Two firms file the same numbers

Timeliness and durability are not separately identified from a reported series

Jason C. Braatz
Independent researcher
jason@braatz.ai

Preprint. Posted 2026-08-29. Not peer reviewed, not submitted.

Declaration of interest. The author is employed by a company building accounting software for very small businesses. This work was conducted independently, on personal time, and without company funding, data or direction.

Abstract

Model the reporting layer of a balance sheet as a low-pass filter on a physical layer that degrades whether or not anyone records it: a share $\boldsymbol{\varphi}$ of each true change passes through at once, the remainder released at rate $\boldsymbol{\alpha}$ from an unrecognised gap, against a physical decay rate $\boldsymbol{\delta}$.

The triples $(\alpha, \delta, \varphi)$ and $(\delta, \alpha, \varphi\delta/\alpha)$ generate the identical reported series. The filter's two roots exchange and the exchange preserves $\varphi\delta$ exactly, so a reported series contains the product of timeliness and decay and nothing further about either. A prompt reporter of a durable asset and a slow reporter of a perishable one file the same numbers to fourteen decimal places while their physical stocks differ by a factor of 250,000. Where the asset's physical scale is unobserved (which is every firm-level series) the ambiguity is not two-valued but a continuum: a factor of **1.67** in that unobserved scale sweeps φ across the entire unit interval.

The corollary is cross-sectional and it is sharp. Classes are ordered by $\boldsymbol{(1-\varphi)\odot\delta}$, divided elementwise by $\boldsymbol{(\alpha-\delta)}$: not by φ. On the four GAAP asset classes, with the decay rates the standards themselves imply, the composite does not blur the intended ranking; it **inverts** it, Kendall $\tau = -1$. Drawing δ independently across classes, the intended ordering survives in **11.5%** of 4,000 ladders. The boundary between the region where such a design works and the region where it does not is exact, is drawn in quantities the design already declares, and sits at a δ-leverage-to- budget ratio of **0.61**; the GAAP ladder sits at **2.58**. This constrains every

cross-sectional conditional-conservatism comparison, because the burden it moves is the burden of showing that asset life is constant across the compared groups.

The repair follows from the theorem rather than from any empirical finding: anything supplying δ from outside the series restores φ. Returns do it, at a price that is a rate rather than a proof and that runs backwards for the slowest assets. **Disclosed useful lives do it for free**, for the two classes the standards put on a schedule, and license a within-life-band design that needs no new data. At $\delta = 0$ there is no parameter to recover, because φ has left the dynamics.

The framework's sharpest prediction (recognition lag ordered by GAAP asset class) was pre-registered, tested on 688 EDGAR-derived events across two sectors declared in advance, powered at 0.95–1.00, and **failed** (Jonckheere–Terpstra $z = -0.290, -0.095$). §9 reports the failure at full length and §10 states exactly what it retracts. The identification result does not excuse it: the one statistic the composite spares is the timing statistic the registration used, and what defeated that registration was a second identification gap upstream of the first.

Keywords: identification · conditional conservatism · reporting lag · impairment · observational equivalence · asset life · preregistration

JEL classification: M41, D80, C18, G14, E01

---

1 · The result

Two firms file identical financial statements for a hundred years. Every reported number agrees to fourteen decimal places. At the end of it, one of them owns an asset worth a third of what it started with, and the other owns an asset worth four parts in a million of what it started with.

Nothing in the filings distinguishes them, and no estimator applied to those filings ever will. This paper proves that, states its reach, works out what it does to the measurement practice that currently reads such filings, and gives the repair.

The mechanism is a filter with two rates. A physical asset degrades at rate δ whether or not anybody writes it down. The reporting layer sees a share φ of each period's degradation immediately (announced capital expenditure, a disclosed impairment, a write-down someone had to sign) and defers the rest into an unrecognised gap that it releases at rate α. Timeliness is φ. Durability is δ.

Solve that filter and the reported series is a linear combination of two geometrics whose exponents are α and δ, and **the two exponents are exchangeable**. Send $(\alpha, \delta, \varphi)$ to $(\delta, \alpha, \varphi\delta/\alpha)$ and every reported number comes back unchanged. The quantity the exchange preserves is exactly $\varphi\delta$.

So a reported series does not make timeliness hard to estimate. **It does not contain timeliness.** It contains $\varphi\delta$, and $\varphi\delta$ is the same number for a firm that reports promptly on a durable asset and a firm that reports slowly on a perishable one. The two are different companies with different prospects and one balance sheet between them.

That would be a curiosity if timeliness were a curiosity. It is not: the recognition speed of economic losses is the object of an entire measurement literature (Basu's asymmetric-timeliness coefficient, Khan and Watts's C_Score, Ball and Shivakumar's piecewise accruals measure, Givoly and Hayn's accumulated negative accruals, Bushman and Williams's DELR) every one of which takes a reported series as input and reads a recognition property off it. If a timeliness parameter reaches that series only in product with an asset-life parameter, then a cross-sectional comparison of any of those measures is a comparison of the product.

And the composite does not merely add noise to such a comparison. §5 works out the ranking a reader can actually compute — $\mathbf{(1-\varphi)} \odot \boldsymbol{\delta} \oslash \mathbf{(\alpha-\delta)}$, with $\odot$ the Hadamard product and $\oslash$ elementwise division, both taken across asset classes — and evaluates it on the four classes US GAAP distinguishes, at the decay rates those same standards imply. The intended ordering does not blur. It reverses, Kendall $\tau = -1$. The class predicted to defer the most defers the least, because it barely deteriorates, and there is little to defer.

That reversal is not an accident of one calibration. §5 gives the exact boundary of the region in which a timeliness-ordered cross-section recovers the ordering it imposed, expressed in two quantities the design already declares: its **budget**, the mean per-rung change in $\log(1-\varphi)$, and the ladder's **δ leverage**, the mean per-rung $|\Delta\log\delta - \Delta\log(\alpha-\delta)|$. The design reads what it ordered, more likely than not, only while leverage stays under about three fifths of budget. The GAAP ladder sits at four times that.

The validity condition for a timeliness-ordered design is therefore a statement about δ — which is the quantity §3 has just shown the reported series does not contain. A researcher cannot establish that such a design is sound without already possessing exactly what the design was built to avoid needing.

The repair is in the theorem's own statement. The series determines $\varphi\delta$, so anything that supplies δ from outside the

series restores φ. Two things do. Returns break the two-point equivalence immediately, and §7 prices that break precisely: identification fades in as a power law in return volatility, never attains the root-T rate at any horizon, and (for assets decaying more slowly than about one per cent a year) *reverses sign*, so more news makes the reading worse. And for the two classes the standards place on an amortisation schedule, the outside determination is already published: **a disclosed useful life is an audited estimate of δ that was not derived from the series whose timeliness is in question.** A design comparing timeliness only within a life band reads φ rather than φδ, is diagonal-safe, carries a built-in negative control, and runs on data that already exists.

Contributions.

1. **An exact observational equivalence, with its proof, its lineage and its reach** (§3). The filter's two roots exchange preserving φδ. The structure is the Bateman function's, the phenomenon is pharmacokinetics' *flip-flop*, and the nearest economic instance is Nerlove's: from which the accounting case differs in exactly the way that matters.
2. **The identified set is a continuum, not two points** (§4). Where the physical scale is unobserved, a factor of 1.67 in that scale spans all of φ, every member reproducing the series to 2×10^{-16}. Bounding the opening gap at ten per cent still leaves 31.7% of the unit interval.
3. **The cross-class corollary and its exact validity boundary** (§5). The observable ranking is $(1 - \varphi) \odot \delta \oslash (\alpha - \delta)$; on the GAAP ladder it is the reverse of the ranking of φ; and the region in which a φ-ordered design succeeds has a closed-form boundary in quantities the design declares.
4. **A constraint on the conditional-conservatism measures** (§6), stated with its three qualifications and with the specific circumstance under which those measures remain sound.
5. **A repair that needs no new data, and its price** (§7). Disclosed useful lives supply δ from outside the series. Returns supply it too, at a rate that degrades as the asset quietens and reverses below $\delta \approx 0.01$.
6. **The recognition hazard's shape, measured rather than assumed** (§8), and the finding that the constant hazard the closed form assumes was never doing structural work, but was holding up the model's *domain*, through a condition on the lag distribution's tail.
7. **A pre-registered severe test and its failure** (§9), registered before the data were touched, replicated in a sec-

ond sector declared in advance, controlled against a label-permutation null, powered at 0.95–1.00, and lost with the stopping rule honoured.

The failed prediction is in the body and in the abstract rather than in an appendix of abandoned approaches: a registered prediction that was tested and lost is a **result**, and §9 reports it as one. §10 states exactly what it retracts, which is more than a reader might expect and less than the whole framework.

The three propositions the filter was originally built inside (composition, decay, atomism) are in **Appendix A**, at the length they have earned. **No result in §§2–9 depends on them.** They are there because the filter came from somewhere and a reader is entitled to see where; they are not there because anything rests on them.

---

2 · The filter

Two layers and one parameter that matters.

The physical layer decays at an entropy rate net of maintenance:

$$\mathbf{E(t+1) = E(t) \cdot (1 - d \cdot (1 - m))}$$

Of each true change, a share $\boldsymbol{\varphi}$ is *observable* (announced capital expenditure, a disclosed impairment, a write-down someone had to sign) and reaches the claim layer at once. The remaining $(1 - \varphi)$ is deferred maintenance and technical debt: real, accruing, and absent from the statements. It accumulates in the gap and is recognised at rate α per period:

$$\mathbf{C(t+1) = C(t) + \varphi \cdot \Delta E + \alpha \cdot gap(t)}, \quad \mathbf{gap(t) = E(t) - C(t)}$$

When the unrecognised gap exceeds a threshold share θ of physical wealth, the deferral becomes unsustainable and the claim layer snaps to the physical one. That discontinuity is the recognition event, and its magnitude is exactly the information that had been withheld.

Write $\boldsymbol{\delta}$ for the effective decay rate $d(1 - m)$ throughout. Where the filter is examined in isolation the recognition mechanism is disabled ($\theta = \infty$), because otherwise the snap timing truncates the measurement window and any lag statistic reports the recognition schedule rather than the filter. Standing parameters: $E_0 = 100$, $d = 0.05$, $m = 0.6$ (so $\delta = 0.02$), $\alpha = 0.05$, $\theta = 0.25$, 400 periods.

Terminology. The discrete event is a **recognition event**; where the referent is literally ASC 350 it is an **impairment loss**. It is deliberately not called a *correction*: ASC 250 reserves that word for the repair of an **error**, which a change in estimate driven by later information is not.

Two conditions bound what this filter models, and both are restrictions rather than assumptions of convenience. The first makes the wedge one-signed: reported value may fall and may not rise. Under US GAAP, the regime this paper's sample files in, there is no upward revaluation of property, plant and equipment and no impairment reversal for goodwill or indefinite-lived intangibles; IAS 36 requires reversal for non-goodwill assets and IAS 16 permits revaluation, so the condition holds for these filers and is not general. The second restricts the domain to degradation on which conservatism has nothing further to bite: no impairment trigger, no estimable expected loss, no observable event to key recognition to. **Where a loss is estimable, recognition is faster than the market and this model predicts nothing.**

φ is not a fudge factor, and the distinction is load-bearing. φ is the *observability of the degradation*: not a free parameter absorbing residual error, and not the reporting entity's willingness to report. That is what carries the model past the efficient-markets objection (Supplementary S2): a sophisticated reader can price what is knowable, and $(1 - \varphi)$ is by construction the share that is not.

One consequence of the filter is worth stating before the theorem, because it is the cleanest thing in the model and it is exact. With the recognition mechanism disabled, substituting $E(t+1) - C(t) = \text{gap}(t) + \Delta E$ into the two recursions gives

$$\mathbf{gap(t+1) = (1 - \alpha)\cdot gap(t) + (1 - \varphi)\cdot \Delta E(t)}$$

so with $\text{gap}(0) = 0$ the gap at every t is $(1 - \varphi)$ times its value on the $\varphi = 0$ path, and since $\Delta E < 0$ throughout, every term shares a sign and the absolute integral inherits the factor exactly:

$$\mathbf{D(\varphi) = (1 - \varphi) \cdot D(0)}$$

A doubling of unobservability doubles the integral of what the statements owe, exactly. The simulation reproduces this to a relative error of 10^{-15} across φ and the test suite asserts it. The recognition *lag*, by contrast, is sigmoidal in $(1 - \varphi)$ (1 period at $\varphi = 0.9$, 3 at 0.8, 14 at 0.5, 26 at 0.0) so the quantity of undelivered information is linear in unobservability while the delay in delivering it is not. The two should not be expected to

move together, and §5 is largely a consequence of their not doing so.

---

3 · The theorem

3.1 · The class index, and why the product is elementwise

A firm holds several classes of asset and the standards treat them differently on purpose. Index the classes *i*. Each carries its own δ_i, φ_i, α_i and θ_i, and §2's recursions become one line in vectors:

$$\mathbf{C}(t+1) = \mathbf{C}(t) + \boldsymbol{\varphi} \odot \Delta\mathbf{E} + \boldsymbol{\alpha} \odot \mathbf{gap}(t),\ \mathbf{gap}(t) = \mathbf{E}(t) - \mathbf{C}(t)$$

The elementwise form is a substantive claim rather than a notational economy: **the reporting layer is diagonal in class space.** A dollar of unrecognised deterioration in a distribution centre does not force recognition against a trademark. Each class's filter reads its own gap and nothing else.

That claim is an assumption and it is false in detail: §11 says where, and proposes the test. Without the class index the identification result below is a remark about a single parameter. With it, the remark acquires a corollary about *rankings*, and rankings are what the empirical literature on this subject estimates.

3.2 · The exchange

Take one class and disable the recognition mechanism. Substituting $\Delta E = -\delta E(t)$ collapses the pair of recursions to a single line:

$$C(t+1) = C(t)(1-\alpha) + E(t)(\alpha - \varphi\delta),\ E(t) = E_0(1-\delta)^t$$

φ appears once, in the product $\varphi\delta$, and nowhere else. With the books opening square — $C(0) = E(0) = E_0$, an asset carried at cost on the day it is acquired — this solves in closed form. Writing $A = 1 - \alpha$ and $D = 1 - \delta$ for the two roots,

$$\mathbf{C(t) = E_0 \cdot [\ \delta(1-\varphi)\, A^t - (\alpha - \varphi\delta)\, D^t\ ] / (\delta - \alpha)}$$

The reported series is a linear combination of two geometrics whose exponents are the reporting rate and the physical decay rate. Now exchange them. Send $(\alpha, \delta, \varphi) \longrightarrow (\delta, \alpha, \varphi')$, which swaps A and D, and ask what φ' must be for the series to return unchanged. The A^t coefficient requires $\delta(1-\varphi) = \delta - \varphi'\alpha$. The D^t coefficient requires $\alpha - \varphi\delta = \alpha(1-\varphi')$. Both reduce to

$$\boldsymbol{\varphi'\alpha = \varphi\delta}$$

Two equations, one unknown, and they agree.

That coincidence is the theorem. The system is overdetermined and consistent, so the exchange is not approximate, not local, and not a matter of conditioning.

> **Observational equivalence.** The parameter triples $(\alpha, \delta, \varphi)$ and $(\delta, \alpha, \varphi\delta/\alpha)$ generate the *identical* reported series. The filter's two roots (the reporting rate and the physical decay rate) are exchangeable, and the quantity preserved by the exchange is exactly $\varphi\delta$.

The numerical confirmation runs alongside the proof rather than in place of it. The largest deviation between a series and its mirror is 8×10^{-16} (the arithmetic, not the model) against 3×10^{-1} when the mirror's φ' is replaced by the value preserving the *unrecognised gap* instead of the reported series. Fourteen orders of magnitude between the right conserved quantity and a plausible wrong one. Admissibility requires $\varphi\delta \le \alpha$, which holds at every parameter setting used anywhere in this paper.

And this is not a property of the particular filter. Any model in which reporting lag attenuates a physical signal will multiply a timeliness parameter by an asset-life parameter somewhere, because the observable is a rate times a duration. What this model contributes is making the product explicit enough to be checked.

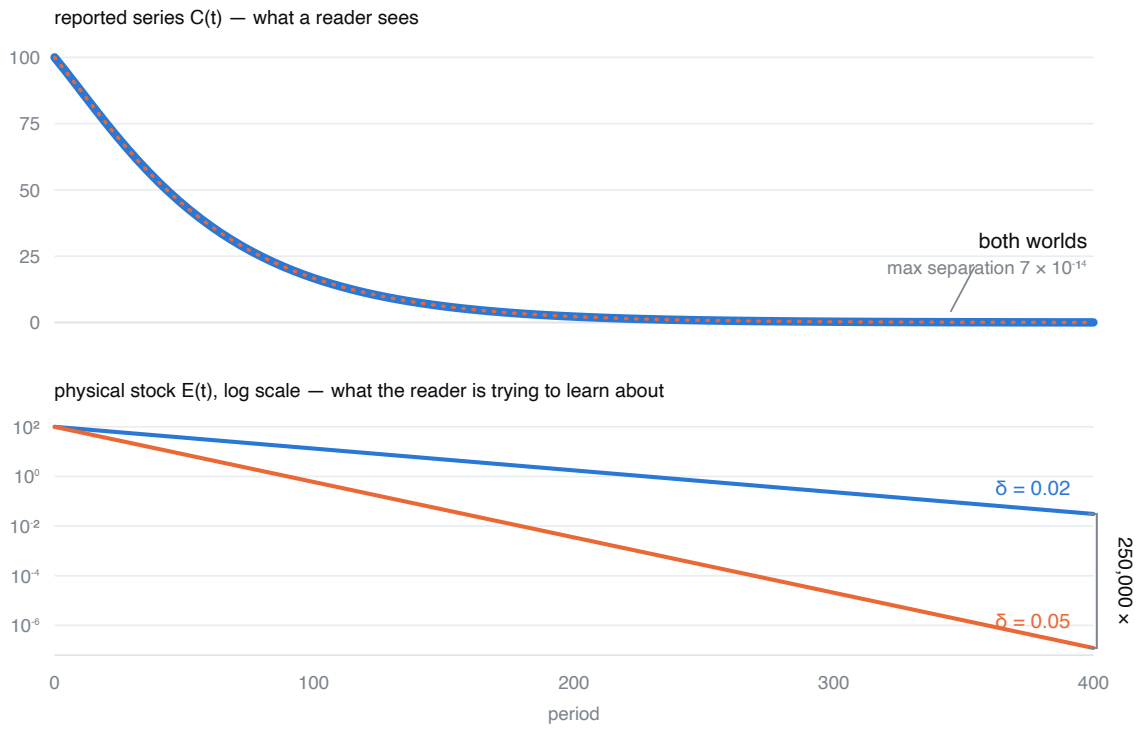

Figure 1 — Two firms, one filing. *Top:* the reported series C(t) for (α, δ, φ) = (0.05, 0.02, 0.60) and for its mirror (0.02, 0.05, 0.24), over 400 periods. The two curves are one curve; the maximum separation anywhere on the path is 7×10^{-14}, which is double-precision arithmetic. *Bottom:* the physical stock E(t) for the same two worlds, log scale. They end a factor of 250,000 apart. **The top panel is what a reader sees. The bottom panel is what the reader is trying to learn about.**

What the two worlds disagree about is the firm, not the filing. The mirror is not a mathematical curiosity with no economic content. It is a slow reporter of a fast-decaying asset: at t = 400 its physical stock stands at 4×10^{-6} of the original world's, a book value sitting above an asset that has all but evaporated. That is a recognisable kind of company, and it files the same statements, to fourteen decimal places, as the prompt reporter of the durable asset.

3.3 · The mathematics is old, and where it comes from decides how it behaves

Subtract E(t) from the closed form and the unrecognised gap is

$$G(t) = E_0 \cdot (1 - \varphi)\, \delta \cdot S(t),\ S(t) = (A^t - D^t)/(\delta - \alpha)$$

a scalar amplitude carrying every trace of φ, multiplied by a shape function invariant under exchanging the roots. **S is the Bateman function**, written down in 1910 for the daughter activity of a radioactive decay chain (Bateman, 1910), and its exchange symmetry is why pharmacokinetics has a name for this: **flip-flop**, the interchange of an absorption rate constant with an elimination rate constant in a one-compartment model, which leaves the concentration–time profile unmoved (Garrett, 1994).

Kuan, Wright and Duffull (2023) place flip-flop as an issue of *local* identifiability, "in that there exists a finite set of parameter values (rather than a single set) that solves the problem": precisely the two-point structure above. Their caution should be carried across with the result: they hold that the competing solutions are "not simply a function of swapping the rate constants" but a partial permutation of the parameter set, with n + 1 of them for an n-compartment model. **The accounting case here, where the exchange *is* a clean swap of two roots, is the simplest member of that family rather than the general one.** The framework underneath is Bellman and Åström's (1970), which defines structural identifiability by what the input–output map determines and tests it through the transfer function, and a transfer function fixes its poles as a set, which is the exchange above stated once and for all.

Economics has its own instance and it sits closer to this paper than either. **Nerlove's supply-response model** (1958) stacks adaptive expectations at rate β on partial adjustment at rate γ; eliminating the two unobservables leaves a reduced form in which every systematic coefficient is a *symmetric function* of β and γ: $[(1 - \beta) + (1 - \gamma)]$ on the lagged dependent variable, $-(1 - \beta)(1 - \gamma)$ on its second lag, $\beta\gamma a_1$ on the price. The conditional mean is exactly invariant under exchanging the two behavioural rates. What distinguishes them is the disturbance, $\gamma[u(t) - (1 - \beta)u(t-1)]$, which is not symmetric.

That the tie is broken by the error process rather than by the systematic part is a property of the Nerlovian model which the filter here does not inherit, and §7 says what breaks it instead. The distinction matters for the practical reason: a researcher who has met this shape in Nerlove will expect the disturbance to rescue identification, and here it does not.

Why the ambiguity bites here and not everywhere. It needs a **free scale parameter** to hide in. A decay chain's parent activity is measured independently, so the exchange changes the observed curve by a detectable factor and nothing is lost. A plasma profile has an unknown volume of distribution. A balance sheet has an unknown $(1 - \varphi)$. Accounting is on the pharmacokinetic

side of that line, and §4 is what it costs.

---

4 · The identified set is a continuum

The two-point ambiguity of §3 is the *favourable* case. It assumes the books open square and the physical scale is known: an asset followed from acquisition, where cost fixes E_0 and the gap opens at zero. Relax either and the result gets worse, and the relaxation is the empirical norm rather than an edge case.

Count first. The closed form has two roots and two amplitudes: four numbers, and that is everything a reported series contains. The model has five parameters: α, δ, φ, the physical scale E_0, and any gap g_0 already open when observation starts. **Five into four does not go, and the shortfall lands on φ.**

Two consequences, both exact. First, opening the books with a gap already in place does not rescue identification: the mirror survives with the same g_0 under the shifted map $\varphi' = [\varphi\delta + g_0(\alpha - \delta)]/\alpha$, and the conserved quantity generalises to $(\varphi - g_0)\delta = (\varphi' - g_0)\alpha$. The obvious escape (*real firms are not observed from acquisition*) makes matters worse rather than better.

Second, and this is the sharp form: **when the physical scale is not observed, φ is not two-valued. It is free.**

Fix a reported series generated at $\varphi = 0.60$ in a world whose books open 15% above the physical asset, and ask what other parameter vectors reproduce it exactly. Assuming a physical scale of 0.76 implies $\varphi = 0$. Assuming 1.27 implies $\varphi = 1$. Every assumption in between implies an intermediate φ, and every one of them regenerates the observed series to $\mathbf{2 \times 10^{-16}}$.

> **A factor of 1.67 in the unobserved physical scale spans the entire unit interval of timeliness.**

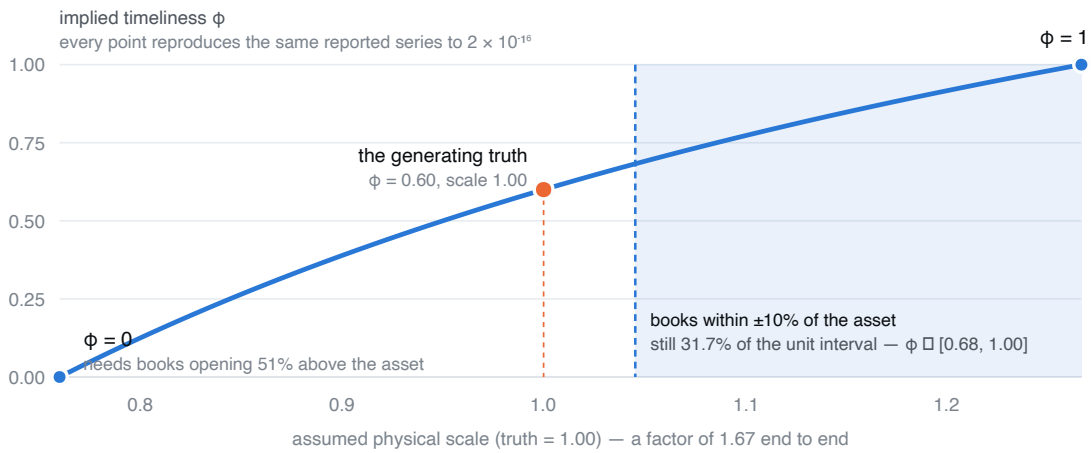

Figure 2 — The identified set. φ against the assumed physical scale, over the range that reproduces one fixed reported series to 2×10^{-16}. The band is flat across the whole sweep: every point on it is exactly as consistent with the filing as every other. The bracketed region is what survives when the opening gap is bounded at ten per cent, still 31.7% of the unit interval. The φ = 0 end requires books opening 51% above the physical asset, which is the state impairment accounting exists to prevent; it is marked, and it is the only thing in the figure that rules anything out.

The family has a closed form, and it is one line. Assume a physical scale c times the truth's, granting both roots. Matching the two amplitudes of the closed form gives two conditions — $c(\alpha - \varphi'\delta) = \alpha - \varphi\delta$ on the D^t amplitude and $c(1 + b') = 1 + b$ on the opening position, writing b for the share by which the books open above the physical asset — so

$$\boldsymbol{\varphi'(c) = [\ \alpha - (\alpha - \varphi\delta)/c\ ] / \delta,\ b'(c) = (1 + b)/c - 1}$$

φ' is a hyperbola in the assumed scale, not a straight line, and it passes through the truth at $c = 1$ by construction. Setting $\varphi' = 0$ and $\varphi' = 1$ gives the endpoints $c_0 = 1 - \varphi\delta/\alpha$ and $c_1 = (\alpha - \varphi\delta)/(\alpha - \delta)$, whose **ratio is exactly $\boldsymbol{\alpha/(\alpha - \delta)}$**: 1.67 at the calibration, and a quantity a reader can evaluate for their own case without simulating anything. The identified set widens as the two rates approach, which is the same corner §7.2 finds worst for the returns repair.

Each member of that family carries its own opening position, and the $\varphi = 0$ end requires books opening **51% above** the physical asset. Bounding that gap at ten per cent (a generous bound) still leaves an identified set covering **31.7%** of the unit interval, which is fatal to a cross-sectional ranking and is not an artefact of an unbounded freedom. **The reported series is consistent with a firm that recognises everything at once and with a firm that recognises nothing.**

The condition under which this bites is worth naming precisely. E_0 is observed when an asset is followed from acquisition. It is *not* observed for a firm-level series, which aggregates assets of many vintages, so the scale that would pin φ is exactly what a cross-sectional study does not have.

What this looks like from inside an optimiser. Fitting the model to a synthetic series recovers φ with a median absolute error of 0.211 when δ is estimated jointly and 0.00073 when δ is pinned at its true value: a **291-fold** difference, with a noise-free series giving 0.211 as well. That is what an exact degeneracy looks like to numerical optimisation. Not a cliff, a canyon.

One channel does break the equivalence, and it is closed. The recognition events' trigger reads gap/E, and across five mirror pairs the event counts differ sharply enough to separate the worlds (16 against 66, 25 against 80, 36 against 133, and two pairs where one side is silent entirely). But **the trigger reads E, which no filing reports.** The information that breaks the degeneracy arrives through a channel the reported series does not have.

---

5 · What a cross-class ranking reads

5.1 · The composite, and the sign of its damage

The class index now earns its keep. What a series identifies is $\varphi_i\delta_i$, so a cross-class ranking computed from reported numbers is a ranking of the product, and a ranking of $\varphi \odot \delta$ is not a ranking of φ unless δ is constant across the classes being ranked.

The model's own measure of how much a class defers sharpens this. The steady-state ratio of unrecognised gap to physical value has a closed form:

$$\mathbf{R}_i = (1 - \varphi_i)\,\delta_i \,/\, (\alpha_i - \delta_i)$$

which simulation reproduces to the transient bound, 2×10^{-4} after 400 periods, against a witness of 1.0 when φ is misstated by 0.1. Across classes with a common α this is a Hadamard product

again, **(1 − φ)** ⊙ **δ**, divided elementwise by (α − δ). **Decay reaches the ranking through two channels, neither of them the parameter of interest.**

R is the model's deferral measure and is not itself an observable: it is a ratio to E, and E is the series nobody reports. Under the mirror it does not shrink or stretch: it **changes sign**, from +0.267 to −1.267, because the mirror world is one in which the books outrun the asset. A quantity that reverses sign under an exchange the data cannot detect is not a quantity a reader recovers. What this section uses R for is what it is good for: computing, inside the model, which way a ranking runs.

> **Timeliness sets the ranking only when durability is constant across the classes being ranked — and the classes accounting standards distinguish are, almost by construction, the classes whose durability differs. That is why the standards distinguish them.**

5.2 · The GAAP ladder inverts

The natural expectation is that a confound of this kind adds noise to a ranking. It does something more specific.

Order the four GAAP classes as this paper's own registration did (property, plant and equipment; finite-lived intangibles; indefinite-lived intangibles; goodwill) and assign the observability shares the registration assumed, falling up the ladder as the standards' willingness to put a class on an amortisation schedule falls. Then assign the decay rates the same standards imply, which fall up the *same* ladder, because a class is placed on a schedule precisely when its decline is predictable enough to schedule. Goodwill sits at the end of both: least observable, and with no degradation schedule at all.

Two of the columns below are evaluated at the recognition rate §9.4 *measures* rather than at the calibration. The **measured** rate is the censored geometric maximum-likelihood estimate on the registered sample, each event carrying the interval from onset of deterioration to charge and right-censored at twenty quarters. The **adverse cut** is that same estimate refit with the 175 events charged one quarter after the peak dropped: aimed at the doubt REG-003 §3.3 registered in advance, which is what makes it adverse.

tier	φ	1 − φ	δ	(1 − φ)δ	**R** at α = 0.05, calibrated	**R** at $\hat{\alpha}$ = 0.408, measured	**R** at $\hat{\alpha}$ = 0.327, adverse cut	R at a common δ
0 · property, plant and equipment	0.80	0.20	0.030	0.00600	**0.3000**	**0.0159**	**0.0202**	0.1333
1 · finite-lived intangibles	0.60	0.40	0.020	0.00800	**0.2667**	**0.0206**	**0.0261**	0.2667
2 · indefinite-lived intangibles	0.40	0.60	0.010	0.00600	**0.1500**	**0.0151**	**0.0189**	0.4000
3 · goodwill	0.20	0.80	0.002	0.00160	**0.0333**	**0.0039**	**0.0049**	0.5333

The right-hand column is the world the design assumed: classes differing in observability and in nothing else. There the deferral measure rises monotonically up the ladder exactly as predicted, Kendall τ = **+1**. The **R** columns are the world the standards describe. There the deferral measure is monotone too: **running the other way.** Kendall τ = **−1** at the calibrated rate and **−0.67** at both the measured rate and the adverse cut, where the first rung alone turns over.

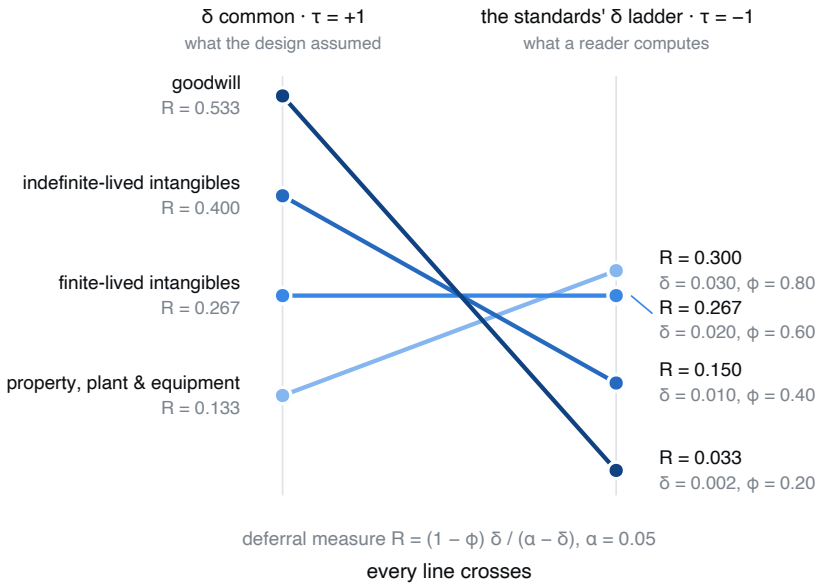

Figure 3 — The design and its own observable are anti-aligned. A slope chart on the four GAAP tiers. *Left axis:* R under a common δ: the world the design assumed, rising up the ladder, τ = +1. *Right axis:* R at the standards' own implied decay rates, falling, τ = −1. Every line crosses. **A confounded design returns noise; this one returns the reverse of what it ordered.**

On the model's own arithmetic, the class predicted to defer most defers least, because it barely deteriorates, and there is little to defer.

5.3 · The condition deciding the direction is a statement about δ

Taking logs of R,

$$\log R = \log(1 - \varphi) + \log \delta - \log(\alpha - \delta)$$

so the deferral measure rises from one tier to the next exactly when

$$\Delta\log(1 - \varphi) + \Delta\log \delta - \Delta\log(\alpha - \delta) > 0$$

The first term is the design. The other two are facts about δ, and on a falling ladder they carry the same sign, so they add. At every rung above, the combined δ contribution (−0.81, −0.98, −1.79) outweighs the design term (+0.69, +0.41, +0.29), and the decomposition predicts the direction of every step. Checking only the log δ term would have got the first rung right for the wrong reason: there the design term is the larger of the two, and the step still falls, **because δ enters twice.**

A researcher cannot establish that a φ-ordered cross-sectional design is sound without already possessing δ — which is the quantity the reported series does not contain.

5.4 · Dispersion loses the ranking; ordering reverses it

Draw 4,000 four-class ladders under the design's own constraint alone (observability falls up the ladder) with δ drawn independently across classes. The deferral measure recovers the registered ordering in **11.5%** of them and exactly reverses it in **1.1%**, mean Kendall τ **+0.32**.

Hold δ common across the four classes instead and redraw: recovery is **100.0%**.

Now impose the standards' falling ladder on the same draw: recovery falls to **1.9%** while exact reversal rises to **23.8%**, mean τ **−0.41**.

So **δ dispersion is what destroys the ranking, and the δ ordering is what turns the wreckage into a reversal.** §5.2's table is what the confound does at one corner of the region; losing the ranking is what it does across the region.

5.5 · The boundary is exact, and it is drawn in quantities the design already declares

Write the design's **budget** as the mean per-rung Δlog(1 − φ), and the ladder's **δ leverage** as the mean per-rung |Δlog δ − Δlog(α − δ)|. Over the same 4,000 draws, the probability that the design fails to recover its ordering, fitted as a logistic in log(leverage / budget), has slope **+1.58** (se 0.081, z = +19.5; the same fit on a permuted outcome returns z = 0.23) and crosses one half at a leverage-to-budget ratio of **0.61**.

> **A φ-ordered cross-section is more likely than not to read what it ordered only while per-rung δ leverage stays under about three fifths of the design budget.**

The GAAP ladder sits at **2.58**.

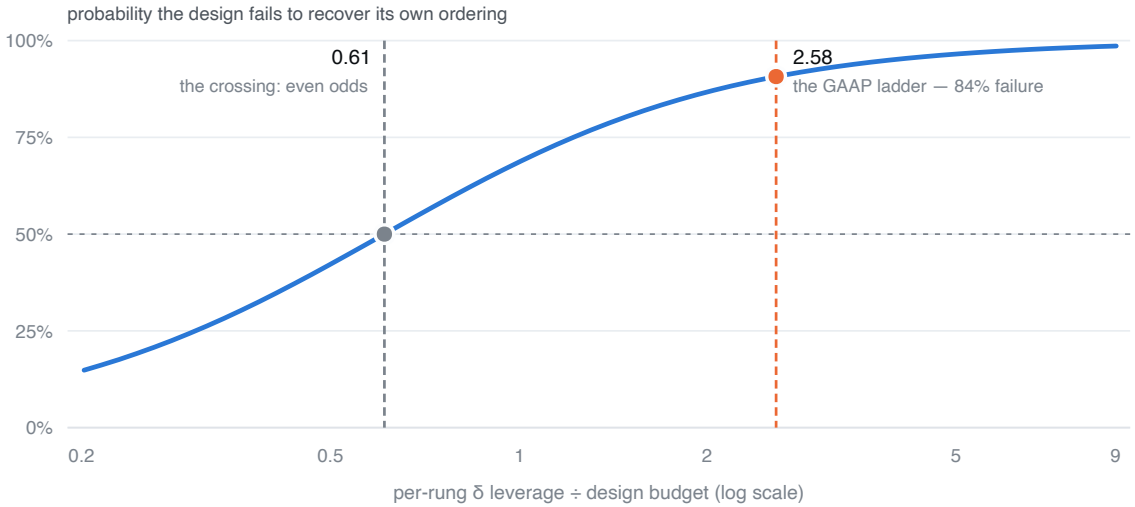

Figure 4 — The validity region, and where the GAAP ladder sits in it. Probability that a φ-ordered design fails to recover its own ordering, against the ratio of per-rung δ leverage to design budget (log x-axis), fitted over 4,000 draws. The crossing at 0.61

is marked; the tabulated GAAP ladder is marked at 2.58, where the failure probability is above four fifths. **This is the figure a reader should take to their own design.** Both quantities are computable before any data are collected.

This is the practical output of the whole paper. A researcher proposing a cross-sectional timeliness comparison can compute both numbers from the design's own stated assumptions, before collecting anything, and read off whether the design is inside the region where it works.

5.6 · Two rungs need no inference at all, and they are the two that break the table

ASC 360 and ASC 350-30-50 require disclosure of useful lives for property and for finite-lived intangibles, and a disclosed life L fixes a write-down rate $1/L$. The first rung falls only when

$$\delta_1 < \alpha\delta_0/(2\alpha - \delta_0)$$

which tends to $\delta_0/2$ as α grows. At the tabulated $\delta_0 = 0.030$ that boundary sits at $\delta_1 =$ **0.0214**, a life of 46.7 years, and the table assigns 0.020: inside by a fourteenth at $\alpha = 0.05$, and *outside* it at the measured rate, where the same boundary is 0.0156. **That is the rung the table's τ turns on, and it turns on the recognition rate's level rather than its shape**: the top three rungs are unchanged at either rate, and §8 puts the shape's contribution to the crossing below it at 0.13%.

Disclosure, however, amortises finite-lived intangibles over materially *shorter* lives than property, so $\delta_1 > \delta_0$ is what a filing actually presents. Measured on the filings themselves (**665 admissible firm-year pairs across 577 firms**, both lives read off one page) the first rung **rises in 65.9%** of them, in both cycles separately (63.3% and 68.2%). Swept uniformly over the rectangle those lives were *assumed* to span (ten to forty years for property, three to twenty for finite-lived intangibles) the same test returns 99.8%, and **86.1% of the disclosed pairs fall outside that rectangle.**

The top rung is a knife edge in its own right. Holding the first three tiers fixed, the deferral measures of goodwill and indefinite-lived intangibles cross at

$$\delta_3^* = K\alpha/(1 + K),\ K = R_2/(1 - \varphi_3)$$

which is **0.0079**, a half-life of eighty-seven periods. Five per cent above it the ladder is no longer monotone and τ moves from -1 to -0.67. The table assigns goodwill 0.002. The exact reversal

therefore needs goodwill to lose half its value no faster than in eighty-seven years, and **the standards do not say that**: a class is left off an amortisation schedule when its decline cannot be *scheduled*, which is a statement about predictability and not about speed.

Unpredictable is not slow, and the difference is measurable in the same direction. Drive a class at $\delta = 0.20$ with probability 0.05 and at zero otherwise (an identical mean decay rate of 0.010, delivered in rare jumps) and its realised deferral is **1.30 times** the closed form evaluated at that mean rate (se 0.002 over 2,000 paths). An unscheduled class defers *more* than a scheduled one of the same average durability, and the lumpy path defers as a smooth class at $\delta =$ **0.0123**, above the crossing rate. Both halves of the inference from an absent schedule push the same way.

5.7 · The binding constraint is the model's domain, not the ordering

R is defined only for $\delta < \alpha$. Past that the deferred gap grows without bound relative to the asset (the ratio reaches 10^{69} by period 400 at $\delta = 0.20$) and there is no steady-state deferral measure to rank.

At the $\alpha = 0.05$ calibrated here, **the entire asserted rectangle lies outside the domain**: every useful life short enough to appear in a filing implies a decay rate at or above the recognition rate. Half of the rectangle is admissible only at $\alpha \approx 0.19$, and all of it above $\alpha = 0.33$.

That made the recognition rate, not the ordering, the quantity to establish first — and §9.4 establishes it. On the registered sample the recognition rate PRE-002's instrument identifies is $\hat{\boldsymbol{\alpha}} =$ **0.408 per year**, 95% interval [0.383, 0.432], on both known biases' inflating side. The calibration used here is low by an order of magnitude, and at the measured rate the asserted rectangle lies *inside* the domain across the whole interval: **0.974** of the 683 disclosed pairs are admissible, and **0.959** at the interval's lower bound. At the adverse cut of 0.327 the rectangle's own fastest disclosed rate of 0.3333 is no longer cleared, and 0.814 of the pairs remain admissible. **The domain restriction is a property of the calibration and not of the disclosure**, and what its remaining margin turns on is the onset bridge rather than the sample.

5.8 · Lag survives what magnitude does not — and cannot be computed

Everything above concerns a *magnitude*, how much a class defers. The registration did not order magnitudes. It ordered **lag**, and the lag statistic does not invert.

tier	lag, standards' ladder	lag, common δ
0 · property, plant and equipment	2	3
1 · finite-lived intangibles	10	10
2 · indefinite-lived intangibles	18	17
3 · goodwill	**36**	22

Monotone under both, and the falling-δ ladder makes the ordering *steeper* rather than flattening it, because the two monotonicities compound: lag falls in φ at every δ, and rises as δ falls at every φ.

Across 400 randomly drawn admissible ladders the lag ordering holds in **100%**, against 1.9% for the magnitude measure. **Part of that margin belongs to the ladder rather than to the statistic.** Drop the durability ordering and draw δ independently, as §5.4 does, and the lag ordering holds in **66.2%** (2,000 ladders, se 0.011) against **11.5%** for magnitude. Lag is the more robust of the two by a factor of six, and it is robust in that ratio rather than in the way a figure of 100% suggests.

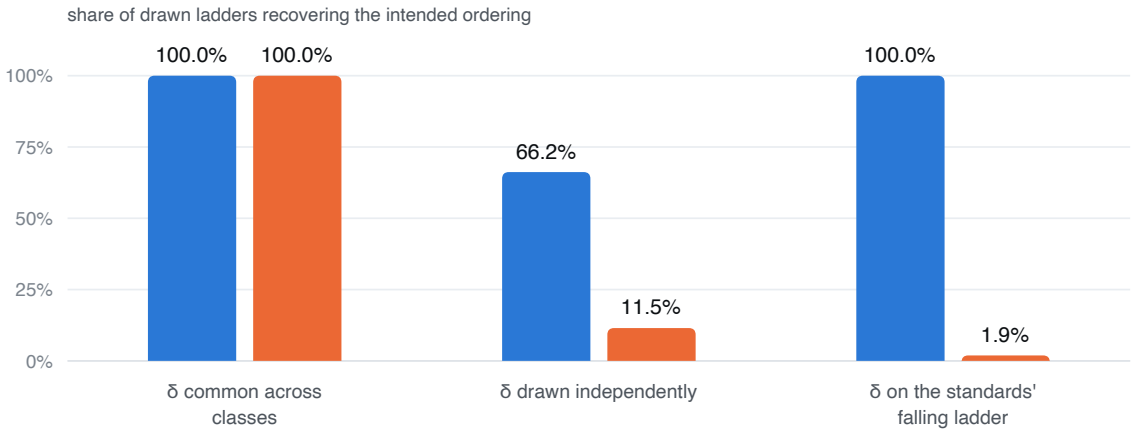

Figure 5 — What survives the confound, and by how much. Recovery rate of the intended ordering for the timing statistic and the magnitude statistic, under three δ regimes: common across classes, drawn independently, and on the standards' falling ladder.

Direct-labelled; no legend needed at two series. The middle group is the honest comparison (66.2% against 11.5%) and the right-hand group is where the GAAP ladder actually sits.

The identification result does not, by itself, wreck a design ordered on lag. Any claim that §9's registered prediction was doomed by the φδ confound is claiming more than the arithmetic gives.

What wrecks it is the next line. **§2's lag statistic is a cross-correlation between ΔE and ΔC, and ΔE is the change in physical value, which no filing reports.** The one statistic the confound spares is the one that cannot be computed from public data. Any empirical instrument must substitute something else (§9's substituted the interval from the onset of a decline in a firm-level signal to the recognition of a charge) and the relation between that substitute and the model's lag has never been written down.

That is a second identification gap, upstream of the first, and it is the one that bit. §10.2 is the discipline it forces.

---

6 · The field's instruments read the same product

The constraint is not local to this model.

Conditional-conservatism measurement estimates how promptly accounting recognises economic losses. The standard instruments — Basu's asymmetric-timeliness coefficient (1997), Khan and Watts's C_Score (2009), Ball and Shivakumar's accrual–cash-flow piecewise measure (2006), Givoly and Hayn's accumulated negative accruals (2000), Bushman and Williams's DELR (2015) — differ in construction and share an input: a reported series, and the requirement to infer a recognition property from it.

If a timeliness parameter reaches a reported series only in product with an asset-life parameter, then a cross-sectional comparison of any of these measures is a comparison of the product. **Two industries with identical recognition practice and different asset lives will score differently, and two industries with identical scores may differ arbitrarily in practice.** The sign of the induced difference is not fixed by the measure; §5.2 shows it can invert a ranking rather than attenuate it.

The literature has already met the confound and read it the other way up. Khan and Watts (2009) report that firms with longer investment cycles score as more conservative, and treat the

association as an economic determinant: a demand for verification that rises with the horizon. Ball, Kothari and Nikolaev (2013) state plainly that firms with shorter asset maturity are expected to exhibit lower timely loss recognition, and read that dependence as the measure behaving *correctly*. Under §5 those are the readings a δ channel would produce whether or not any recognition practice differed at all. **The data cannot separate the two accounts**, which is the whole of the present claim; it may well be both.

One paper has the mechanism itself, in signed form, and it is the nearest accounting-native ancestor this result has. Beaver and Ryan (2005) model unconditional conservatism *preempting* conditional conservatism, and name the channel exactly: "the unconditionally conservative nature of accelerated depreciation creates unrecorded goodwill for tangible assets that preempts conditional conservatism as long as shocks to the market value of those assets are not negative enough to use up that goodwill." A depreciation schedule suppressing measured timeliness: in accounting, in print, in 2005.

What it is *not* is an identification claim. Preemption is a signed comparative static with a stated mechanism, and it leaves the two parameters separately meaningful; the claim here is that the reported series does not contain the difference between them. The lineage is closer than the pharmacokinetics of §3.3 and is owed the same acknowledgement.

The distinction from the standing econometric critiques is what makes this worth stating at all. The objections to the asymmetric-timeliness coefficient (truncation from conditioning on the sign of returns, scale effects, return-variance dependence) are **estimator** problems, and their remedies are the estimator's: controls, fixed effects, interactive corrections, a debiased functional form.

> **A degeneracy is not a bias.** When two parameter vectors generate the identical series, no control recovers the difference between them, because there is nothing in the series to recover it from. The existing repairs are aimed one level up from where the problem is.

The sharpest of those critiques deserves separating from this one precisely, because it shares a title-word with it and almost nothing else. Dutta and Patatoukas (2017) decompose the asymmetric-timeliness coefficient into a component surviving when recognition is symmetric and a component attributable to conservatism, and show the first is positive whenever the return

distribution is skewed, while the second moves with three properties of the news process (expected returns, cash-flow persistence, and the skewness itself) at a fixed degree of conservatism.

Two things make that a different claim. Their confounders are properties of the **news process**; the confounder here is a property of the **asset**, its decay rate, set against a reporting rule, and their firm is a cash-flow stream with no capitalised asset in it to carry one. And their recognition parameter stays recoverable in their own setting, from the spread between bad-news and good-news accrual variances, which is the repair they propose. **A claim that a better statistic can repair is a claim about a statistic.** The claim here is that the reported series is itself invariant, so no statistic computed from it separates the two worlds.

The notation overlaps and should not be allowed to mislead: in Dutta and Patatoukas, δ is the fraction of bad news recognised: this paper's φ. Here δ is the physical decay rate, and has no counterpart in their model.

Three qualifications, stated because the claim is worth stating exactly rather than broadly. First, the mapping from this filter to each of those estimators is not established here: they are not fitting this model, and the composite they read need not be $\varphi\delta$ exactly. Second, the magnitude-versus-timing distinction of §5.8 matters and these measures sit on both sides of it: Basu's coefficient is a slope on returns rather than a delay, and is closer to this paper's magnitude case than its timing one. Third, **the theorem is proved for the reported series alone**, and the returns-based measures condition on a second series, which does break the equivalence. §7 gives that result and its price, and the claim here is correspondingly narrow: what is said is said about a comparison of *reported series*, and a design holding returns is repairing the problem rather than inheriting it.

What the paper claims is that the burden has moved.

> A cross-sectional conservatism ranking now requires an argument that asset life is constant across the compared groups, a correction for it, or an auxiliary series that identifies it, and the ranking most often compared, across GAAP asset classes, is the one where the first of those is least defensible.

The shape of this argument is not new, and its best-known instance is one field over. Fisher and McGowan (1983) argued that an accounting rate of return cannot be used to infer economic profitability, because the reported ratio depends on the depreciation schedule and the firm's growth rate as well as

on the economic return it is supposed to measure: a reporting-rule parameter and an asset-life parameter confounded inside a published number, forty years before this paper, and it detonated an industrial-organisation literature.

Two things about its reception instruct the present one. Their demonstration was numerical rather than a theorem, and the analytical core belongs to earlier work (Kay, 1976). And the sweeping inference they drew (that accounting returns carry almost no information about economic ones) did not survive: it was rebutted on the arithmetic and on the representativeness of the chosen examples (Long and Ravenscraft, 1984), and superseded by a literature recovering conditional usefulness once growth and capitalisation policy are corrected for. **The claim here is deliberately the narrower kind**: an exact equivalence with a stated domain and a repair in §7, rather than a verdict of futility. The ancestor is cited for the shape of its confound; its fate is cited as the reason not to overreach with one.

---

7 · The repair

The way out is visible in the theorem's own statement. The series determines $\varphi\delta$, so **anything that supplies δ from outside the series restores φ.** Two things do, and they are priced very differently.

7.1 · Returns break the equivalence, and the price is a rate

The first repair is the one the field already holds. A market series drawn from the same asset breaks the two-point equivalence immediately and by a wide margin. Under the exchange the mirror firm's asset decays at α rather than δ, so two worlds whose books agree to fourteen decimal places differ in return by $\alpha - \delta$ in *every period*: three percentage points a year, indefinitely. The ambiguity of §3 does not survive contact with a second series drawn from the asset.

It survives the continuum of §4, and the reason is one line. A return is a ratio, and the residual degeneracy is a degeneracy in the unobserved physical *scale*. Grant an analyst both roots exactly (strictly more than returns supply) and the one-parameter family is untouched: φ still sweeps $[0, 1]$ with the reported series reproduced to 2×10^{-16}, and every member emits the *identical* return series, bit for bit. **A scale divides out of a ratio, so no**

quantity of returns data bears on the parameter the scale is concealing.

What breaks the continuum is not the returns but the **news** they carry. The degeneracy is a property of a noiseless economic path: when the asset's value decays geometrically and does nothing else, the reported series has a single geometric driving term and a scale factor absorbs any rescaling of φ. Let the value receive innovations and the realised rate of decline varies period to period; matching the driving term then requires $c\alpha = \alpha$ and $c\varphi' = \varphi$ *simultaneously*, which forces $c = 1$. Regressing the reported series on its own lag, the return-implied path and that path's first difference recovers α, E_0 and φ to 10^{-16} at a return volatility of 0.15, against a design matrix that is exactly singular at zero volatility.

> **So the answer is yes — something outside the reported series does restore φ — and the price is a rate rather than a proof.** Identification does not switch on at the first innovation. It fades in.

Over a twelvefold range of return volatility both the design's collinearity and the standard error on $\hat{\varphi}$ degrade as power laws in σ, with weak-identification bias visible in the mean by $\sigma = 0.025$. **Neither exponent is a constant of the model and neither should be read as one.** Across nine (α, δ) settings spanning the four decay rates §5.2 attributes to the standards, the collinearity exponent runs from −1.07 to −0.38 and the standard error's from −0.78 to −0.09. What holds in all nine is the sign: **identification always degrades as the asset quietens.**

7.2 · Two rates, two different jobs — and one of them changes sign

The decay rate governs how strongly identification *responds* to volatility; the gap between the two rates governs its *level.* These must be kept apart or they read as a contradiction: a sweep at one volatility says the decay rate hardly matters, a sweep across volatilities says it decides everything, and they are statements about different quantities.

Level. Hold δ fixed and sweep $\alpha - \delta$ over a sixteenfold range: the standard error on $\hat{\varphi}$ moves by a factor of **6.8**, as $(\alpha - \delta)^{-0.70}$. Hold $\alpha - \delta$ fixed and sweep δ over a fifteenfold range: it moves by **1.24**, and in the direction that favours *slow* decay, since a slow asset stays alive to be observed.

> **The unreadable case is not the slow asset. It is the firm whose book amortisation rate sits close**

> **to its asset's true rate of decline** — hard for the plainest reason in econometrics: the two numbers being told apart are nearly the same number. At a gap of 0.002 the standard error is 0.13, so φ is readable to ±0.26. The whole interval.

At a realistic amortisation rate the goodwill decay rate is no harder to read than property's, 0.021 against 0.023.

Response. At the fixed gap above, the volatility exponent runs from **−0.39** at a property-like $\delta = 0.030$ to **+0.16** at a goodwill-like $\delta = 0.002$. **A change of sign, not a flattening.** Below roughly $\delta = 0.01$, a noisier asset is read *less* accurately, not more: which is the one place in this paper where the repair runs backwards, and it is precisely the corner the standards decline to schedule.

The sample cannot compensate either. The standard error **never attains the root-T rate at any horizon**: quadrupling the panel from 50 to 200 periods buys a factor of 1.22 where root-T would buy 2.00, and from 400 to 1,600 periods it buys nothing measurable at all. Every term in the estimating equation (signal, regressors and accrual noise alike) is proportional to the asset's remaining value, so once the asset has decayed the later periods are not noisy observations but absent ones.

> **The information about recognition speed is a property of the asset: how much its value moves, and how long it goes on existing. The analyst chooses neither.**

A design cannot buy its way out of either term. The panel saturates within a few half-lives whatever the volatility, and the response to news flattens and then reverses as decay slows, so more years and more news fail *together* rather than in sequence.

7.3 · The design this licenses already exists, and it was run in 2000

Beaver and Ryan (2000) decompose the book-to-market ratio into a persistent **bias** component and a **lag** component by regressing it "on the current and six lagged security returns with fixed firm and time effects," taking the firm effect as bias and the returns-associated portion as lag. The regression is Ryan's (1995) (the book-to-market ratio on current and lagged market-value changes with firm and time effects) and the *bias reading* is not. Ryan's model assumes conservatism away by construction: his assumption (A8) "eliminates the possibility of conservative accounting," and his firm effects enter as a control for what that assumption

leaves unmodelled. Beaver and Ryan supply the reading that turns a lag regression into a two-component decomposition.

That is this section's repair, run empirically twenty-six years ago: a second series used to separate a persistent understatement from a delay, which is exactly the separation §3 shows a reported series alone cannot make. The theorem supplies a warrant the design did not have. The measurements above supply its boundary: the strength of the separation belongs to the asset, and is weakest where the amortisation rate sits near the asset's true rate of decline, a condition a firm effect cannot report.

Two things follow, pointing opposite ways. **The first is a defence of the field's specification arrived at from outside it.** The variation identifying φ in this filter is return variation, which is the same variation Basu's regression requires in order to run at all; an instrument conditioning on returns is drawing on exactly the right information, and the return-variance corrections the literature reached for empirically are operating on the identification-strength parameter rather than on a nuisance. **The second is where that leaves the assets anyone argues about**, and §7.2 has just said: the corner where every term is worst is a quiet asset whose book amortisation rate sits close to its true rate of decline.

7.4 · Disclosed useful lives are the second repair, and they are already published

This one does not require the asset to be noisy, and for most classes the data already exists. It is the accounting form of a move the pharmacokinetic literature has long made: break the tie with information the profile itself does not contain. Kuan, Wright and Duffull (2023) observe that it is precisely "in the absence of intravenous data" that covariates describing elimination can load onto absorption parameters, and their own proposals (a mechanistic model of the two processes, or an estimated cutoff at which the rate constants exchange) share that shape. An outside determination of one root releases the other, and the scale with it.

For two of the four classes, the standards already supply that outside determination. Finite-lived intangibles and depreciable property carry **disclosed useful lives and amortisation schedules**: an estimate of the physical decay rate, made by the firm, audited, published, and *not derived from the series whose timeliness is in question.* Pinning δ rather than estimating it jointly is precisely the 291-fold improvement of §4.

A design that uses disclosed useful lives as an independent δ, and compares timeliness only within a life band, is reading φ rather than $\varphi\delta$ — and it runs on the sample §9 already

collected: **151 property events across 98 firms** on the repaired tier-0 tag list, against 55 across 38 on the list as first collected, with 110 of the 151 joining to a disclosed life.

Three properties recommend it over the design this paper registered. It is **diagonal-safe**: no comparison crosses a class boundary, so §3.1's diagonality assumption is not load-bearing. It holds δ approximately constant by construction, which is the condition §5.5 identifies. And it has a **built-in negative control** (the same comparison *across* life bands, where the theorem says the ranking should degrade), which is the kind of prediction that can embarrass the framework rather than decorate it.

And here is the cost, stated before anyone else states it. Across the one-year life bands the design requires, those 110 events occupy sixteen bands and **exactly one clears the registered floor of 30** (the minimum number of events a life band must carry before a within-band comparison is run) thirty-six events from twenty firms at a five-year life, with none clearing on firms rather than events. Filling the coverage §9's two cycles leave between them raises the join to **133 of the 151** and leaves the same single band clearing: the second band reaches twenty-seven against a floor of thirty. **The design is sound and the sample is not yet large enough to run it.** That is a data-collection problem with a known size, which is a materially better position than the one §9 reports.

The analogy marks its own weak joint, and the joint is real. An intravenous dose is exogenous in the strong sense: a different physical administration of the same compound, whose elimination rate is set by physiology and not by the analyst's question. **A disclosed useful life is chosen by the same management whose timeliness is being measured.** Audited and published is not the same as exogenous, and a firm that reports slowly may also amortise slowly.

Three things bound the consequence: useful lives are anchored by industry convention and by tax and regulatory schedules; they are sticky within a firm across the horizon over which timeliness is measured; and the design can be run on industry-median lives rather than firm-specific ones, at the cost of resolution. **The sign of any residual endogeneity is toward finding *less* timeliness variation than exists, not more**: which is the direction that makes a positive finding harder rather than easier.

7.5 · Where the repair does not reach

The repair does not reach the two unamortised classes, and it fails them for different reasons.

Indefinite-lived intangibles are tested for impairment rather

than amortised (ASC 350-30-35-15), so no life is disclosed and there is nothing to pin δ to. **That is a gap in the evidence, and one a standard could close.**

Goodwill's is not a difficulty of measurement. At $\delta = 0$ the physical layer does not move; $\Delta E = 0$; the term $\varphi \odot \Delta E$ vanishes identically; the gap is **exactly** zero at every φ (not small, zero, at all eleven values swept) and no recognition event occurs at any φ in 400 periods.

> **At zero decay, φ is not ill-conditioned. It is absent from the dynamics.**

That limit is narrower than it looks, and the narrowing matters. The run requires two conditions, not one: $\delta = 0$ *and* an asset whose value does not otherwise move. Set $\delta = 0$ and let the value receive news, and the gap reopens and φ is recovered exactly (to 3×10^{-15}) from the reported series and returns together. **The limit belongs to a motionless asset, not to a slowly-decaying one.** An asset whose value never changes for any reason is not goodwill: impairment testing exists because goodwill's value *does* change, and the standards decline to *schedule* that change, which is not the same as denying it.

So what survives about goodwill specifically is a fact about this model rather than about goodwill. Within the deterministic filter, the class the standards decline to amortise produces no recognition events, so the model has nothing to say about it, and the registered test drew its largest single share from that class. That stands, and §9 pays for it. But the cause is the model's determinism (§11, Limitation 3), not goodwill's nature. A filter admitting stochastic degradation would speak about goodwill as readily as about anything else, and would find it neither the hardest class nor the easiest.

> **Which classes this model can speak about is decided by the δ ladder before any hypothesis about φ is entertained.** None of this was known when the registration was written, and all of it was derivable.

---

8 · The recognition hazard's shape, measured rather than assumed

$R = (1 - \varphi)\delta/(\alpha - \delta)$ is derived by summing a geometric, and a geometric is the one lag distribution whose hazard does not depend on how long the gap has been open. Everything §5 reads

off that expression inherits whatever the assumption was doing. §9.4 measures the shape on the registered sample (a discrete lag distribution fitted to the recognition intervals) and **rejects the constant hazard upward**: discrete Weibull $\hat{k}$ = **1.210**, 95% interval [1.135, 1.285], stable under truncation at eight, twelve and sixteen quarters. This section is what that costs.

8.1 · The closed form survives, as a transform

Let a gap cohort created at time s be recognised after a lag $T \geq 1$ periods. The gap is the flow convolved with the lag's survival function, and with $E(t) = E_0(1-\delta)^t$ and $z = 1/(1-\delta)$,

$$\mathbf{R = (1-\varphi)\,\delta\,\Sigma_{a \geq 1}\, z^a\, P(T \geq a) = (1-\varphi)\cdot(\,\Pi(z) - 1\,)},\ \Pi(z) = \mathbf{E}[z^T]$$

because $\Sigma_{a\geq1}\, z^a P(T \geq a) = z(\Pi(z)-1)/(z-1)$ and $z - 1 = \delta/(1-\delta)$. The generating function is evaluated **outside the unit disc**, so it is a moment generating function rather than a Laplace transform, and that is precisely why it can fail to exist. With T geometric at rate α, $\Pi(z) = \alpha z/(1-(1-\alpha)z)$, which at $z = 1/(1-\delta)$ is $\alpha/(\alpha-\delta)$, and the published form returns exactly.

> **The constant hazard was never doing structural work. It was supplying a closed form for one transform.**

Three checks, each of which could have ended the section. The general form reproduces the published one at a geometric lag to $\mathbf{2 \times 10^{-13}}$. An age-structured simulation — the gap carried as cohorts, each aged one period and multiplied by its own $P(T \geq a+1)/P(T \geq a)$, no closed form anywhere in the loop — reproduces it at the fitted lag distribution to $\mathbf{2 \times 10^{-13}}$ against the 2×10^{-4} transient bound, while **rejecting** the substitution $\alpha \leftarrow 1/\mathbf{E}[T]$ at ten times that bound. And $R(\varphi)/R(0) = (1-\varphi)$ to **zero**, at every φ on a tenth-grid: **φ is a pure scale under age-dependence exactly as it is under memorylessness**, so §2's proportionality result and every ranking statement resting on the $(1-\varphi)$ channel alone are untouched.

8.2 · What the assumption *was* doing is holding up the domain

R is finite exactly when $\Pi(1/(1-\delta))$ is, which is a condition on the **radius of convergence** of the lag's generating function and therefore on its **tail**: not on its mean, and not on any single rate. For a geometric lag that radius is $1/(1-\alpha)$ and the condition

reads $\alpha > \delta$, which is §5.7's domain verbatim. For a lag whose hazard *rises*, the survival function outruns every geometric and the generating function is entire: at $\hat{k} = 1.21$ the transform is finite at every decay rate swept, up to $\delta = 0.80$ per year where it is 7.5×10^{25} and still a number.

The general statement is classical: for distributions with monotone hazard, the transform converges below the limit inferior of the hazard rate and diverges above its limit superior (Barlow, Marshall and Proschan, 1963, Theorem 6.3). A constant hazard puts both limits at α. An increasing one puts the first at infinity, and a *decreasing* one puts it at zero, so a lag distribution with a fattening tail would admit **no steady-state deferral measure at any positive decay rate at all.**

> **The interval [1.135, 1.285] is therefore doing more than rejecting a null. It is what makes the model well-posed across the disclosed range.** Had the same fit returned $\hat{k} < 1$, the closed form would have had no domain to be restricted to.

This does not rescue the asserted rectangle, and the statistic that would look as though it did is withdrawn rather than reported: if the domain is everything, the share of the rectangle inside it is one by construction, and a share of an empty complement is arithmetic rather than evidence. What replaces it is the *level* of R, which is defined everywhere and moves in both directions.

8.3 · The correction is negligible where this paper's ladder sits and material where the filings sit

Against a geometric with the *same mean*, the measured distribution defers **less**, at every decay rate: the direction registered in advance from the standard reliability bound for a distribution new-better-than-used in expectation (Marshall and Proschan, 1972). The size is a different question from the sign:

disclosed life	δ per year	R at the measured shape	R under a constant hazard	overstatement
40 years	0.025	0.0248	0.0249	**0.6%**
20 years	0.050	0.0523	0.0530	**1.2%**
10 years	0.100	0.1180	0.1215	**2.9%**
5 years	0.200	0.3152	0.3445	**9.3%**
3 years	0.333	0.9145	1.3156	**43.9%**

(φ held at 0.5 throughout, since it is a common scale.)

Across §5.2's four-tier ladder, whose fastest rate is 0.030, the worst overstatement is **0.67%**. Across the rates ASC 360 and ASC 350-30-50 disclosure actually spans it reaches **43.9%**, at the three-year life the second of those routinely carries. The mechanism is visible in the transform: z^a with $z > 1$ weights the tail, and the tail is what a constant hazard gets wrong. This is not an artefact of approaching a singularity: the constant-hazard form's pole sits at 0.435 per year, a 2.3-year life, outside the rectangle.

An effective rate exists and it is not a constant. Writing $\alpha\_eff(\delta) = \delta\, \Pi(z)/(\Pi(z) - 1)$ returns the published form verbatim. But α_eff runs from **0.437** per year at a forty-year life to **0.476** at a three-year one: **nine per cent** of itself across the asserted rectangle, in the direction that a faster-decaying class behaves as though recognition were faster. Across the four-tier ladder it moves by six parts in a thousand, which is why the magnitudes there barely move.

> **A recalibration is therefore available and is not a repair.** Any comparative static holding α_eff fixed while moving δ is using the wrong derivative, and a single recognition rate quoted for a cross-section of asset lives misstates one end of it.

The crossing is insensitive to the shape and sensitive to the level. §5.6's $\delta_3^* = K\alpha/(1 + K)$ generalises exactly, to $\Pi(1/(1 - \delta_3^*)) = 1 + K$. REG-004 named the two channels in advance and they oppose: a lower R_2 lowers K and pushes the crossing down, a flatter transform pushes it up. They very nearly cancel: the shape moves δ_3^* by **0.13%**, from 0.00755 to 0.00754. The *level* moves it by 4.3%, and moves the first rung besides, which §5.2's table states.

One thing the shape does not do is break the exchange. An age-dependent world sits 5×10^{-4} from its own constant-hazard match in the reported series (four orders of magnitude below the 3×10^{-1} separating the right conserved quantity from a plausible wrong one in §3.2) and exactly as far from that match's mirror, which is forced rather than discovered. **§3's degeneracy is not repaired by age-dependence.**

8.4 · The shape is identified from a series, and the price of admission is four significant figures

§3 proves an impossibility by counting: four numbers against five parameters. That count was taken under a constant hazard. §8.1 replaced it with an arbitrary lag distribution, which is not one more parameter but an **infinite-dimensional** object, so the count must be taken again.

It leaves a trace, and the trace has a size. REG-005 registered the question (four falsifiers and five ladders) before the code existed. The measurement is made in the *favourable* setting of §3, so a null would have held a fortiori and a positive result carries the condition that the asset is followed from acquisition.

The best constant-hazard world reproduces the measured world's reported series to 3.9 × 10^{-4} per quarter over ten years at a ten-year life, rising to 4.1 × 10^{-3} at a three-year one, with the best mimic an admissible firm at every rate and every φ swept. That number is the whole answer in one figure: **it is the precision a reported series must carry to reject the constant hazard at all.**

precision of the reported series	shapes it cannot separate from $\hat{k}$ = 1.21	width	against the 0.150 width of §9.4's [1.135, 1.285]
10^{-6} per quarter	1.21 alone	0.00	—
10^{-4}	**[1.16, 1.26]**	**0.100**	**0.67 ×**
10^{-3}	[0.60, 1.87]	1.27	8.5 ×
10^{-2}	the whole range swept	≥ 1.40	≥ 9.3 ×

(The lower two rows run into the boundary of the pre-registered sweep, so their widths are lower bounds; on a sweep extended to [0.2, 3.0] the 10^{-3} interval is [0.50, 1.86] and the reading is unchanged. The search's own floor at the true shape is 2.7 × 10^{-8}, thirty-seven times below the finest tolerance reported, so the top row measures the model and not the optimiser.)

At one part in ten thousand the reported series is a better instrument for the shape than the event dates are — 0.100 against 0.150, from hand-collected impairment lags. At one part in a thousand it is an order of magnitude worse. The identification is real and it is expensive.

And what lies inside the interval matters more than how wide it is. At one part in a thousand the registered sweep's set reaches **k = 0.60**: a *decreasing* hazard, which §8.2 says admits no steady-state deferral measure at any positive decay rate. **A series matched to a tenth of a per cent per quarter cannot separate the world in which this model is well-posed from one in which it has no steady state at all.** That is a sharper limit than any width.

A longer series does not help, and the reason is where the information sits. The interval is not monotone in the observation window: 1.40 over five years, 1.26 over ten, **0.98 over twenty**, and 1.32 over a hundred. Once the gap reaches its steady state each further quarter repeats a single number, so extending the window adds redundancy to a mean and dilutes the transient

carrying the shape. This is an identification property rather than a sample-size one: the general result for any finite-dimensional approximation to a lag space, whose approximating families are meagre in it and whose approximation error "cannot, in other words, be made asymptotically negligible" (Sims, 1971, §5), said of exactly the rational lag distributions of which a constant hazard is the lowest-order member (Jorgenson, 1966).

Three recognition rates now live in this paper and they are three different quantities. The series-matching constant is nearly flat across the rectangle at 0.438 per year: the reciprocal mean lag $1/\mathbf{E}[T] = 0.435$ to within a per cent; α_eff rises with the decay rate because the transform weights the tail; and $\hat{\alpha} = 0.408$ is the geometric maximum-likelihood summary of the event dates. They agree to **five parts in ten thousand** at a twenty-year life and part company at a three-year one, where they differ by **15%** and move in opposite directions from $\hat{\alpha}$. Least squares on the series matches the mean, the transform matches the tail, and the likelihood matches the event dates: three functionals of one distribution, with no obligation to coincide.

> **§3's exchange survives into all of this, and it is forced rather than discovered.** The mimic search returns the mirror pair at an identical objective, which it must. What the mirror costs is the mimic's own parameters: at a forty-year life the set of worlds fitting within one part in a million of the best spans **0.128 in each root and 0.577 in φ**. The best-fitting constant hazard is not a world. It is a pair, and φ inside it is as free as §4 says it is.

Two predictions registered in advance were wrong, in the same direction. REG-005 predicted the shape would be invisible below one part in a thousand and that the shape interval would be a hundred times the event-date interval, reasoning from the density of rational lag distributions in lag space and from the classical ill-conditioning of exponential sums (Lanczos, 1956, as quantified by Varah, 1982). The measured residue is four times larger than that reading allows and the interval is nine times, not a hundred. **Approximation theory describes what a family can do in the limit; this is one distribution over one horizon, and it leaves more behind than the general argument suggests.**

---

9 · The severe test: registered, run twice, and lost

What §§3–8 bear on here, stated before the result so it cannot be read as an excuse afterwards. The test below ordered asset classes by expected timeliness. §5.2 shows a *magnitude* reading of such a ladder is inverted by the composite; §5.8 shows the *timing* reading (the one this registration used) is not. **So the identification result does not explain this failure, and it is not offered as though it did.** It bears on this section at one point: the model's lag is defined against a physical series no filing reports, so the instrument below measures a substitute. **"Severe" is Mayo's (1996) word and is used in her sense** (a test that would very probably have caught the prediction being wrong had it been wrong) and what makes this one severe rather than merely early is the registration discipline of Nosek, Ebersole, DeHaven and Mellor (2018): the prediction, the instrument and the falsifiers were committed and pushed before the data were touched.

9.1 · What was predicted, and when

The framework's sharpest available prediction follows directly from §2:

> **Recognition lag scales with the unobservability of degradation.**

To test it, unobservability must be identified with something measurable. The identification chosen was **GAAP asset class**, on the reasoning that the categories accounting standards decline to place on an amortisation schedule are precisely the categories whose degradation is hardest to observe. The classes are those of the FASB *Accounting Standards Codification*: Topic 360 for property, plant and equipment, Topic 350 for intangibles and goodwill. That yields a four-tier ordering predicted in advance to be monotone in lag, shortest at property and longest at goodwill.

Every event in this test is a recognised impairment, which places the sample on the *boundary* of §2's domain restriction rather than inside its complement: a charge is the moment degradation became estimable, so §2 governs the accumulation that precedes it and the event marks where that accumulation ends: which is why an interval is measurable on these events and on no others.

The registration preceded the data. PRE-001 was committed **alone**, at commit 9722342, and pushed, before any lag was computed; the analysis code did not yet exist. The git history is the timestamp, and it is the entire evidence that the prediction preceded the outcome: which is why the registration and the code were deliberately not batched into one commit.

A disclosure, because that discipline does not fully cover the test reported here. PRE-001 timestamps the *prediction*. The result in §9.3 comes from PRE-002, which specifies a different *instrument*, and PRE-002's registration shipped in the same commit (d655501) as the implementation of that instrument. The result itself came later, so the git history still establishes that PRE-002 was registered before its outcome existed. What it does **not** establish is that the instrument's details (the onset rule's window, the tie-break direction, the materiality floor) were fixed before anyone had seen what they produced. **That is precisely where the remaining researcher degrees of freedom live.** No claim is made that they were exploited, and the author's account is that they were not; the point is that this is an *account* rather than a demonstration, whereas for PRE-001 it is a demonstration. A reader is entitled to weight the two differently.

9.2 · The first instrument failed, and so did its diagnosis

The first test (PRE-001) returned a null in both universes, and the two nulls did not agree with each other. The pilot retained 120 events across 72 firms and gave Jonckheere–Terpstra **z = −0.177** (one-sided p = 0.570), with goodwill's median lag sitting *below* property's: 4.0 quarters against 5.0, the reverse of the predicted ordering. The replication universe, declared in PRE-001 §4.2 before the pilot was run, retained 202 events across 106 firms and gave **z = +0.634** (p = 0.263), with goodwill and property tied at 3.0. A weak negative and a weak positive, neither significant: the signature of a measurement carrying no information about the ordering rather than of an effect in either direction.

The instrument was then examined and it had a defect. Across the 322 events retained by both universes there was **zero right-censoring**, 69% of pilot lags fell at six quarters or fewer, and **1,047 charges were discarded for having no measurable onset.** An onset rule requiring an unbroken decline in a firm-level signal measures the volatility of that signal, not the phenomenon: it can only find an onset when the signal happens to fall monotonically, which is common over short windows and vanishingly rare over long ones. **The lag distribution was pinned against a ceiling the instrument itself imposed.**

This diagnosis was not permitted to rescue the result. A second, separately numbered registration (PRE-002) was written with a different onset instrument (peak-to-charge), a label-permutation negative control, a power curve to be reported whatever happened, the significance level tightened to 0.025 for the second look, and (decisively) an explicit **stopping rule** stating in

advance that there would be no third instrument.

9.3 · The second instrument worked, and the prediction failed anyway

Pilot — retail trade (SIC 5200–5999). 244 events across 121 firms.

tier	n	median lag (quarters)	IQR	mean
0 · property, plant and equipment	21	5.0	3.0–9.0	7.05
1 · finite-lived intangibles	34	4.0	1.0–8.0	5.71
2 · indefinite-lived intangibles	34	5.5	1.2–9.0	6.12
3 · goodwill	155	5.0	1.0–9.0	5.93

Jonckheere–Terpstra **z = −0.290**, permutation p = **0.590**. median(t_3) − median(t_0) = **0.0 quarters**, CI [−4.0, +2.0].

Replication — computer and data processing services (SIC 7370–7379). 444 events across 190 firms.

tier	n	median lag	IQR	mean
0 · property, plant and equipment	34	5.0	1.2–9.8	6.62
1 · finite-lived intangibles	102	4.5	1.2–10.0	6.12
2 · indefinite-lived intangibles	46	6.0	2.2–11.0	6.85
3 · goodwill	262	5.0	1.0–10.0	6.46

Jonckheere–Terpstra **z = −0.095**, permutation p = **0.520**. median(t_3) − median(t_0) = **0.0 quarters**, CI [−4.0, +2.5].

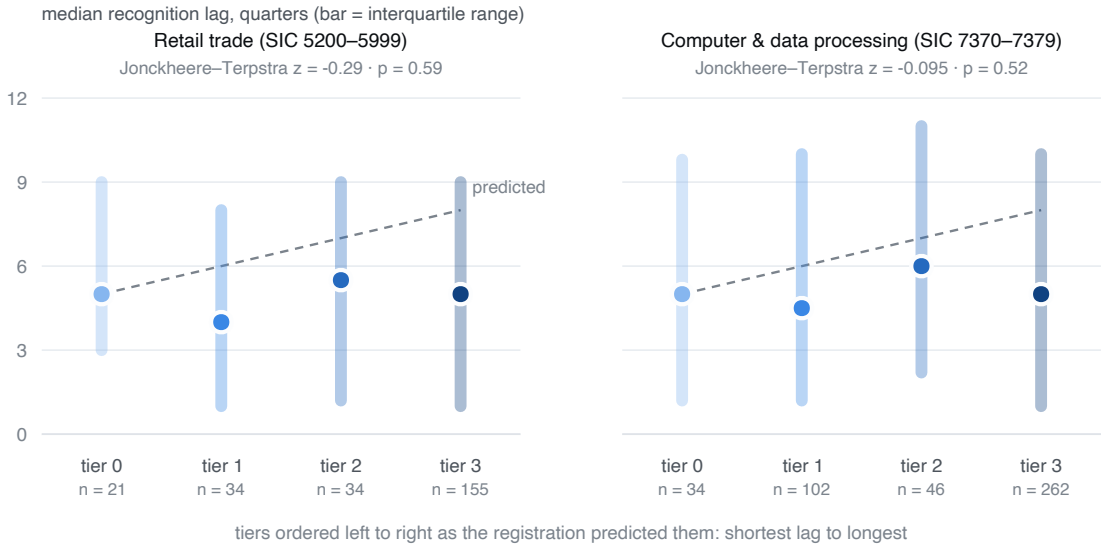

Figure 6 — The gradient that was not there. Median recognition lag by tier with interquartile range, both universes side by side, ordered left to right as the registration predicted them. The predicted line is drawn as a reference. Tiers use the ordinal blue ramp, not the categorical palette, because they are a ladder. Both panels are flat; both z-statistics are negative. **The design would have detected a one-quarter-per-tier gradient with probability 0.95 in retail and 1.00 in computer services.**

The instrument was demonstrably better this time, and that is what makes the null bite:

	PRE-001 (streak onset)	PRE-002 (peak onset)
events retained, both universes	322	**688**
charges discarded for no onset	**1,047**	**0**
right-censored	0%	7.8% pilot, 14.2% replication
pilot IQR width, tier 3	3.0–7.0	1.0–9.0

The lag distribution now spans the registered range instead of piling against a ceiling. **And the ceiling is not where the gradient went:** goodwill, the tier predicted to lag longest, carries the *least* censoring of the four in the replication and the second-least in the pilot, so the twenty-quarter cap is not hiding its tail.

Negative control. Tier labels permuted 1,000 times with the lag distribution held fixed: the null z is +0.007 (sd 1.025) in the pilot and −0.002 (sd 1.000) in the replication. **The pipeline does not manufacture a gradient**, and the empirical p-values do not lean on a normal approximation, so they are untroubled by tier sizes of 21/34/34/155.

Power, reported because a null without its detectability attached is not a result: 0.65 / 0.87 at half a quarter per tier, **0.95 / 1.00 at one quarter per tier**, 1.00 / 1.00 at two.

Three qualifications bound that sentence, and they are the paper's own rather than a referee's. The power simulation assumes the onset instrument measures lag without error (it resamples the observed lag distribution and adds a noiseless per-tier shift) so any measurement error in the onset attenuates a true gradient and the true power is lower by an unquantified amount. The 688 events come from 311 firms and every statistic here treats events as independent, so the effective sample is smaller and the reported power is an upper bound. And **one quarter per tier was never derived from the model**: it is a plausible round number,

and it could not have been derived, because deriving one requires the very φ-to-tier bridge §10.2 concludes was unsound.

What the result supports, at the strength the evidence carries: a well-powered, pre-registered, replicated test found no gradient, and if a gradient of the size tested exists it is unlikely to have been missed twice. **What it does not support** is that the framework's own predicted effect has been ruled out, because the framework never named that effect in the units the test measured. Calling this *evidence of absence* would contradict §10.2 in the same paper: the bridge cannot be too broken to license a prediction and sound enough to license its refutation.

The shape of the failure, as sharply as the data allow. Both z-statistics are negative, so the point estimates ran *opposite* to the predicted ordering in both universes, as they had in the PRE-001 pilot. And the ladder does not merely fail to be monotone: it is wrong in a specific and instructive place: **tier 2, indefinite-lived intangibles, carries the longest median lag in both universes, with goodwill below it** (5.5 against 5.0; 6.0 against 5.0). The predicted ordering appears in **5.7%** of firm-clustered resamples in the pilot and **5.8%** in the replication.

The stopping rule fired. There is no third instrument. A hypothesis requiring one on the same data is a hypothesis being fitted.

9.4 · The same sample answers two questions it was not collected for

The stopping rule bars a third instrument for the *lag gradient*. It does not bar asking the sample questions it was never asked. Two were registered in REG-003, committed and pushed before the instrument existed. **Neither is a re-test of §9.1's prediction and neither may be read as one.**

The sample rebuilt at 99.0% agreement. `companyfacts` serves each firm's latest view of its own history, so a re-pull is not the original pull. Rebuilt: 695 events across 313 firms against 688 across 311, three of four tier counts identical in the pilot, censoring at 7.7% against 7.8%. The registered reconciliation rule (95% agreement in total n with no tier moving by more than 20%) admits this at 99.0% and 1.4%.

The peak-to-charge recognition rate is 0.41 per year, and the calibration was low by an order of magnitude. Each event carries the interval from onset to charge, right-censored at twenty quarters: which is α's definition, measured once per event, by an instrument built to look at something else. The censored geometric maximum likelihood estimate is $\hat{\alpha} = 0.1227$ **per quarter (se 0.0046), 0.408 per year, 95% interval [0.383, 0.432]**. Retail

gives 0.433 and computer services 0.394. The three sensitivities registered with PRE-002 give 0.397, 0.499, 0.413; administrative censoring at eight, twelve and sixteen quarters gives 0.396, 0.398, 0.404. **Every cut lands in the same regime**, and the calibrated 0.05 is outside the interval of all of them.

Two biases push this up and one pushes it down; the direction of each was registered before the number. A gap that opened and was never recognised leaves no filing, so conditioning on a charge over-represents short intervals. If revenue peaks after economic value has turned, the measured interval is short of the true one. Against those, the sample contains no lag of zero, so fitting on a support including it understates $\hat{\alpha}$. **The one unregistered cut removing the mass where the onset bridge is least credible (the 175 events charged one quarter after the peak) gives 0.327, still an order of magnitude above the calibration.** The result does not rest on its most suspect quarter.

The shape was fitted rather than assumed. A discrete Weibull in Nakagawa and Osaki's (1975) parameterisation (whose hazard is increasing exactly when its shape parameter exceeds one, so the nesting is a boundary case and not a coincidence of fit) gives **$\hat{k}$ = 1.210, 95% profile interval [1.135, 1.285]**, excluding the constant hazard the model assumes. The non-parametric hazard shows why: a quarter of the sample is recognised one quarter after the peak, and the rest faces a hazard rising from 0.09 to about 0.25 over the following five years. **The longer a gap has been open, the likelier it is to close**: the opposite of the memorylessness a single α encodes. §8 is what that costs the model.

9.5 · And the reporting layer is not diagonal

§3.1's diagonality assumption predicts that recognition in one class does not force recognition in another, so events should be independent across classes within a firm-quarter. Taking each firm's per-class impairment frequency as given and redrawing which quarters they land in (10,000 draws within each firm's own eligible-quarter set) firm-quarters carrying two or more classes are:

universe	observed	null mean	central 95%	observed/expected	two-sided p
retail	**30**	7.3	[3, 12]	**4.12×**	0.0002
computer services	**44**	21.8	[15, 29]	**2.02×**	0.0002

Both universes, same direction, at the resolution 10,000 draws

can report; the design detects an injected excess of five per cent of events with probability 1.00. **The assumption is false, it is false in both sectors, and §3.1 said in advance that it would be.** The strongest pairwise coupling is goodwill with indefinite-lived intangibles in retail (5.86×) and goodwill with finite-lived intangibles in computer services (2.22×), and it is the intangible-with-goodwill cells that replicate across both sectors, all four surviving Holm correction.

The mechanical reading has to be excluded before the economic one is available, and it is two readings rather than one. The ordering is imposed by ASC 350-20-35-31, which requires any other asset or asset group of a reporting unit to be tested before goodwill, and by 35-32, which extends that to every asset tested. On one channel the rule creates joint *testing*: ASC 350-20-35-3C(f) names testing a significant asset group for recoverability as an event requiring an interim goodwill test, so one trigger fires two tests. On the other it suppresses joint *recognition*: the other charge is recognised first and reduces the reporting unit's carrying amount, and the goodwill charge is the excess of that amount over fair value, so the prior charge is subtracted one for one. **Under the ordering alone the two charges are substitutes at the margin, and this sample shows them as complements.**

Signing the net requires the two charges at the reporting-unit level, which US filings do not disclose. REG-006 registered an entity-level test of the suppressing channel and **it failed as registered**: no consistent sign in either sector. The natural fallback is the triggering *disclosure*, and REG-007 measures why that route does not identify either: ASC 350-20-50-2(a) compels a description only *for each goodwill impairment loss recognized*, so a triggering-event population assembled without that restriction is selected on the outcome under study. REG-008 sharpens the instrument to the sentence and the *named* reporting unit; the separation from the placebo more than doubles (0.103 against 0.030) while the joint-versus-goodwill-only difference stays at +0.014 (p 0.60) in a design that could have detected 0.068. **The reason is countable: not one of the 281 firm-years taking both charges writes a sentence naming a reporting unit, a trigger, and any of the standard's own (f)-family language.** The disclosure does not carry the quantity the decomposition needs.

This is a finding and not a caveat. §3.1's diagonality is a modelling convenience that the world declines to honour, in both sectors, by factors of two and four. Any successor to this model owes an off-diagonal term, and §7.4's within-life-band design is recommended partly because it does not need one.

10 · What may now be claimed, and what may not

10.1 · The demotion, stated exactly

Not supported: that recognition lag scales with the unobservability of degradation, where unobservability is identified with GAAP asset class, in US-listed retail trade (SIC 5200–5999) or computer and data processing services (SIC 7370–7379), among registrants filing in 2013–2024, on charges recognised 2012 Q2 – 2026 Q2, at the firm level, at effect sizes of one quarter per tier or larger.

Demoted: the lag-scaling claim moves from *a prediction the framework makes* to *a prediction the framework made, at one level of aggregation, with one bridge, and lost.* Any surviving version must state its measurable and its bridge before it is tested again.

Unaffected: every result in §§2–8 and Appendix A. Those are properties of a stated model, established by proof and by simulation and held in place by a test suite. A model result is not made false by the failure of an empirical identification, and it is not made true by one either.

And that last line is a problem, not a reassurance. If nothing in §§2–8 was at risk, then nothing in §§2–8 was on test. The accounting is therefore:

- **What was at risk and lost:** the conjunction of the model, the bridge, and the firm-level unit of observation. That conjunction was the framework's *entire empirical content* as of §9. It is gone.
- **What was never at risk:** everything in §§2–8, because those are theorems and simulations. Calling them "unaffected" states a fact about their logical type, not a survival.
- **What follows: this framework currently has no confirmed empirical claim.** It has a theorem with derived consequences, two auxiliary measurements that landed (§9.4, §9.5), one registered prediction that failed, and a stated method for building the next one.

The framework would have been *confirmed* had the gradient appeared. It did not, and the correct posture is not that the theory survived but that **the theory has not yet been given a test it can pass.** Designing one is §7.4's business, and it is unfinished work rather than a conclusion.

Three post-hoc conjectures about where the conjunction broke are recorded in the repository and excluded from this paper's argument deliberately: each arrived after the number,

none is evidence for anything, and any that is ever tested must be registered from scratch. One is worth naming here only because it generalises into a discipline rather than a defence.

10.2 · The bridge discipline

The registration contained a tier table and no **bridge proposition**. It never wrote down, as a deniable claim, the sentence connecting the model to the world:

> *φ, the observability of degradation, is identified with the presence or absence of a GAAP amortisation schedule, because…*

Had that sentence been written, its weakness would have been visible before the data were touched. The observability of *degradation* and the observability of the *accounting treatment* are different quantities, and they may even be anti-correlated: goodwill carries no schedule, but its impairment is triggered by conspicuously public signals: a share-price fall, a missed segment, a lost contract. The physical condition of a distribution centre carries a schedule and is visible to essentially nobody outside the firm.

> **A quantity in the model was matched to a quantity in the world that shares its name and not its meaning.**

The lesson is not about accounting. **A bridge from a parameter to a measurable must itself be stated as a proposition and checked**, and this programme now requires it of every registration.

10.3 · The comparison this paper is not entitled to make

This paper is not entitled to the argument that the framework has escaped the trap that closed on Odum's emergy programme: that emergy's fatal defect was making no risky predictions, and that losing a registered bet therefore counts as a kind of methodological success.

That argument is withdrawn, on three counts a sceptical reader would have reached first. It selects its own reference class: introduce a comparator that made no predictions and any loss becomes a comparative victory. It is an assessment of the author's conduct rather than of the world. And it arrives at the end of the section reporting the failure, so it would be the last thing a reader carried away. **A paper does not get to grade its own integrity.**

What remains after the withdrawal is a fact and not an evaluation. A prediction was registered before the data were seen, it was tested, and it failed. Whether that is worth anything is a judgement this paper leaves to whoever is reading it.

---

11 · Limitations

1. **The severe test failed and this paper does not know why.** Three post-hoc explanations exist (the theory is wrong; the bridge was wrong; the unit of observation was wrong) and **the data do not distinguish them.**
2. **The unit mismatch is real and unfixed.** The impairment charge is asset-level; the deterioration signal used was firm-level. A firm can impair a failing reporting unit while consolidated revenue rises. Fixing this requires segment-level disclosures (ASC 280) and is a different project with a different registration, which **may not cite the present failure as support for anything.**
3. **The filter is deterministic and single-firm.** No stochastic degradation, no heterogeneity, no interaction between firms, no market. **The determinism is not innocuous, and §7.5 is where it bites**: the goodwill limit reported there is a consequence of the physical layer being noiseless and dissolves once that layer is allowed to move for reasons other than a schedule. Admitting stochastic degradation is the single change most likely to alter what this model says, which is why it is named here rather than in a list of extensions.
4. **φ and θ are swept rather than measured — and for φ, §3 says why no one can do better from a series.** α *is* measured, but for the quantity PRE-002's instrument dates rather than for the model's α. The bridge from that rate to the model's α is exactly the proposition §10.2 requires of every registration, and this paper has not written it. §8 settles what the shape rejection costs; the domain restriction it was holding up is the surprise.
5. **The diagonality of the reporting layer is an assumption, it was testable, and it is false** (§9.5). The Hadamard form is an approximation whose error is now measured, and §9's treatment of events as independent draws overstates their information content by a factor this paper can state. What the design cannot do is separate an economic coupling from the sequencing the standards impose. The disclosure

route that would separate them is closed twice over, by a selection argument and by a count.

6. **Appendix A's coupling Λ^{-1} and the SDG 7.3.1 series are the same quantity dimensionally, not empirically.** The correspondence licenses "this dimension is one institutions already report," not "this series measures the model."
7. **A Duhem–Quine problem is present and is narrower than the usual invocation.** A failed test here cannot distinguish a false theory from a bad observability proxy: genuine, and §9 is an instance. It should not be confused with heterogeneity in decay rates across industries, which is an ordinary identification problem with ordinary remedies and not a philosophical one.
8. **The framework claims necessary conditions, not uniqueness.** Any adequate account must distinguish physical from claim components, must let the second lag the first asymmetrically, and must make the residue accumulate. This construction satisfies those conditions; it is not argued to be the only one that does.

---

12 · Relation to existing work

Most of the placement in this paper is made where the results are, because a result and its ancestor belong on the same page. What follows is what does not fit there.

On the identification result. The mathematics is Bateman's (1910); the phenomenon is pharmacokinetics' flip-flop (Garrett, 1994; Kuan, Wright and Duffull, 2023); the general framework is Bellman and Åström's (1970). The nearest economic instance is Nerlove's (1958), and the nearest accounting-native ancestor is Beaver and Ryan (2005), whose preemption channel is this mechanism in signed comparative-static form (§6). The shape of the argument (a reporting-rule parameter confounded with an asset-life parameter inside a published number) is Fisher and McGowan's (1983), whose fate is cited as the reason not to overreach with one.

On the reporting layer as a filter. The qualitative claim that historic-cost accounting relocates volatility across time rather than suppressing it is Bleck and Liu's (2007), and it is not claimed here. What §2 adds is the parameterisation: the same claim indexed by a continuous observability parameter, with smoothing and concentration measured separately. **That table's headline result is nineteen years old.**

On recognition delay. Goodwill write-offs are known to lag economic impairment by three to four years on average, with the delay extending to ten for a third of firms (Hayn and Hughes, 2006); impairment timing has been modelled as a first-event hazard with covariates (Potepa and Thomas, 2023); conditional conservatism is measured throughout as a timeliness coefficient. **In none of it is the recognition lag's *shape* estimated rather than assumed.** §9.4 estimates it and §8 is what the estimate costs the model that motivated it.

On the composition claim underneath the filter. Soddy is the origin, and P1 is his observation made axiomatic; the present contribution is not the distinction between physical wealth and claims on it but its *dynamics*. **Georgescu-Roegen is a hostile witness inside this framework's own bibliography**: the most-cited authority in the tradition this work draws on, and an explicit refuser of the physical-to-monetary reduction a naive reading of Appendix A's Λ proposes. That refusal is not answered here. It is *adopted*: measuring the wedge between physical throughput and financial claims does not assert that energy determines value.

On stock-flow-consistent modelling. The two-layer structure is not new as a structure: it is Godley and Lavoie's (2007), where every flow has a source and every stock a counterpart and the accounting closes by construction. What is new is treating the wedge *between* the layers as an *information* quantity with a release rate, and proving what that structure does to identification.

---

13 · Data and code availability

Every number in this paper is produced by committed code in a public repository, and every table above names the script that generates it. The reproduction is two commands; the six quantities not printed by either are enumerated in the repository's `REPRODUCTION.md` along with how to obtain them.

The registered materials are `docs/preregistration/PRE-001`, `PRE-002` and `REG-003` through `REG-008`, each committed and pushed before the instrument it governs existed, save the disclosure in §9.1. Raw EDGAR-derived event data is committed at `data/pre-002-events.json`.

A caveat on reproduction that is a property of the source rather than of this repository. SEC `companyfacts` serves each filer's *latest* view of its own history, so a re-pull is not the original pull. §9.4 reports the rebuild at 99.0% agreement against a reconciliation rule fixed in advance. A reader rebuilding in 2030 should

expect a comparable drift and should apply the same rule rather than expecting identity.

Appendix A · The framework the filter was built inside

Three propositions about the composition of wealth, and the coupling they oblige. This material motivated the filter and states the domain within which §2's two layers are the right two layers. It is an appendix, at this length, because ***nothing in §§2–9 depends on it.*** *The identification result holds for any two-layer filter of the stated form, whatever one believes about the composition of wealth, and a result that needs a metaphysics is weaker than one that does not. The full development, including the three-leg defence of the coupling Λ and the invariance evidence, is in the repository at* `docs/appendix/A-framework.md`.

A.1 · What a first principle is, and what it is not

An appeal to "first principles" is worthless without a definition of the term, and the gap is a **type error** rather than a wording problem. **An axiom is a proposition** (truth-apt, deniable, the kind of thing that can be false. **A model is a structure**) it has interpretations, not a truth value. A structure cannot be promoted to a proposition by describing it more emphatically. The tensor is not the axiom; the axiom is the proposition that wealth has the structure the tensor formalises.

And **"undeniable" must go.** An axiom nobody can deny is a definition, and definitions generate no empirical content. The useful notion is computing's: an **invariant**: never proved undeniable, proved *preserved*, within a **stated domain**.

The test separating a first principle from a result: **denying a first principle produces a different science; denying a result produces a wrong number.**

A.2 · The three propositions

P1 · Composition. Every unit of wealth is a compound of a physical component and a claim component, obeying different laws: thermodynamic and arithmetic respectively. *Domain:* units of wealth having any physical referent. Silent on purely contractual objects whose referent is another claim.

P2 · Decay. The physical component degrades absent maintenance. No store is inert. *Domain:* physical

referents over horizons long relative to their maintenance cycle. Silent on the short run, where degradation is negligible against measurement noise.

P3 · Atomism. Measured aggregates are folds over units. No aggregate is more fundamental than its constituents. *Domain:* any measurement presented as a property of an economy rather than of a population.

Three, not ten, and each stated so a competent economist can say *no* to it and mean something specific. **P3 is where this framework's commitment actually bites**, and the strawman version must be avoided: neoclassical economics does *not* deny physical depreciation: δK appears in every growth model from Solow forward. What P3 puts at issue is whether an aggregate can be treated as a primitive carrying its own laws.

P1 concerns composition and is silent about time. P2 concerns time and presupposes only that a physical component exists, a compound whose physical component were inert would satisfy P1 and violate P2. P3 concerns the relation between measurements at different scales; an economy of inert single-component units would satisfy P3 while violating P1 and P2. **No proposition is derivable from the others.**

A.3 · The propositions are deniable, and the repository proves it for one

P2 fails at complete maintenance. Effective decay is the entropy rate *net of maintenance*, so a fully maintained asset has no dynamics at all and the model collapses to an identity. That regime is reachable by setting one parameter.

And the framework's central mechanism switches off at φ = 1. Perfect observability annihilates the entire phenomenon: lag 0, deferred information 0.0, zero recognition events. Note what this does *not* say: P2 still holds at φ = 1 and the physical layer still decays, from $E_0 = 100$ to 0.031 over 400 periods. **What vanishes is the gap, and therefore everything this paper is about.**

These are not embarrassments to be hidden behind a stronger word. A switch-off regime demonstrates that the framework's subject matter is contingent on a quantity that could take another value, **which is the minimum a claim must satisfy to be empirical rather than definitional.** It is not itself a refutation and none is offered here.

*A companion result is cited rather than reproduced: **Paper II** of this programme reports that a levy whose base cannot observe an*

accrual is inert regardless of its rate. ***The identification of the two mechanisms that the shared theme invites is withdrawn.*** *Put to a cross-scale check it does not hold: what this filter fails to recognise is deferred, held in the gap and released at rate α, while what a levy's base fails to recognise is never assessed at all, and a levy has no parameter playing α's part. The two results share the question and not the operator. The check, and the fact that the withdrawal was written down before it was run, are recorded at* `docs/RESULT-END-TO-END-001-E1.md`.

A.4 · The boundary on P3, and it is a fifty-year-old theorem

P3 is the proposition a competent economist attacks, and the attack has a name. Sonnenschein (1972, 1973), Mantel (1974) and Debreu (1974) proved that aggregate excess demand inherits from individually rational agents only continuity, homogeneity of degree zero and Walras's Law: not downward slope, not uniqueness, not stability. Aggregate demand can take essentially arbitrary shape. **The best-established result about aggregation in economics is that aggregation destroys structure**, and a proposition asserting that measured aggregates are folds over units owes it an answer rather than a citation.

The answer is one distinction, and it is not a hedge.

> **SMD is a theorem about maps. P3 is a claim about states.**

What SMD constrains is the aggregate excess demand *function* (an object taking prices and returning quantities, assembled from individual demand functions) and what it establishes is that essentially nothing about the individual functions survives the assembly. What P3 asserts folds is the extensive **state**: how much physical stock is held, how much claim is recorded against it, and at what rate each moves. Summing steel is not summing preferences. The two claims are therefore not in tension. They are complementary halves of one statement, and the conjunction is sharper than either half alone:

> **Aggregation preserves the extensive state and destroys the behavioural map.**

SMD is the second clause, proved fifty years ago inside the mainstream. P3 is the first. What makes the pair worth having is

that it sorts measurements by what they can carry. A discipline that aggregates in order to recover behaviour (a technology from a production function, a propensity from a consumption function, an elasticity from a demand curve) is estimating the object that does not survive the sum. The standard response has been to impose enough distributional structure on the population that the aggregate is well-behaved (Hildenbrand, 1994; Grandmont, 1992), which is a legitimate research strategy and is also an admission that the structure is imposed rather than inherited.

Three limits on that answer, and the third is this paper's own result.

1. **A state that folds is not thereby a state anyone can observe.** Folding is a property of the object; observability is a property of the measuring layer, and §§2–9 are about a measuring layer failing to see something, as is Paper II. This framework's own results are the reason to doubt that the folded state is available to anyone.
2. **"Extensive" is doing real work, and rates are not extensive.** δ, φ and α do not fold by addition. They fold, where they fold at all, as weighted combinations whose weights are themselves state: which is exactly why **§5's cross-class ladder reads the composite $(1-\varphi)\odot\delta$ rather than φ**, and why ranking by the parameter can invert the ordering rather than blur it. §5 is what P3 looks like when the rate is forgotten.
3. **Diagonality across classes within a firm-quarter is assumed, and §9.5 rejects it.** The fold is *degraded* at precisely the scale where the accounting is done: degraded rather than severed, because what was rejected is a property of the reporting filter and not of the extensive state, and because the departure is now a measured quantity rather than an open exposure. What is not available is its cause: the design cannot say whether the coupling is economic or an artefact of the order the standards impose on the tests.

This section is the surviving argument of a fourth manuscript, written and not carried forward; `docs/papers/README-v1.md` *records why, and* `docs/papers/paper-IV-composition/paper-IV.md` *carries its own header saying where each of its three surviving parts went.*

Supplementary material

The following are in the repository rather than in this paper, because each is a record a reader may want and none is an argument this paper makes.

- **S1 · What was tested and survived.** The programme's ledger of checks that held, each with what would have killed it: including the row that overturned this paper's own preferred reading of its null. `docs/SURVIVALS.md`.
- **S2 · The efficient-markets reply, withdrawn in full.** Why a sophisticated reader pricing what is knowable does not dissolve the filter, and why the first version of that argument was wrong. `docs/ABANDONED.md` §1.
- **S3 · The crisis framing, and the paper it belongs to.** §2's threshold mechanism has a systemic reading which this paper deliberately does not take; the positioning work is at `docs/papers/paper-III-dual-tensor/POSITIONING-001-crash-risk.md`.
- **S4 · The φ-identifiability conditioning study** that preceded the theorem, on synthetic data only: `docs/notes/NOTE-001-phi-identifiability.md`. It is **not** evidence about §9's null, which used an entirely different, non-parametric estimator.

References

How to read this list. **The edition cited is the edition *consulted* — the copy in the author's possession — not the earliest printing a catalogue happens to list. Where the original's date does argumentative work, because the entry is a translation or because a claim about priority rests on it, the entry is** dual-dated** `original/consulted`. A reprint that changes no pagination is a *printing*, not an edition, and is not dual-dated. Where the copy read was a working paper or an accepted manuscript rather than the typeset article of record, the entry is dual-dated in the other direction: `consulted/published`.**

Each entry carries a mark recording how far verification reached, because a bibliographic record and a text are different objects: a work can exist exactly as cited and still not contain the sentence attributed to it. Bibliographic verification was carried out on 2026-08-10; the six aggregation entries §A.4 rests on were verified on 2026-08-26. Per-entry findings are in the note attached to the entry they describe.

✓ — checked against a publisher page, a library-catalogue record, a Crossref record or the issuing body's own documentation,

not recalled. ✓✎: additionally checked against **the author's own copy**, by reading that copy's title page and colophon. The ✓✎ entries are the ones where doing so changed the citation. ✓⧗ (bibliographically verified, but the **text** consulted is a pre-publication version; any quotation is attributed to the version read and may not appear in the article of record; three entries carry it. ⧗ *alone*) the bibliographic record is verified and the **text was not read**; the characterisation rests on named secondary sources, and the entry says so in its own note. Three entries carry it. Two entries carry **no mark at all**, each stating in its own note why it is unmarked, and those two are the only unmarked entries in the list.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics*, 24(1), 3–37. ✓ *(Cited for the asymmetric-timeliness result its title carries verbatim,* the asymmetric timeliness of earnings. ***Read at source****; the volume, year and page range are confirmed against the typeset article. Nothing is quoted from it.)*

Ball, R., Kothari, S. P., & Nikolaev, V. V. (2013). Econometrics of the Basu asymmetric timeliness coefficient and accounting conservatism. *Journal of Accounting Research*, 51(5), 1071–1097. ✓ *(§6 cites it for its stated expectation that firms with shorter asset maturity exhibit lower timely loss recognition, and for reading that dependence as the measure behaving correctly.* ***Read at source****, and the characterisation held: the passage sits in their §5.2, headed "Other Determinants of Conditional Conservatism," and the expectation is stated of "companies with short operating cycles, short investment cycles, or short asset maturity." The determinant reading is theirs and is explicit: they conclude that the measure "is unbiased under the null hypothesis of zero asymmetry, and that under the alternative hypothesis it captures conditional conservatism," in direct rebuttal of the invalidity critiques. Asset maturity is one of several examples they give of a comparative static, not their headline; §6 is worded accordingly.* ***No page is cited****: the text consulted was a full-text copy reporting itself as the published article, not the typeset original, and the MIT deposit (handle 1721.1/87767:* not* 87766, which is the different 2013 paper in *The Accounting Review*) refused every retrieval route attempted, so nothing is quoted beyond the two phrases above and no absence is claimed of the typeset article.)*

Bateman, H. (1910). The solution of a system of differential equations occurring in the theory of radioactive transformations. *Proceedings of the Cambridge Philosophical Society*, 15(V), 423–427. ✓ *(Cited for the function that bears its name and for nothing else; §3.3 characterises only the functional form, which is standard. The bibliographic record is from catalogue listings rather than the au-*

thor's own copy, and no text is quoted.)

Beaver, W. H., & Ryan, S. G. (2000). Biases and lags in book value and their effects on the ability of the book-to-market ratio to predict book return on equity. *Journal of Accounting Research*, 38(1), 127–148. ✓ *(Cited for the bias/lag decomposition its title carries verbatim (*biases and lags in book value. *§12 identifies this as the closest prior art to §5's filter, so the entry is load-bearing against this paper rather than for it.* ***Read at source****; §7 quotes their method) regressing the ratio "on the current and six lagged security returns with fixed firm and time effects", from* ***p. 128****. The sentence recurs at p. 135 as "six lagged* annual *security returns", which is not the wording quoted. That design is this paper's returns repair carried out twenty-six years earlier, and the decomposition is theirs: Ryan (1995) supplies the regression and assumes conservatism away.)*

Beaver, W. H., & Ryan, S. G. (2005). Conditional and unconditional conservatism: concepts and modeling. *Review of Accounting Studies*, 10(2–3), 269–309. ✓ *(****Read at source.*** *Cited in §6 as the nearest accounting-native ancestor of the present confound: their preemption mechanism runs a depreciation schedule against measured conditional conservatism explicitly. One sentence is quoted, from their development of the tangible-asset case. Theirs is a signed comparative static and not an identification claim, and §6 says so rather than recruiting it.)*

Bellman, R., & Åström, K. J. (1970). On structural identifiability. *Mathematical Biosciences*, 7(3–4), 329–339. ✓ *(Cited in §3.3 for the definition of structural identifiability and the transfer-function criterion, which is what the source is characterised on. Characterised at abstract level; the pole-set consequence drawn in §3.3 is this paper's statement of the mechanism, not theirs.)*

Barlow, R. E., Marshall, A. W., & Proschan, F. (1963). Properties of probability distributions with monotone hazard rate. *The Annals of Mathematical Statistics*, 34(2), 375–389. ✓ *(§8 cites Theorem 6.3, which is quoted in the paper in the form used here: the moment generating function is finite below the limit inferior of the hazard rate and infinite above its limit superior. That theorem is what turns this model's $\alpha > \delta$ into a statement about the recognition lag's tail. The paper's equation (6.2) (a hazard bounded between two constants bounds the survival function between the corresponding exponentials) is the same result in the form a reader may find more familiar.)*

Bleck, A., & Liu, X. (2007). Market transparency and the accounting regime. *Journal of Accounting Research*, 45(2), 229–256. ✓ *(Read in full text; the copy consulted carries the journal's own title page (vol. 45 no. 2, May 2007, DOI 10.1111/j.1475-679X.2007.00231.x)*

so it is the typeset article and not a pre-publication version. §2 and §12 both cite it against this paper: it states §2's volatility result nineteen years earlier.)

Bushman, R. M., & Williams, C. D. (2015). Delayed expected loss recognition and the risk profile of banks. *Journal of Accounting Research*, 53(3), 511–553. ✓

Debreu, G. (1974). Excess demand functions. *Journal of Mathematical Economics*, 1(1), 15–21. ✓ *(§A.4 cites it as one of the three SMD papers, for the theorem and not for a passage. Crossref record verified; nothing is quoted.)*

Financial Accounting Standards Board. *Accounting Standards Codification*, Topic 350 (*Intangibles) Goodwill and Other*; Topic 360 (*Property, Plant, and Equipment*; Topic 280) *Segment Reporting*. ✓

Garrett, E. R. (1994). The Bateman function revisited: a critical reevaluation of the quantitative expressions to characterize concentrations in the one compartment body model as a function of time with first-order invasion and first-order elimination. *Journal of Pharmacokinetics and Biopharmaceutics*, 22(2), 103–128. ✓ *(Cited in §3.3 as the bridge between the Bateman function and the flip-flop phenomenon, which its own title and abstract establish. Characterised at abstract level; nothing is quoted.)*

Georgescu-Roegen, N. (1971). *The Entropy Law and the Economic Process*. Harvard University Press. ✓✎ *(The copy consulted is the Harvard Paperback second printing, 1974, ISBN 0-674-25781-2; a printing is not an edition, so no dual date.)*

Grandmont, J.-M. (1992). Transformations of the commodity space, behavioral heterogeneity, and the aggregation problem. *Journal of Economic Theory*, 57(1), 1–35. ✓ *(§A.4 cites it, with Hildenbrand (1994), for the strategy of restoring aggregate regularity by restricting the population's heterogeneity. Crossref record verified; the* ***text was not read****, and the characterisation claims no more than the title states.)*

Godley, W., & Lavoie, M. (2007). *Monetary Economics: An Integrated Approach to Credit, Money, Income, Production and Wealth*. Palgrave Macmillan. ✓✎ *(Copy consulted confirms first published 2007, ISBN 978-0-230-50055-6.)*

Hildenbrand, W. (1994). *Market Demand: Theory and Empirical Evidence*. Princeton University Press. ✓ *(Frontiers of Economic Research series; verified against the publisher's own page. §A.4 cites it for the dispersion-of-characteristics route to a well-behaved aggregate. The* ***text was not read****.)*

Khan, M., & Watts, R. L. (2009). Estimation and empirical properties of a firm-year measure of accounting conservatism. *Journal of Accounting and Economics*, 48(2–3), 132–150. ✓ *(§6 cites it both for C_Score and for its reported association between longer*

investment cycles and higher measured conservatism. Characterised at abstract level.)

Kuan, I. H. S., Wright, D. F. B., & Duffull, S. B. (2023). The influence of flip-flop in population pharmacokinetic analyses. *CPT: Pharmacometrics & Systems Pharmacology*, 12(3), 285–287. ✓ *(Cited in §3.3 for the classification of flip-flop as a failure of global rather than local identifiability.* ***Read at source*** *(PubMed Central PMC10014047). They write of* local *identifiability rather than of a failure of global identifiability, and §3.3 uses their adjective. Their concluding sentence qualifies the "finite set" formulation as "not just a finite set of parameter values but a partial permutation of the set"; §3.3 carries that qualification too. Two short phrases are quoted; no page is cited, the article running to three pages without internal pagination in the deposit consulted.)*

Dutta, S., & Patatoukas, P. N. (2017). Identifying conditional conservatism in financial accounting data: Theory and evidence. *The Accounting Review*, 92(4), 191–216. ✓⧗ *(Cited in §6 as the nearest existing claim and the one most needing separation. The* ***text*** *consulted is the open UCLA Anderson working-paper version, read in full for the decomposition, the three named confounders and the accrual-variance-spread repair; the displayed algebra rendered unreliably in that copy, so nothing interior to their coefficient B is asserted here and no page is cited. Pagination of the published article is verified bibliographically and has* ***not*** *been checked against the working paper's.)*

Fisher, F. M., & McGowan, J. J. (1983). On the misuse of accounting rates of return to infer monopoly profits. *American Economic Review*, 73(1), 82–97. ⧗ *(Cited in §6 for the shape of its confound (a reporting-rule parameter against an asset-life parameter inside a published ratio) and for the fate of the inference drawn from it.* ***Not read****; the record is verified and the characterisation rests on the Long and Ravenscraft comment below, whose working-paper version was read in full, and on secondary accounts. No quotation is taken from it.)*

Hayn, C., & Hughes, P. J. (2006). Leading indicators of goodwill impairment. *Journal of Accounting, Auditing & Finance*, 21(3), 223–265. ✓ *(§8 cites it for the two figures its abstract states: goodwill write-offs lag the economic impairment of goodwill by an average of three to four years, and for a third of the companies examined the delay extends up to ten. Cited for the existence and length of the delay, which it establishes, and not for its shape, which it does not model. The page range is the publisher's landing-page range and was not checked against the typeset issue.)*

Jorgenson, D. W. (1966). Rational distributed lag functions. *Econometrica*, 34(1), 135–149. ✓ *(§8.4 cites it for the density result*

(that an arbitrary distributed lag may be approximated to any desired accuracy by a rational lag function, of which a constant hazard is the lowest-order member), which is the reason REG-005 predicted the shape would be invisible. Verified at abstract level against the Econometric Society's own record, from which the approximation claim is taken verbatim; the body was not read and nothing else is attributed to it.)

Lanczos, C. (1956). *Applied Analysis*. Englewood Cliffs, NJ: Prentice-Hall. *(§8.4 cites the exponential-decomposition example at pp. 272–280 for the classical ill-conditioning of exponential-sum fitting.* ***Not read****, and the entry carries no verification mark for that reason: every copy located was lending-restricted and no full text was obtained. The page range is from the NIST Statistical Reference Datasets documentation of the same example, and everything §8.4 draws from it is drawn through Varah (1982), which quotes it with page citations. Nothing here rests on it alone.)*

Marshall, A. W., & Proschan, F. (1972). Classes of distributions applicable in replacement with renewal theory implications. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Volume I. Berkeley: University of California Press. *(§8 cites it for the new-better-than-used-in-expectation bound that signs the correction: an NBUE distribution's moment generating function is dominated by that of the exponential with the same mean.* ***Deliberately unmarked.*** *The Berkeley Symposium volume was not consulted and the page range is omitted rather than guessed; the attribution is taken from the reliability literature that rests on it. §8 does not depend on the citation, since the direction it predicts is also measured directly, but the reasoning that produced the prediction is this lemma's and is credited as such.)*

Mantel, R. R. (1974). On the characterization of aggregate excess demand. *Journal of Economic Theory*, 7(3), 348–353. ✓ *(§A.4; Crossref record verified, nothing quoted.)*

Nakagawa, T., & Osaki, S. (1975). The discrete Weibull distribution. *IEEE Transactions on Reliability*, R-24(5), 300–301. ✓ *(§9.4's fitted lag distribution is this one, in the survival parameterisation* $P(T \geq t) = q^{(t}k)$ *the source defines, whose discrete hazard is increasing exactly when* $k \geq 1$ *and which nests the geometric at* $k = 1$*. Named here because a distribution fitted and reported ought to say whose it is.)*

Potepa, J., & Thomas, J. (2023). Goodwill impairment after M&A: acquisition-level evidence. *Journal of Financial Reporting*, 8(2), from p. 131. ✓⧗ *(§8 cites it as the closest existing treatment of impairment* timing*: their design tracks each acquisition to its first impairment within a ten-year window and stops, which their own text describes as effectively a hazard model. It is cited for what it*

is and for what it is not (a covariate model with no baseline shape estimated), which is the gap §9.4's fit occupies. The text consulted is the authors' working paper rather than the typeset article, hence ✓⧗; the end page could not be confirmed from an open source and is omitted rather than guessed.)

Kay, J. A. (1976). Accountants, too, could be happy in a golden age: The accountant's rate of profit and the internal rate of return. *Oxford Economic Papers*, 28(3), 447–460. ⧗ *(Cited in §6 for the analytical result that precedes Fisher and McGowan's numerical demonstration.* ***Not read****; record verified, characterisation from secondary sources.)*

Long, W. F., & Ravenscraft, D. J. (1984). The misuse of accounting rates of return: Comment. *American Economic Review*, 74(3), 494–500. ✓⧗ *(Cited in §6 for the rebuttal. The* ***text*** *consulted is the open FTC Bureau of Economics Working Paper No. 94, June 1983, read in full; the published comment is verified bibliographically and has not been read. Nothing is quoted.)*

Ryan, S. G. (1995). A model of accrual measurement with implications for the evolution of the book-to-market ratio. *Journal of Accounting Research*, 33(1), 95–112. ✓ *(Cited in §7.3 for the regression Beaver and Ryan (2000) adopt.* ***Read at source****, with the Autumn 1995 erratum at 33(2), 417, which corrects two typesetting errors in equation (5) (a "+" printed for the "=", and ΔMV_{i,t−10} printed for BV_{i,t−10}) and changes no coefficient, hypothesis or result. The erratum's γ term is absent from the equation §7 relies on, which is Beaver and Ryan's (4). Ryan's assumption (A8) "eliminates the possibility of conservative accounting," and his firm effects are a control for what that leaves unmodelled; the bias/lag reading is Beaver and Ryan's. In 1995 the journal carried the name given here.)*

Ryan, S. G. (2006). Identifying conditional conservatism. *European Accounting Review*, 15(4), 511–525. ✓ *(Cited in §6 solely to distinguish a near-identical title.* ***Read at source.*** *The characterisation holds and is now supported by the body rather than the abstract: the word "econometric" does not occur in the article, and "identify" and its cognates are used throughout in the empirical sense of detecting conservatism in data. Nothing is quoted.)*

Sims, C. A. (1971). Distributed lag estimation when the parameter space is explicitly infinite-dimensional. *The Annals of Mathematical Statistics*, 42(5), 1622–1636. ✓ *(§8.4 cites it for the conclusion that finite-dimensional approximations to a lag space are meagre in it and that their approximation error "cannot, in other words, be made asymptotically negligible" (quoted from §9, p. 1634) and for his naming of "the finite-dimensional parameter spaces of rational lag distributions (see Jorgenson (1966))" as exactly the approximating class the result covers, p. 1628. The text consulted is an*

optically-recognised scan rather than the typeset original; the volume, issue and page range are confirmed against the journal's own table of contents. ***The journal is the* Annals of Mathematical Statistics*, not* Econometrica*, and the title word is "explicitly", not "essentially": this entry is commonly miscited on both counts.****)***

Varah, J. M. (1982). *On fitting exponentials by nonlinear least squares.* Technical Report TR-82-02, Department of Computer Science, University of British Columbia. ✓ *(§8.4 cites it for the quantified form of Lanczos's example, and it is the route by which Lanczos is cited at all.* ***Read in full.*** *It attributes the observation to "Lanczos (1956, pg. 279)" and locates the data at p. 273, and reports the Hessian's smallest eigenvalue falling by roughly three orders of magnitude per additional exponential term. A later journal version is believed to exist and was* ***not*** *verified, so the technical report is what is cited and what was read.)*

Nerlove, M. (1958). *The Dynamics of Supply: Estimation of Farmers' Response to Price.* Baltimore: Johns Hopkins Press. ⧗ *(Cited in §3.3 for the combined adaptive-expectations/partial-adjustment model, whose reduced form is the closest economics-native analogue of this paper's exchange symmetry.* ***Not read.*** *The bibliographic record is verified; the reduced-form algebra attributed here (symmetry of every systematic coefficient in $\beta \leftrightarrow \gamma$, asymmetry of the disturbance) was derived and checked in this repository (*`scripts/wt085_returns_conditioning.py`*, E7) rather than taken from Nerlove's text, and the entry claims nothing beyond the model's structure.)*

International Energy Agency & United Nations Statistics Division. *SDG Indicator 7.3.1: Energy intensity measured in terms of primary energy and GDP.* Reported as World Bank series `EG.EGY.PRIM.PP.KD`, *Energy intensity level of primary energy*, compiled for *Tracking SDG 7: The Energy Progress Report* by the IEA, IRENA, UNSD, the World Bank and the WHO. ✓

Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, 41(1–2), 133–145. ✓

Mayo, D. G. (1996). *Error and the Growth of Experimental Knowledge.* University of Chicago Press. ✓✎ *(The copy consulted is the University of Chicago Press edition of 1996, which uses* severity *374 times and* severe test *232 times and is where the severity requirement is introduced. Mayo (2018),* Statistical Inference as Severe Testing * (a later restatement) has not been read. Both the edition-consulted rule and the first-appearance rule select 1996, and they agree here because the book he read is also the origin.)*

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National*

Academy of Sciences, 115(11), 2600–2606. ✓

Odum, H. T. (1996). *Environmental Accounting: Emergy and Environmental Decision Making*. John Wiley & Sons. ✓✎ *(Copy consulted gives © 1996 John Wiley & Sons, Inc., New York, ISBN 0-471-11442-1. This entry was re-pointed away to* Environment, Power, and Society *(Columbia, 2007) when the first library sweep did not find the 1996 book, then restored when it did. The sweep, not the citation, was wrong.)*

Quine, W. V. O. (1951). Two dogmas of empiricism. *Philosophical Review*, 60(1), 20–43. ✓

Soddy, F. (1926/1961). *Wealth, Virtual Wealth and Debt: The Solution of the Economic Paradox* (3rd ed.). Omni Publications. ✓✎ *(The copy consulted is the third edition, LCCN 60-53331, printed in the United States under a Britons Publishing Company, London, title page, and described on that page as a reprint of the second edition of 1933 "containing new material and Foreword to the American Nation". The first edition is George Allen & Unwin, London, 1926; the copy's own* Preface to the First Edition *is dated January 1926 and its Addition to the Second Edition refers to "the book, which first appeared in 1926". The term* virtual wealth *is in the 1926 title, so 1926 is the earliest appearance this paper can support; Soddy's own footnote points back to* Cartesian Economics *(Hendersons, 1922) as prior work on the subject, and whether the term itself originates there has NOT been checked: the 1922 pamphlet is not in the author's library. No claim of priority is made in the text, so none is made here.)*

Sonnenschein, H. (1972). Market excess demand functions. *Econometrica*, 40(3), 549–563. ✓ *(§A.4; Crossref record verified, nothing quoted.)*

Sonnenschein, H. (1973). Do Walras' identity and continuity characterize the class of community excess demand functions? *Journal of Economic Theory*, 6(4), 345–354. ✓ *(§A.4; Crossref record verified, nothing quoted.)*

Terpstra, T. J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae*, 14, 327–333. ✓

Use of AI assistance

Anthropic Claude Opus 5, at high reasoning effort, was used throughout as a research and drafting assistant: literature retrieval, adversarial review, code review and prose drafting. All claims, results and final text are the author's, and every computational

result is produced by committed code in the repository named in §13.
