on this axis too. We conclude that the filter bounds the precision of $\hat{k}$ rather than manufacturing the correlation.

A fourth caveat: capability is not fixed here

The model takes k constant, whereas chess players improve, and this cohort is rich in newcomers: the median player gains +19 Elo between the month-end rating and the reference-window estimate six

months on (`lichess_worked_example.py` prints this per cohort). The convergent channel therefore estimates a slightly *later* k than the one the entry cohort carried, and that mismatch enters R as though it were manufacture. The horizon dependence makes the effect visible. Lengthening the window from +6 to +12 months takes R from 0.38 to 0.50, which is the opposite of what better averaging alone would give: a longer window averages the early fluctuation away more thoroughly, but it also lets the player change more, and here the second effect wins. So the two horizons yield two upper bounds rather than a refinement, and 0.38 is an upper bound in this further sense, beyond the attenuation argument. A cohort of entities whose capability is genuinely fixed would tighten the bound; that design is available in the Music Lab data of Section 11, where quality is fixed by construction, and it is unavailable here.

A fifth caveat: the channel is later, not different, and what that assumes

The discriminator of Section 3 asks for an estimate that does not pass through the amplifier's early lock-in, and the $\hat{k}$ used here is the *same* rating channel read months later rather than a different observable. Its claim to independence is therefore not sampling design but a property of the channel: that the rating dynamics are not themselves a runaway amplifier, so that the early fluctuation decays before the reference window. If that assumption failed, month-end and $\hat{k}$ would share the amplified early noise, the correlation would be inflated, and R would be biased *downward*, against the upper-bound reading the other caveats support; this is the one caveat that pushes in that direction. It is, however, the assumption most defensible in precisely this domain. A chess-rating update is additive in rating with a bounded step and a gain (the K-factor; in Lichess's Glicko-2, the rating deviation plays that role) that *shrinks* with games played, so the rating map is mean-reverting towards performance, not multiplicative in the rating itself: it lies outside the normal form of Eq. (1), and hundreds of intervening games average an early run of luck away by construction. The consistency of the two horizons' correlations (0.79 and 0.71) with the drift account of the fourth caveat points the same way, though the sharper signature, super-linear variance growth, is one a single month cannot resolve, as already noted. The upper-bound reading of R is thus conditional on the rating system not amplifying, an assumption argued, not proven, here; a $\hat{k}$ from a genuinely different observable (over-the-board ratings, engine-evaluated move quality) would remove the condition, and is the natural strengthening.

13. Conclusion

The amplifier of position has two faces, and the paper's work is to keep them apart. Its *dynamics* (A) are legible from a trajectory, though only in the restricted circumstances Section 8 sets out: on a homogeneous population directly, and elsewhere through a within-system contrast, since the homogeneity itself is not verifiable from the trajectory. It manufactures dispersion multiplicatively, and where the status carries a per-unit ceiling this is non-monotone in gain because its own saturating insurance eventually caps the spread; a free token has no downturn. Its stationary tail, for maintained, revocable statuses, is a power law whose index is fixed in closed form by the ratio of toppling rate to per-step amplification, so how contestable a status is sets how heavy its distribution becomes. Without the feedback it does not develop at all. It moves the ensemble while stranding a non-trivial share of its members in dead ends. Its *seed* (B), whether what got inflated was capability or position, real skill or amplified noise, is by contrast not legible from a trajectory at all. The map carries capability and position only through their sum, and it carries a lucky early draw and an inherited head start through one identical term, so the decomposition of the amplified input is not a function of the trajectory: the persistent component is recovered, its composition is not, and because the two readings induce the same law over observables, no quantity of data closes the gap. The split is not unknowable, only unknowable from this observable, and what a different observable returns is a proportion. The first of those is close to definitional once the amplifier is granted to be blind to provenance, and is refutable there, since a capability-dependent gain or noise would break it; the second is not definitional under any reading. That is a theorem, and it is why “forward” is a label applied afterwards rather than a direction the process ever chose. A lower threshold (Appendix A) sharpens the same point from the opposite end of the distribution. An arbitrary early loss is made unrecoverable, so a buffer, not a competence, decides who survives a shock.

We claim no new mechanism: the cascade is standard, the multiplicative machinery older still, and the propositions are elementary on purpose. Nor do we claim that real hierarchies are position amplifiers, which our data do not establish and which the trajectory signatures cannot establish on a heterogeneous population. We do not even claim that hierarchy is noise, since that is the one thing the non-identifiability result forbids a trajectory to support, and the single hierarchy we examine with an exogenous channel turns out mostly earned. The work is corrective, and three corrections are worth restating on their own.

The first is negative and, we think, the most useful. An excess of early-lead persistence over the diffusive law does not identify amplification. Entities that merely differ in a persistent trait produce that excess at zero gain, and because that null carries a free parameter it has no characteristic lock-in speed at all: it can be tuned to match any amplifier and any preferential-attachment alternative. Studies that read a fast-freezing order as evidence of a compounding mechanism are therefore reading something the data do not say. What replaces the magnitude test is a contrast internal to one system, and we give a worked one: the same votes of the same candidates read as a revocable position and as a monotone free token, so that any dispersion in candidate quality enters both arms alike and cancels, with a placebo to quantify what does not.

The second is the boundary itself. Grant the amplification, and the composition of what was amplified is still not a function of the trajectory: the persistent component is recovered, the split into capability and position is not, and since the competing accounts induce the same law over observables, more data cannot help. The boundary specifies what to go and measure instead: three exogenous channels can supply the missing component, and what they return is the manufactured share $R = 1 - corr^2$, a proportion rather than a verdict.

The third is the demonstration that the first two are not academic. Online chess offers a cohort wearing every appearance the amplifier reading predicts; a convergent skill estimate turns the verdict over to about 62% revealed skill. The reading we would have defended is the one the data refuse, which is what makes the case worth reporting.

What is constructive here is offered in the same spirit: a criterion for which positional currencies self-limit and which run away, a tail index fixed by how contestable a status is rather than by how fast it grows, and a statement of when amplification is required for structure at all. Each is stated so that a single observation could refute it. Whether a given hierarchy's order was earned or merely broke a symmetry somewhere remains answerable only with a handle from outside its own history, and unanswerable without one. Drawing that line, and marking the tests that cannot cross it, is the work we offer.

Appendix A. A complementary lower threshold: buffer asymmetry

The amplifier explains the *origin* of a gap: it breaks the symmetry and manufactures dispersion from noise. It says less about why the gap *persists* and widens rather than reverting. A second mechanism, complementary and equally capability-blind, does that, and reaches the paper's conclusion, position over

competence (“competence” here being the same fixed input k that Sections 4–5 call capability or quality), by an independent route. Where Proposition 1 derived an upper carrying capacity, symmetry suggests a *lower* one: a critical threshold S_{crit} , the minimum standing below which an entity’s flow is consumed by mere subsistence and none is left to grow. Below it the entity enters a survival mode in which its competence is switched off, all effort going to staying in place, and it climbs back to productive growth only after clearing a recovery band $S_{rec} > S_{crit}$ (a hysteresis, so the floor is a sticky trap, not a line one bounces over). This is the poverty trap of development economics [31], the buffer-stock logic of consumption under liquidity constraints [32], and the scarcity that consumes cognitive capacity [33], in the same normal form.

The consequence is an *asymmetry of consequence*: the buffer, the distance above S_{crit} , decides whether a given shock is a scratch or is fatal. The *same* shock a large buffer absorbs pushes a thin buffer into the trap, where competence stops paying. Figure 1 makes this concrete. In a twin test (two entities with *identical* competence and the *identical* shock sequence, differing only in starting buffer) a single early shock that the large buffer absorbs drops the thin one into survival mode, and it ends several times lower for no difference in competence or luck, only in position (left). At the population level the floor *decouples* outcome from competence just as the amplifier does: the rank correlation between final wealth and true competence falls (in the calibration, from 0.93 with no floor to 0.56 with it) while the correlation with the starting buffer rises to match it (from 0.13 to 0.56), so outcome comes to track position as much as ability (middle). And the tail of dead ends of Section 7 hardens into an *absorbing* trap, turning the distribution bimodal (right). As elsewhere, only the signs are claimed; the magnitudes depend on S_{crit} and the survival rate (`critical_threshold.py`).

Two mechanisms, then, reach the same anti-meritocratic conclusion from opposite ends of the distribution. The amplifier, at the top, inflates an arbitrary early lead into a runaway; the threshold, at the bottom, makes an arbitrary early loss unrecoverable. Neither needs a difference in competence to produce a difference in outcome: the first manufactures the gap, the second locks it in, and together they bracket the claim that where one stands can matter more than how able one is.

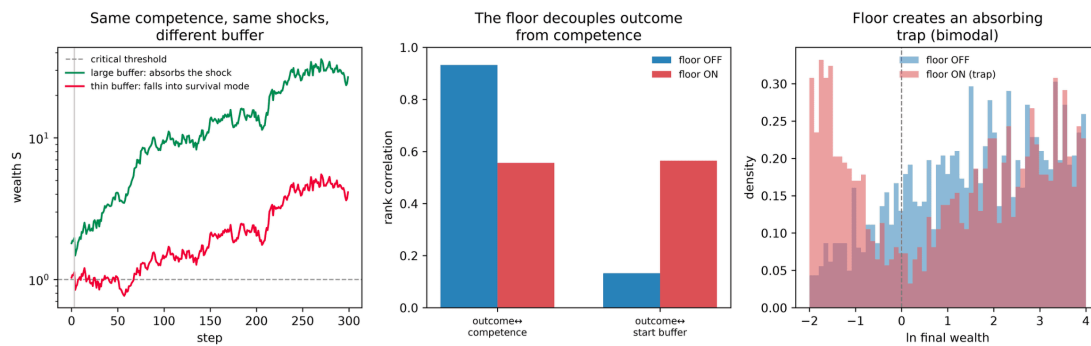


Figure 1. A lower absorbing threshold decouples outcome from competence, by buffer asymmetry. *Left:* two entities with identical competence and the identical shock sequence; the same early shock a large buffer absorbs drops a thin buffer into survival mode, and it ends far lower, with the same competence and the same luck, differing only in position. *Middle:* across the population, adding the floor drops the outcome–competence rank correlation ($0.93 \rightarrow 0.56$) while raising the outcome–starting–buffer correlation ($0.13 \rightarrow 0.56$), so position comes to matter as much as competence. *Right:* the floor turns the dead-end tail of Section 7 into an absorbing trap, and the final distribution bimodal.

Reproducibility

All code, bundled data, and a single-command runner are available at <https://github.com/dkrse/symmetry-breaking-amplifier>. One entry point reproduces every figure and printed number: `./run_all.sh` bootstraps a local virtual environment and runs the offline simulations and bundled-data tests in about half a minute with no network access; `--power` adds the Monte-Carlo power analysis, `--online` the Wikipedia RfA and GitHub passes, and `--lichess` the online-chess worked example (three monthly Lichess dumps ^[30], ~150 MB, fetched once and cached). The bootstrap confidence intervals of Sections 11 and 12 are reproduced by the `--bootstrap 2000` option of `wiki_rfa_toppling.py` and `lichess_worked_example.py`, and the survivorship sweep by `lichess_worked_example.py --survivorship`. The heterogeneity placebo of Section 11 and the within-month trajectory signatures of Section 12 are printed by those same two scripts on every run, so no number in either section is asserted without being computed. The tail-index check of Section 6 is `kesten_tail.py` and the horizon-scaling fit of Section 2 is `optimal_gain_fit.py`; both are offline and are grouped with the Monte-Carlo power analysis under `--power`, which takes about 22 minutes in total and is dominated by the tail-index check. Every run reports its own per-script wall-clock. All randomness is seeded (`numpy.random.default_rng`) and every figure and table regenerates

deterministically; the mapping from each result to its script is documented in the repository (README.md, scripts/ALGORITHMS.md). Every dataset is public and free, with source URLs, exact download commands, and SHA-256 integrity hashes recorded in scripts/data/DATA_SOURCES.md; the bundled inputs are listed in scripts/data/SHA256SUMS, which run_all.sh verifies before running.

Statements and Declarations

Funding

The author received no specific funding for this work.

Potential Competing Interests

The author declares no competing financial or non-financial interests.

Ethics

Not applicable. The study involves no human or animal subjects and uses only publicly available secondary datasets; ethics approval was not required.

Consent for publication

Not applicable. The research involves no participants.

Data Availability

All data are publicly available. The analysis uses the Wikipedia Requests-for-Adminship corpus from the Stanford SNAP collection (wiki-RfA; <https://snap.stanford.edu/data/wiki-RfA.txt.gz>), the December-2015 Crunchbase open funding-rounds export (<https://github.com/notpeter/crunchbase-data>), repository star timestamps from the GitHub GraphQL API, the Salganik–Dodds–Watts Music Lab dataset (Princeton release, <https://doi.org/10.34770/y56c-ym90>), and the Lichess Open Database (monthly standard-rated PGN dumps, <https://database.lichess.org/>, CC0). Every bundled input is SHA-256 hashed in scripts/data/SHA256SUMS; the on-demand Lichess dumps (150 MB, fetched at run time rather than bundled) are hashed in scripts/data/DATA_SOURCES.md. Each source URL and its exact download command are recorded there as well.

Code Availability

All code to reproduce the paper is openly available at <https://github.com/dkrse/symmetry-breaking-amplifier>

CRedit authorship contribution

Kristian Sestak: Conceptualization, Methodology, Formal analysis, Software, Data curation, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing.

Acknowledgements

The author thanks the maintainers of the Stanford SNAP collection, the open Crunchbase data mirror, the GitHub platform, the Lichess Open Database, and the Music Lab dataset for making these resources publicly available.

Footnotes

¹ This is a preprint. The manuscript is under peer review at *Physica A: Statistical Mechanics and its Applications* (submitted July 2026) and has not yet been peer reviewed.

References

1. ^{a, b, c}Salganik MJ, Dodds PS, Watts DJ (2006). "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market." *Science*. **311**(5762):854–856. doi:[10.1126/science.1121066](https://doi.org/10.1126/science.1121066).
2. ^{a, b}Muchnik L, Aral S, Taylor SJ (2013). "Social Influence Bias: A Randomized Experiment." *Science*. **341**(6146):647–651. doi:[10.1126/science.1240466](https://doi.org/10.1126/science.1240466).
3. ^{a, b}van de Rijt A, Kang SM, Restivo M, Patil A (2014). "Field Experiments of Success-Breeds-Success Dynamics." *Proc Natl Acad Sci U S A*. **111**(19):6934–6939. doi:[10.1073/pnas.1316836111](https://doi.org/10.1073/pnas.1316836111).
4. ^{a, b}Heckman JJ (1981). "Heterogeneity and State Dependence." In: Rosen S, editor. *Studies in Labor Markets*. University of Chicago Press. pp. 91–139. <https://www.nber.org/books-and-chapters/studies-labor-markets/heterogeneity-and-state-dependence>.
5. ^{a, b}Heckman JJ (1981). "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process." In: Manski CF, McFadden D, editors. *Structural Analysis*

- sis of Discrete Data with Econometric Applications. MIT Press. pp. 179–195. <https://eml.berkeley.edu/~mcfadden/discrete/ch4.pdf>.
6. ^{a, b, c}Gould RV (2002). "The Origins of Status Hierarchies." *Am J Sociol.* **107**(5):1143–1178. doi:[10.1086/341744](https://doi.org/10.1086/341744).
<https://www.jstor.org/stable/10.1086/341744>.
 7. ^{a, b, c}Sah RK, Stiglitz JE (1986). "The Architecture of Economic Systems: Hierarchies and Polyarchies." *Am Econ Rev.* **76**(4):716–727. <https://ideas.repec.org/a/aea/aecrev/v76y1986i4p716-27.html>.
 8. ^{a, b}Merton RK (1968). "The Matthew Effect in Science." *Science.* **159**(3810):56–63. doi:[10.1126/science.159.3810.56](https://doi.org/10.1126/science.159.3810.56).
 9. ^{a, b}Barabási A-L, Albert R (1999). "Emergence of Scaling in Random Networks." *Science.* **286**(5439):509–512. doi:[10.1126/science.286.5439.509](https://doi.org/10.1126/science.286.5439.509).
 10. ^{a, b}Kawakatsu M, Chodrow PS, Eikmeier N, Larremore DB (2021). "Emergence of Hierarchy in Networked Endorsement Dynamics." *Proc Natl Acad Sci U S A.* **118**(16):e2015188118. doi:[10.1073/pnas.2015188118](https://doi.org/10.1073/pnas.2015188118).
 11. ^{a, b}Deffuant G, Roubin T (2022). "Emergence of Group Hierarchy." *Physica A.* **607**. doi:[10.1016/j.physa.2022.128422](https://doi.org/10.1016/j.physa.2022.128422).
 12. ^ΔFisher RA (1930). *The Genetical Theory of Natural Selection*. Clarendon Press. doi:[10.5962/bhl.title.27468](https://doi.org/10.5962/bhl.title.27468).
 13. ^ΔKimura M (1968). "Evolutionary Rate at the Molecular Level." *Nature.* **217**:624–626. doi:[10.1038/217624a0](https://doi.org/10.1038/217624a0).
 14. ^ΔPeters O (2019). "The Ergodicity Problem in Economics." *Nat Phys.* **15**:1216–1221. doi:[10.1038/s41567-019-0732-0](https://doi.org/10.1038/s41567-019-0732-0).
 15. ^ΔSimon HA (1955). "On a Class of Skew Distribution Functions." *Biometrika.* **42**(3–4):425–440. doi:[10.1093/biomet/42.3-4.425](https://doi.org/10.1093/biomet/42.3-4.425).
 16. ^ΔYule GU (1925). "A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S." *Philos Trans R Soc Lond B Biol Sci.* **213**:21–87. doi:[10.1098/rstb.1925.0002](https://doi.org/10.1098/rstb.1925.0002).
 17. ^{a, b}Levy M, Solomon S (1996). "Power Laws are Logarithmic Boltzmann Laws." *Int J Mod Phys C.* **7**(4):595–601. doi:[10.1142/S0129183196000491](https://doi.org/10.1142/S0129183196000491).
 18. ^{a, b}Sornette D, Cont R (1997). "Convergent Multiplicative Processes Repelled from Zero: Power Laws and Truncated Power Laws." *J Phys I.* **7**(3):431–444. doi:[10.1051/jp1:1997169](https://doi.org/10.1051/jp1:1997169).
 19. ^{a, b}Gabaix X (1999). "Zipf's Law for Cities: An Explanation." *Q J Econ.* **114**(3):739–767. doi:[10.1162/003355399556133](https://doi.org/10.1162/003355399556133).
 20. ^{a, b}Bouchaud J-P, Mézard M (2000). "Wealth Condensation in a Simple Model of Economy." *Physica A.* **282**(3–4):536–545. doi:[10.1016/S0378-4371\(00\)00205-3](https://doi.org/10.1016/S0378-4371(00)00205-3).

21. ^a ^b ^c ^d Kesten H (1973). "Random Difference Equations and Renewal Theory for Products of Random Matrices." *Acta Math.* **131**:207–248. doi:[10.1007/BF02392040](https://doi.org/10.1007/BF02392040).
22. ^a West GB, Brown JH, Enquist BJ (1997). "A General Model for the Origin of Allometric Scaling Laws in Biology." *Science*. **276**(5309):122–126. doi:[10.1126/science.276.5309.122](https://doi.org/10.1126/science.276.5309.122).
23. ^a Dunbar RIM (1992). "Neocortex Size as a Constraint on Group Size in Primates." *J Hum Evol.* **22**(6):469–493. doi:[10.1016/0047-2484\(92\)90081-J](https://doi.org/10.1016/0047-2484(92)90081-J).
24. ^a Aghion P, Tirole J (1997). "Formal and Real Authority in Organizations." *J Polit Econ.* **105**(1):1–29. doi:[10.1086/262063](https://doi.org/10.1086/262063).
25. ^a Hill BM (1975). "A Simple General Approach to Inference About the Tail of a Distribution." *Ann Statist.* **3**(5):1163–1174. doi:[10.1214/aos/1176343247](https://doi.org/10.1214/aos/1176343247).
26. ^a Boehm C (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Harvard University Press. doi:[10.4159/9780674028449](https://doi.org/10.4159/9780674028449).
27. ^a Leskovec J, Krevl A (2014). "SNAP Datasets: Stanford Large Network Dataset Collection." <http://snap.stanford.edu/data>. Wikipedia Requests for Adminship (wiki-RfA): West R, Paskov HS, Leskovec J, Potts C (2014). "Exploiting Social Network Structure for Person-to-Person Sentiment Analysis." *Trans Assoc Comput Linguistics*. **2**:297–310. doi:[10.1162/tacl_a_00184](https://doi.org/10.1162/tacl_a_00184).
28. ^a Salganik M (2026). "Success and Failure in Cultural Markets [Dataset]." Princeton University. doi:[10.34770/y56c-ym90](https://doi.org/10.34770/y56c-ym90).
29. ^a Salganik MJ, Watts DJ (2008). "Leading the Herd Astray: An Experimental Study of Self-Fulfilling Prophecies in an Artificial Cultural Market." *Soc Psychol Q.* **71**(4):338–355. doi:[10.1177/019027250807100404](https://doi.org/10.1177/019027250807100404).
30. ^a ^b Lichess (2013–present). "Lichess Open Database (standard rated games)." <https://database.lichess.org/>.
31. ^a Bowles S, Durlauf SN, Hoff K, editors (2006). *Poverty Traps*. Princeton University Press. doi:[10.1515/9781400841295](https://doi.org/10.1515/9781400841295).
32. ^a Carroll CD (1997). "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis." *Q J Econ.* **112**(1):1–55. doi:[10.1162/003355397555109](https://doi.org/10.1162/003355397555109).
33. ^a Mullainathan S, Shafir E (2013). *Scarcity: Why Having Too Little Means So Much*. Times Books. ISBN [978–0-8050-9264-6](https://doi.org/10.1002/9780805092646). <https://openlibrary.org/isbn/9780805092646>.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.