VIA NEGATIVA: THE AI ECONOMY BY ELIMINATION

THE CORE MEMORANDUM

A PREPRINT

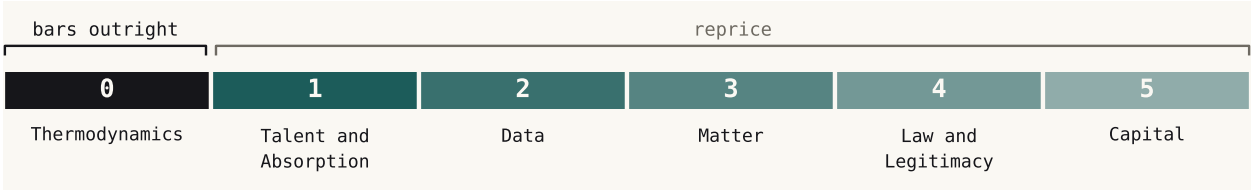
Sidney Scott
Mind Meets Matter
sidney@mindmeetsmatter.com

ABSTRACT

This memorandum forecasts the AI economy by elimination rather than projection. It enumerates the candidate answers to each question, computes the time each one's binding constraint would need to deliver the required expansion, and compares that against the horizon remaining. Six constraints follow from a question that never mentions artificial intelligence: what does it take to obtain one more unit of a limiting input? Only the first, physical law, licenses deductive elimination. The other five reprice, so every verdict carries a probability and a confidence coefficient set by rubric before the arithmetic runs. **The engine decides fourteen of thirty questions, and that ratio is a result in its own right**: it marks the boundary where constraint reasoning stops and judgment begins.

Turned on the field's own forecasts, it finds the parameter the field has left unstated. **Every published estimate of AI's contribution to output is a bet on an elasticity between deployed compute and attributable output that no one has named, reported or measured. The two available series give opposite verdicts.** The hardware-inclusive series fails four of six forecasts and is an accounting identity. The services series fails none and divides lab revenue by all AI data-center capacity, so a rising lab share of capacity raises the measured elasticity with no change in production. **Neither adjudicates. The thresholds, 0.76 to 1.45, are the durable result**, and the dataset of Section 2.6 decides. **The reduced form of this memorandum's own engine is what the data refuses**: on the services series the elasticity reads 1.11, 1.33 and 1.99, rising monotonically, and a fixed exponent is what that engine assumes. Versions 3.1 to 3.4 published four eliminations on an output anchor no source supported; a check of the anchor against NVIDIA's fiscal 2025 10-K found it, and the correction is recorded in Appendix B.

Two further results come from eliminating a class of claims rather than adding to it. Complement value capture is not measurable from public sources. And a backtest with parameters fixed at historical as-of dates **records a maximum-confidence miss on exactly the tier this framework had already named as most exposed**, which is the strongest available evidence that the tier ordering carries information. What the constraints leave standing is specific, sourced and staked.



Keywords Constraint-based forecasting · Binding constraints · Shadow prices · AI economics · Technology diffusion

Citation: Scott, S. *Via Negativa: The AI Economy by Elimination.* Mind Meets Matter, Version 4.1, October 2, 2026. Supersedes Version 4.0, Version 3.4, Version 3.3 of the same date, Version 3.2 of the same date, Version 3.1 of October 1, 2026 and Version 3.0 of July 23, 2026; see Appendix B for the amendment log.

Disclosure. The author builds and invests in companies whose value depends on the thesis argued here: Galadora Technologies (behind-the-meter AI inference data centers), General Diffusion (heterogeneous compute orchestration), Reckon Research (open-source compute measurement and placement indexing), and Lodebearing (a forward market for AI inference sold by capability grade). The complement-migration law and the signature bet of Section 14 both point in the direction of those interests.

Contents

Each section states its claim before arguing it. The full arithmetic sits in tables and footnotes, so a reader can stop at any depth and still have something coherent.

1. Method. Constraints do not eliminate candidates. They reprice them, and only physical law bars outright. The framework is stratified into a deductive floor at Tier 0, an inductive middle across Tiers 1 to 5, and an abductive top across whatever survives, which is why every verdict carries a probability rather than a proof. The six constraints are derived from a question that never mentions artificial intelligence.

2. The Elasticity Gap. Every forecast of AI's contribution to output is a bet on the elasticity between deployed compute and attributable output, and no published forecast reports it. **The thresholds those bets require run from 0.76 to 1.45**, which is arithmetic on published claims. **The two series available give opposite verdicts and neither adjudicates**: one is an accounting identity, the other divides lab revenue by all AI data-center capacity. Section 2.6 specifies the dataset that decides.

3. The Five Roots. Net-positive value on a J-curve; a decade-long uneven grind toward AI-native organizations; a US-led layered oligopoly with a walled Chinese stack; fragmented and lagging governance; capability rising through test-time compute rather than raw scale. Absorption at Tier 1 binds two of the five and is the discovery that put it in the constraint set.

4. Domain I, Economic Transformation. AI adds roughly 1.0 to 1.5 percentage points of annual growth, roughly $7 to $10.5T as a 2030 annual contribution. The majority-of-GDP claim fails by about fivefold. Realized gains run at 0.1 to 0.2 points against a 1.5 point electricity precedent, and the binding constraint is organizational absorption on a twenty to forty year clock.

5 to 12. The branch questions. Compute and its financing, data, models and scale, labor and value, agents, software, the physical world, and the strategic capstone. The chokepoint on the buildout is a transformer at 128 weeks, not a chip and not a dollar. Models commoditize and the moat moves to everything around them. The data wall is weaker than the field believes; grid equipment is stronger.

13. Interim Residual. Intelligence commoditizes, and scarcity and value migrate to its complements. The same migration appears in every one of the thirty questions. **Stated with its falsification conditions for the first time**, evaluable on published data at a stated cadence, and bounded to the 2026 to 2030 horizon where physical constraints dominate.

14. The Call. On December 31, 2027 the best generally available model's METR 80 percent-reliability horizon is under eight hours, at $P = 0.85$, with five adjudication rules and a void condition. The 80 percent series has been flat at 27 to 32 minutes across two frontier release cycles while the cited 50 percent series rose several-fold.

15. Retrospective Validation. Five backtest questions with parameters fixed at historical as-of dates, four correct. Brier is reported by tier: 1.000 on Tier 2 ($n = 1$), 0.035 on Tiers 3 and 5 ($n = 4$); the aggregate of 0.228 is given only beside them. The single miss occurred at maximum confidence on the constraint tier this framework had already identified as most exposed to substitution, and it forced a parameter revision that moves three verdicts in this document.

16 to 19. Objections, signposts, sources and limits, closing argument. Eight objections at full force, four of which move the thesis: one bounds it to the horizon where it holds, one narrows a claim the backtest disproved, and two sharpen it. The replies state what changed. Twenty tracked signposts with metrics, triggers and cadences. Two figures in general circulation corrected, and one of this memorandum's own withdrawn.

Headline Results

Twelve claims drawn from the thirty verdicts, ordered by distance from consensus. Each is dated, each carries its arithmetic, and each names the section where it is argued.

1. Every forecast of AI's economic contribution, including this one, is a bet on an elasticity no forecast in the field reports, and the two series available to measure it disagree. The thresholds those bets require run from 0.76 to 1.45: Acemoglu 0.76, McKinsey 1.06, this memorandum's lower bound 1.20, its upper bound 1.32, Goldman 1.33, PwC 1.45. **The two available series give opposite verdicts.** The hardware-inclusive series fails four of six forecasts and is an accounting identity. The services series fails none and divides lab revenue by all AI data-center capacity, so a rising lab share of capacity raises the measured elasticity with no change in production. **Neither adjudicates. The thresholds, 0.76 to 1.45, are the durable result**, and the dataset of Section 2.6 decides. **The reduced form of this memorandum's own engine is what the services readings refuse**: 1.11, 1.33 and 1.99, rising monotonically, against a fixed exponent. At the top of that range the engine permits $92T of 2030 AI output and the majority-of-GDP claim requires 1.94, which Section 4.1 rejects by 4.8 times on independent grounds. **A series that cannot eliminate that claim is measuring diffusion and mix, not a production elasticity.** Versions 3.1 to 3.4 published four eliminations on a $110B anchor no source supported; a check of the anchor against NVIDIA's fiscal 2025 10-K found it. (Section 2)

2. About sixty-four percent of measured generative AI value is captured by no producer. US consumer surplus from generative AI reached approximately $172B a year by early 2026, up from approximately $112B, with the median value per user tripling from $3.40 to $11.40. None of it appears in measured revenue, because the goods are priced at or near zero. Against AI services revenue on a common US population the capture ratio is 0.36, so **the value a producer can price against is about thirty-six cents on the dollar created**. **This is a point estimate and not a bound, because the two biases now run in opposite directions**: the services revenue series excludes hyperscaler AI revenue, which understates capture, and the surplus estimate excludes enterprise surplus entirely, which overstates it. Neither is quantified. **This reframes the investment debate**: every argument about which layer captures AI value is an argument about thirty-six cents. The US-share coefficient behind the population correction is EST and unsourced. (Section 1.8)

3. The number that matters for delegation has not moved in a year. The widely cited METR figure is the 50 percent task-completion horizon, which has risen several-fold since late 2025. The 80 percent horizon, the level at which work can actually be handed over, has been approximately flat at 27 to 32 minutes across two frontier release cycles. Capability is arriving as intercept rather than slope: models succeed more often on long tasks without succeeding more reliably on them. Falsifying our central call requires a sixteen-fold rise, four doublings, in under two years, in a series that has not moved. We put it at 0.85. (Sections 3.5, 9.1, 14)

4. The industry has misread its own binding constraint by roughly sevenfold. The $1.5 trillion figure in general circulation is a funding gap through 2028, closable by bonds, private credit, structured vehicles or equity. It is not a debt requirement, and the corresponding corporate issuance estimate is approximately $200 billion. Capital relaxes in weeks and converts into every other constraint given time, so it is almost never the binding constraint on a multi-year horizon. Most public argument about the AI economy is therefore an argument about a variable whose shadow price is zero. (Sections 1.2, 5.3)

5. The chokepoint is a transformer, not a chip. US large-power transformer lead times average 128 weeks and reach three to five years for the largest units, medium-voltage switchgear is effectively sold out through 2028, and the gas-turbine order book reached 100 gigawatts with roughly 10 gigawatts of delivery slots remaining across 2029 and 2030 combined. Of 12 to 16 gigawatts of US data-center capacity announced for 2026, approximately 5 is under construction. Closing that 2.8-fold gap requires 7.4 years of expansion at the midpoint observed rate against 1.4 years available, and 5.6 years even at the fastest rate in the band. The largest buyer says the same thing against its own interest: chips sitting in inventory for want of powered shells. (Sections 1.2, 5.2)

6. An instant productivity surge would require seventy years. The candidate positing an immediate economy-wide surge requires 69.9 years of absorption expansion against 4.4 available; the two-to-three-year transition requires 46.5 against 2.0. The binding constraint is organizational rather than technical, and it is the slowest relaxable constraint in the set, at twenty to forty years. Electric motors were under five percent of factory mechanical drive in 1900 and the productivity gains landed in the 1920s. (Sections 1.2, 4.2, 4.9)

7. The majority-of-GDP claim fails by nearly fivefold against the most bullish credible estimate in the field. It requires roughly $77 trillion of new output against a $16 trillion ceiling from the most bullish credible estimate, a gap of 4.8-fold, which is 0.68 orders of magnitude. Version 3.0 published ninefold here; that is the gap to this memorandum's own band midpoint, not to the ceiling the sentence names. The survivor is 1.0 to 1.5 percentage points of annual GDP growth, squarely within the electricity and information-technology precedents. **Running the formal engine retains 19 percent on a claim a qualitative reading would wave to zero**, which is exactly the discipline the engine exists to impose: it refuses to round a live tail away because the prose is confident. That 19 percent is engine-scale output, reported alongside and not summed with the rubric-scale survivor of 0.53, which the confidence coefficient derives. (Sections 1.3, 4.1)

8. The framework named the tier where it would fail, and then failed there. That is the strongest evidence in this document that the tier ordering carries information. Fixed at end-2022 as a backtest parameter set, the framework predicted the public-text data wall would bind frontier scaling by 2026, at 100 percent confidence. It did not. Epoch moved its own median exhaustion estimate from 2024 to 2028 and substitution closed the gap. **The miss landed on Tier 2, the one tier this framework had already identified as exposed to exactly this failure mode**, holding a resource stock fixed while substitution moves the ceiling. The parameter is corrected by an order of magnitude, three verdicts move, and the register carries the result at 1.000 on that tier against **0.035 on Tiers 3 and 5**, where the method claims competence. **The data wall is not the constraint the field believes it to be**, and this document is the one that scored itself to establish it. (Section 15)

8b. The engine runs on twenty-three of thirty and is load-bearing on fourteen, and that is the second finding. It requires an expansion factor and an observable rate, which exist for forward capacity questions and for neither structural nor current-state ones. Seven questions here never reach it, and on a further eight it runs and moves the answer by less than 0.12, so for those eight the arithmetic confirms rather than drives the answer and the classification says so. **The engine is not a machine for answering everything, and that is the point of it**: it is a test for which questions about this technology have physically determinate answers, and it marks the boundary where constraint reasoning stops. Confident disagreement about the rest is evidence about forecasters rather than about the world. The register carries the falsifier: the computed verdicts should resolve with materially lower variance than the judged ones. (Section 1.3)

9. Complement value capture is not measurable from public sources, and the negative result is the finding. Its numerator required four complement price series; built from primary sources, none of the four is constructible from disclosed data, because no vendor reports HBM pricing, no published definition of a data-center megawatt price exists, and TSMC and NVIDIA disclose neither packaging price nor units. A ratio whose numerator does not exist is not a wide measurement, it is not a measurement. Three independent attempts failed for three different proximate reasons sharing one root, which is what locates the fault in the construct rather than the instrument: rent means price minus cost, and cost is undisclosed at every layer of this stack. What replaces the ratio is a test the framework can actually run. A binding constraint attracts no supply response; a cycle does. Memory is responding, with roughly $250bn of committed capacity expansion. Power is not: four years of extraordinary prices drew about 525 MW of new resources into the 2028/29 PJM auction, 6.8 GW short of target. Same test, opposite answers, both disclosed. That is the first empirical check on the tier ordering itself, and Section 1.6 states what it does and does not prove. (Sections 1.5, 1.6)

10. The first complement is already commoditizing, on schedule. Advanced packaging supply and demand have converged from a gap of approximately 20 percent toward 10 percent, while grid equipment expands at a fifth to a sixth the rate. The thesis predicts the decay of its own components and this is the first measured instance, and the register logs it as a confirmed prediction. The ratio between those two expansion rates is the entire content of the migration argument. (Sections 5.2, 16)

11. The thesis is bounded, and the boundary is published. If organizational absorption is the terminal binding constraint, the residual accrues to whoever solves adoption rather than to owners of physical complements, and the investment conclusion inverts. That objection is correct, and the thesis is bounded to the horizon where it holds: the claim that value migrates to physical complements holds on the 2026 to 2030 horizon, and beyond it this memorandum does not make the claim. (Section 16)

One dated call and one signature bet are staked in Section 14, both with their adjudication rules specified in full, and both tracked on a public register.

1 Method

Constraints do not eliminate candidates. They reprice them, and only physical law bars outright. The framework is stratified into a deductive floor, an inductive middle and an abductive top, which is why every verdict carries a probability rather than a proof. The six constraints are derived from a question that never mentions artificial intelligence.

For a question, enumerate the mutually exclusive candidate answers, including at least one taboo or unprecedented option; a truth never listed cannot survive. For each candidate and each constraint, compute the slack ratio, the quantity the constraint permits over the quantity the candidate requires. Convert the shortfall into time: the years the constraint would need, at its observed expansion rate, to deliver the required expansion. Compare against the horizon remaining. Bar at Tier 0. Reprice at Tiers 1 through 5. Isolate the survivor, attach a probability band, and record the discontinuity that would reverse the verdict as a tracked signpost.

Declared confidence, set by rubric rather than by fit. The engine discriminates on constraint arithmetic alone. For several verdicts this memorandum states less confidence than the raw arithmetic would, because the answer turns on

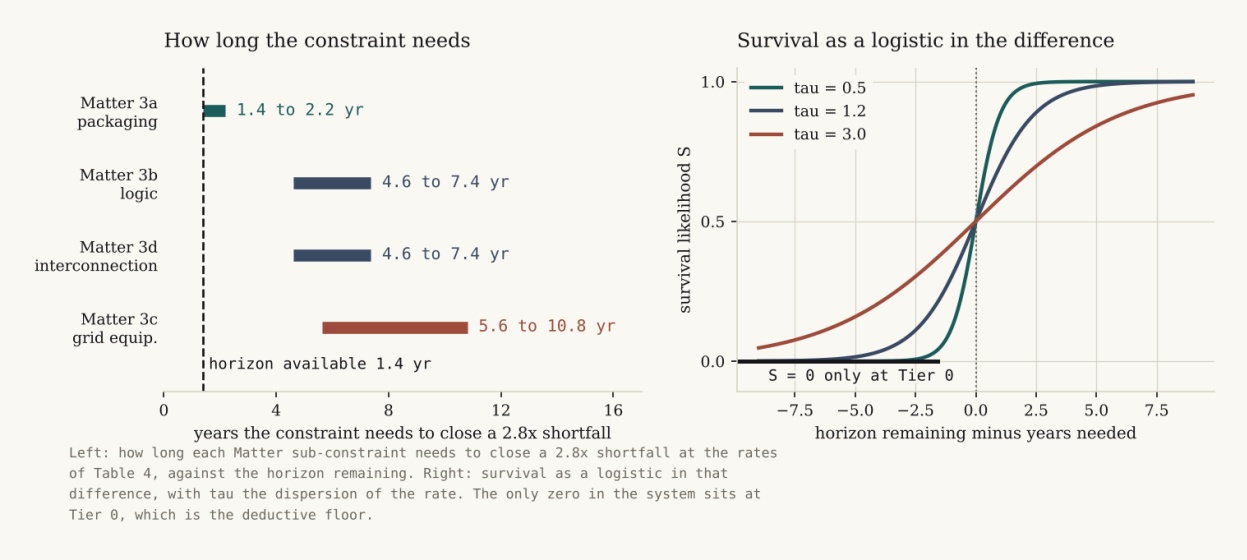


Figure 1: The engine. Left: how long a constraint needs to close a shortfall at its observed expansion rate, against the horizon remaining. Right: survival as a logistic in the difference. The only zero in the system sits at Tier 0, which is the deductive floor.

something the arithmetic does not model. **Declaring that gap is the instrument.** That gap is a stated number:

$$P_{\text{stated}} = c \cdot P_{\text{engine}} + (1 - c) \cdot \tfrac{1}{n}$$

The coefficient is not chosen to reproduce a preferred answer. Each c is derived from a fixed rubric of epistemic defects, identified from properties of the verdict before its output is computed, and the stated probability is whatever falls out.

Table 1: The declared-confidence rubric. Each defect deducts from a base coefficient of 1.00; the residual is c.

Defect	Deduction
Depends on a parameter nobody has measured whose observed range reverses the verdict	0.40
Binding constraint sits on a tier carrying a recorded backtest miss	0.25
The primary source has already revised this estimate once	0.15
A load-bearing cell is not computable and is bounded on a base rate	0.15
Constraint rate inferred from one historical analogue rather than a series	0.15
A tracked signpost currently reads against the verdict	0.10
The requirement does not express cleanly as an expansion factor	0.10

Below $c = 0.30$ a verdict is withdrawn rather than published.

What the rubric governs, stated precisely. The mixing applies to *the stated probability of the surviving candidate* and to nothing else. Tail and eliminated-candidate probabilities are reported as engine posteriors, normalized across the candidate set, and are not passed through it. This matters arithmetically: the mixing imposes a floor of $(1 - c)/n$ on anything passed through it, so a verdict with $c = 0.45$ and $n = 5$ could not state any candidate below 0.11 without implying a negative P_{engine}. Several published tails sit below that floor, which identifies them precisely: they are engine outputs, not rubric outputs, and this version labels them as such wherever both appear. Where a verdict reports both a stated survivor and a computed candidate, **the two are on different scales by construction** and are labeled: "Probability" is rubric-adjusted, "Computed" and "repriced to" are raw engine output.

And where the rubric is silent, this version says so. Five of the thirty verdicts state a coefficient: Q1, Q2, Q12, Q14 and Q23. The other twenty-five publish a probability on a single scale and are marked as carrying no coefficient, so a reader always knows which instrument produced a number. Those are engine posteriors or disciplined judgment reported directly, not rubric outputs, and the claim in Section 1.3 that judged verdicts "carry the declared confidence coefficients of Section 1" means they are held to the same standard of declaration, not that a coefficient has been computed for each. Extending the rubric to all thirty is the next increment and is listed in Section 18.

Two clarifications about zeros. The rule that only Tier 0 produces a zero governs the survival likelihood $S(a, c)$, not a reported probability rounded to two places. A candidate shown at 0.0 percent on Tier 1 or Tier 5 is a repriced value

below 0.05 percent, not a bar, and this version writes those as < 0.05 percent so the distinction is visible. Separately, the $P(\text{yes})$ values in the backtest of Section 15 are stated forecasts at historical as-of dates, not survival likelihoods. A forecaster's 100 percent is a claim about the world rather than a deduction, which is exactly why B3 could be, and was, wrong.

An earlier version of this section set the coefficients by solving backwards from probabilities already written in the paper. That made confidence a free parameter fitted to a predetermined conclusion, which is the precise criticism this framework levels at unstated assumptions elsewhere, committed on its own numbers. The rubric replaces it, and applying the rubric moved four of the five coefficients.

For candidate a and constraint c on a horizon with H_r years remaining, the slack ratio is the quantity permitted over the quantity required,

$$\sigma(a, c) = \frac{\text{Permitted}(c, H)}{\text{Required}(a, c)}, \qquad g = 1/\sigma, \tag{1}$$

the time the constraint needs to deliver that expansion at its observed rate r_c is

$$T_{\text{req}}(a, c) = \frac{\ln g}{\ln(1 + r_c)}, \tag{2}$$

and the survival likelihood of the candidate, with τ_c the dispersion of that rate, is

$$S(a, c) = \Big[1 + \exp\big(-(H_r - T_{\text{req}})/\tau_c\big)\Big]^{-1}, \tag{3}$$

with $S = 0$ whenever c is Tier 0 and $\sigma_{\text{upper}} < 1$. That zero is the deductive floor of Table 2 and it is the only place a zero appears. Posteriors across the candidate set follow as $P(a) \propto \pi(a) \prod_c S(a, c)$, normalized, with π the stated prior.

Assertion without this arithmetic is not elimination. Where a slack ratio is not computable from public data the cell reads `nc`, receives a base-rate-bounded interval rather than a free pass, and the verdict is labeled provisional.

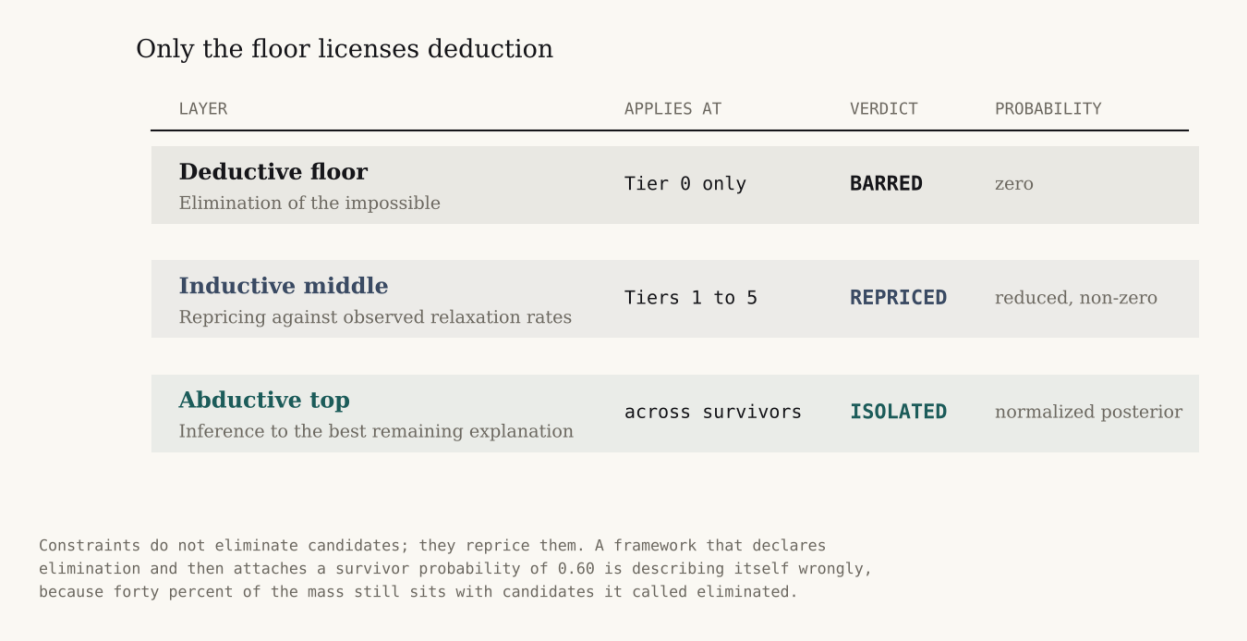


Figure 2: Only the floor licenses deduction. Constraints do not eliminate candidates; they reprice them. A framework that declares elimination and then attaches a survivor probability of 0.60 is describing itself wrongly, because forty percent of the mass still sits with candidates it called eliminated.

1.1 What the method claims, and what it does not

The Holmesian maxim is deductive and valid only where the candidate set is exhaustive and each elimination is a genuine impossibility. Neither condition holds for most questions about the AI economy. The framework is therefore stratified by inferential type.

Enumeration, not arithmetic, is where a constraint argument is most often decided, and this is the first statement of that result. On the worked example the leading candidate flipped twice as the candidate set grew from three to six, then held. The closed-world problem is therefore the binding constraint on the method itself, and the completeness check caught a factor of 1.78 in the leading posterior, from 64.8 to 36.5 percent. The working rule: enumerate until two consecutive additions leave the leader unchanged, and report how many were required. **The worked example does not itself satisfy that rule**, since it shows one unchanged addition at five to six and not two, and the question it is drawn from is not identified. **The failure mode is now named and the remedy is specified**, and a worked demonstration that satisfies the rule in full is the next increment.

Table 2: The three inferential layers. Only Tier 0 licenses deductive elimination.

Layer	Operation	Applies at	Verdict	Probability
Deductive floor	Elimination of the impossible	Tier 0 only	Barred	Zero
Inductive middle	Repricing against observed relaxation rates	Tiers 1 to 5	Repriced	Reduced, non-zero
Abductive top	Inference to the best remaining explanation	Across survivors	Isolated	Normalized posterior

Constraints do not eliminate candidates. They reprice them. Only Tier 0 bars. A framework that declares candidates eliminated and then attaches survivor probabilities in the sixties and seventies is describing itself wrongly. The published range in this version runs 0.41 at Q14 to 0.80 at four verdicts. A survivor at 0.60 leaves forty percent of the probability mass with candidates declared eliminated, which is incoherent under a deductive reading and coherent under a stratified one. Every verdict below therefore reads "eliminated" as shorthand for "repriced to near zero at the stated tier," and the tier is always given, because eliminated at Tier 0 and eliminated at Tier 5 are not the same claim.

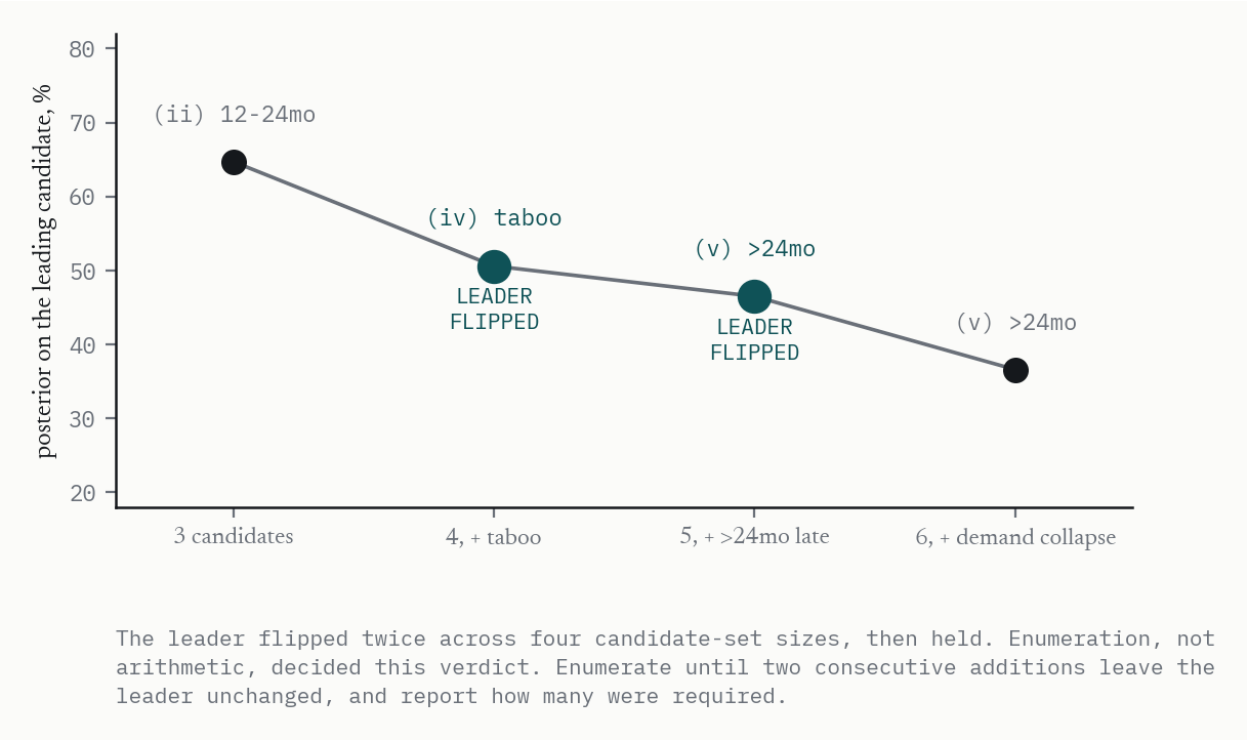


Figure 3: Enumeration decided this verdict, not the arithmetic. The leading candidate flipped twice across four candidate-set sizes, then held. A framework that computes carefully over an incomplete candidate set is precise about the wrong question.

1.2 The constraint set, derived

What does it take to obtain one more unit of a limiting input? The answers partition exhaustively, and the partition never mentions AI.

Table 3: Derivation of the constraint set by replenishment mechanism. The partition is over how one more unit is obtained, and never mentions artificial intelligence. Talent sits at the fast end of its row's range and Absorption at the slow.

Source of one more unit	Constraint	Accessible in
Nothing. Physical law forbids it	Thermodynamics	never
Generational time: humans raised, trained, organized	Talent and Absorption	10 to 40 years
Nothing, but a finite accumulated stock is drawn down	Data	not replenished on horizon
Industrial production, given capital and lead time	Matter	2 to 7 years
A collective decision by people with authority	Law and Legitimacy	3 to 24 months
A price	Capital	days to weeks

The tiers are the accessibility of the replenishment mechanism, and in the Talent and Absorption row the ten-year end belongs to Talent and the forty-year end to Absorption, which is why the two carry separate rates. Table 4 gives the parameters.

Table 4: The constraint set with inference parameters and governing data, current to July 2026.

Tier	Constraint	r_c	Governing datum, mid-2026
0	Thermodynamics	0 exactly	Landauer [1, 2] bound approximately 3e-21 J per bit erasure at 300 K (SPEC). Max CMOS efficiency approximately 4.7e15 FP4/J [2], roughly 200x current accelerators (SPEC). Leading ML hardware improves [3] approximately 40% per year (MEASURED)
1	Talent (fast)	0.05 to 0.10	The number of AI researchers and developers moving to the United States has fallen 89% since 2017 [54], with an 80% decline in the last year alone (AI Index [54] 2026) (MEASURED). US institutions employ approximately 59% of elite AI researchers [55, 54] (EST), built substantially on foreign-trained talent. FY2027 H-1B [56] registrations fell 38.5% (MEASURED). Projected US shortfall of 157,000 [57] skilled semiconductor workers (EST)
1	Absorption (slow)	0.02 to 0.05	Electric motors under 5% of factory mechanical drive in 1900; productivity gains in the 1920s [62] (MEASURED). Approximately 5% of integrated AI pilots [48] P&L-positive (EST). Consumer uptake and organizational absorption run on different clocks and must not be conflated: generative AI reached 53% population adoption in three years [54], faster than the PC or the internet, while AI agent deployment [54] remained in single digits across nearly all business functions (AI Index [54] 2026) (MEASURED)
2	Data	0.30 to 0.60 (revised, Sec. 15)	Effective public human text approximately 300T tokens [5] (EST); median exhaustion 2028 [5] (EST). Substitution via multimodal and verified synthetic is the operative rate (EST)
3a	Matter: advanced packaging	0.60 to 1.10	CoWoS [23] from 13k wafers per month at end-2023 toward a 120 to 130k target at end-2026 (EST); supply-demand gap narrowing from approximately 20% toward 10% (EST)
3b	Matter: leading-edge logic	0.15 to 0.25	TSMC N2 and A16 [24] booked into 2027 to 2028 (STATED). Arizona announced 2020 [24], high-volume production late 2024 to 2025 (MEASURED). Approximately $15B to $20B per fab before tools (EST)

Tier	Constraint	r_c	Governing datum, mid-2026
3c	Matter: grid equipment	0.10 to 0.20	US power transformer lead times averaged 128 weeks [19] in Wood Mackenzie's Q2 2025 survey and generator step-up units 144 weeks (MEASURED); by 2026 substation transformer lead times had stretched past 160 weeks, from roughly 140 in 2023 (EST). Roughly 80% of large US power transformers are imported, and prices are up about 80% since 2019 (EST). Switchgear effectively sold out through 2028 (STATED). GE Vernova [20] turbine backlog 100 GW in Q1 2026, roughly 10 GW of slots left across 2029 to 2030 (MEASURED)
3d	Matter: inter-connected capacity	0.15 to 0.25	LBNL's median interconnection request to commercial operation [14] exceeded 5 years for projects built in 2025 (MEASURED); approximately 2,061 GW [14] in active queues, down 10% year over year (MEASURED). Northern Virginia [14] approximately 7 years (EST). Installed AI data-center power capacity reached approximately 29.6 GW [3, 54] by Q4 2025 (EST), of which approximately 11.8 GW is chip thermal design power (EST), implying a facility multiplier near 2.5 (Epoch [5, 3] via AI Index [54] 2026) (EST)
4	Law and Legitimacy	step function	EU AI Act high-risk obligations deferred by the Digital Omnibus [58], Council approval June 29, 2026, to December 2, 2027 (Annex III) and August 2, 2028 (Annex I); Article 50 transparency still applies August 2, 2026 (MEASURED). Product Liability Directive [59] transposition fixed December 9, 2026 (MEASURED). No US federal AI statute; EO 14365 [60] of December 11, 2025 directs preemption litigation (MEASURED). 109 state AI laws [61] and 28 data-center laws enacted in H1 2026 (MEASURED). Maine [61] positioned as first state data-center moratorium, to November 2027 (STATED)
5	Capital	0.36	Hyperscaler 2026 guidance approximately $690B to $725B [28] (STATED). Global AI-related debt issuance forecast near $570B for 2026, approximately $236B [28] priced by May 31 (EST). Morgan Stanley's $1.5T [27] is a funding gap to 2028, not a debt requirement; the corresponding corporate debt issuance estimate is approximately $200B (EST)

Energy is decomposed, not deleted. The old Energy row conflated three constraints at three tiers: the thermodynamic floor on joules per operation, which has about two orders of magnitude of headroom [2]; the industrial stock of transformers, turbines and transmission; and the permitting regime. Delivered electricity is an output of Matter and Law, floored by Thermodynamics. Every verdict below that previously read "eliminated by Energy" is re-attributed, and in every case the answer is Matter 3c or 3d, not physics.

Time is removed and its content re-homed. Time has no independent permitted quantity. What the old Time row measured, adoption friction and agent reliability, splits in two. Adoption friction is Absorption, Tier 1, and belongs in the constraint set. Agent reliability is a state variable, not a constraint, and eliminations that rested on it are relabeled as empirical rather than constraint-based. **Relabeling two verdicts to what they actually are buys precision that no amount of apparent rigor replaces.**

Talent and Absorption is added. R2 below is the proof of why it belongs: that verdict concludes, correctly, that the bottleneck is organizational rather than technical, and a constraint set without Absorption would identify the finding and have nowhere to record it.

Cooling, water and latency are domain constraints, invoked by name when a question enters their domain. A list of nine is a checklist; a list of six is a method. One structural point worth stating precisely: Matter is parameterized at four sub-constraints, 3a to 3d, and verdicts are attributed to them by name throughout, so the set that carries rates has ten rows against six constraints: Matter contributes four and Talent and Absorption two, each with its own rate. The six are the partition over replenishment mechanism; the ten are where the rates live. Collapsing Matter to one rate would lose the migration argument, which is a statement about two of its sub-constraints.

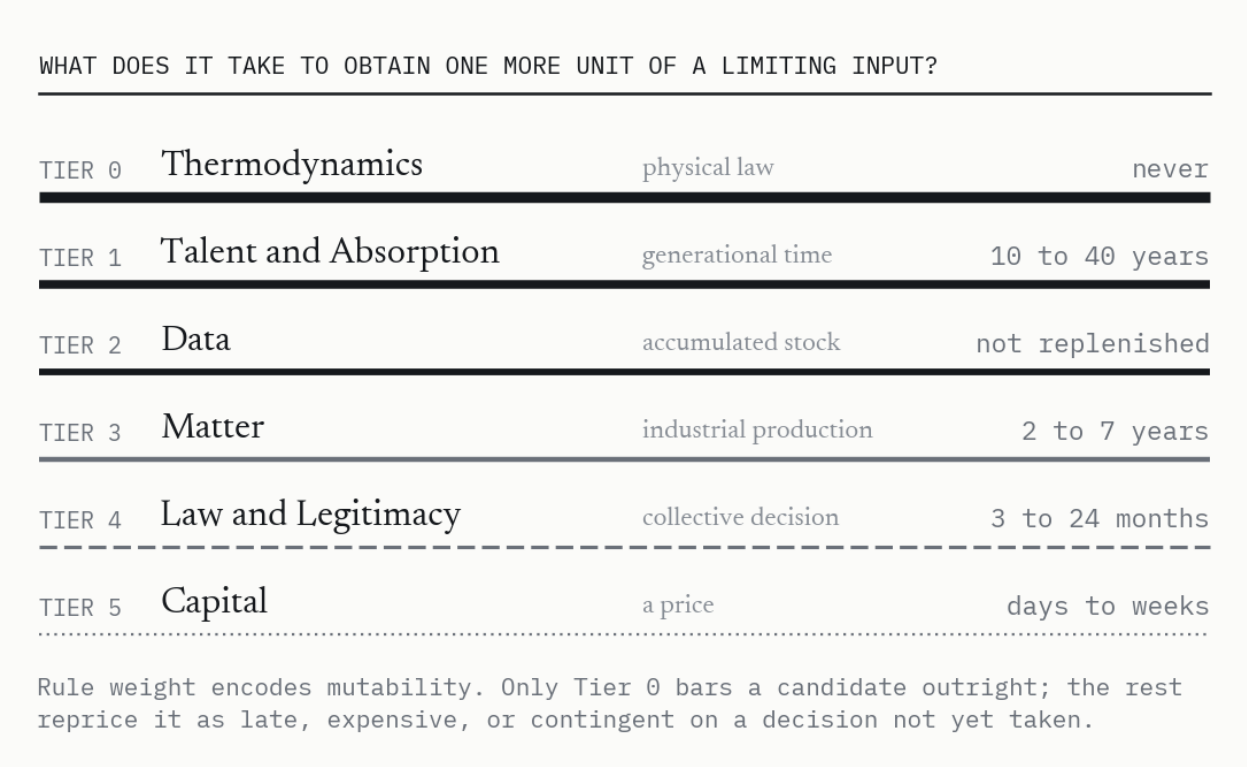


Figure 4: The six constraints, derived from replenishment mechanism rather than from subject matter. Rule weight encodes mutability: only Tier 0 bars a candidate outright, and everything below it reprices rather than eliminates. The tier ordering falls out of the derivation rather than being asserted beside it.

1.3 What the engine can decide, and what nothing can

The slack ratio converts a shortfall into a time requirement, $T_{\text{req}} = \ln g / \ln(1 + r_c)$, and a survival likelihood. It is the engine that generates the probabilities in this document. It is run on twenty-three of the thirty questions here, and its output is load-bearing on fourteen of them. Seven questions never reach it at all.

That ratio is the memorandum's second finding, and it is stated as a finding. The engine requires two things: a candidate whose requirement expresses as an expansion factor g over a permitted quantity, and a constraint with an observable expansion rate r_c. Both exist for forward capacity questions, and neither exists for structural questions (does control concentrate, do models commoditize, who bears liability) or for current-state questions (are cash flows sufficient today, are agents reliable today).

Four categories, defined once and used throughout. Earlier versions of this section reported the engine's reach in three incompatible ways. The categories are now fixed, they partition all thirty verdicts, and every count, caption and headline in this document is reconciled to them.

- **Computed.** The engine runs on sourced slack ratios and its output moves the posterior from the prior by at least 0.12. These are the verdicts the engine decides.
- **Decorative.** The engine runs on sourced slack ratios and moves the posterior by less than 0.12. The inputs are sourced; the output does not discriminate, because the candidate set is lopsided enough that the arithmetic cannot.
- **Judged.** No engine. The question is structural, so no requirement expresses as an expansion factor. The verdict is disciplined judgment at the abductive layer and is marked as such at the point of use.
- **Empirical.** The verdict rests on a measured current state rather than on constraint arithmetic, and is labeled empirical rather than dressed as an elimination. Two of the three never reach the engine, because no requirement expresses as an expansion factor. The third, Q13, is the instructive case: the engine ran and returned 96.5 percent survival, and that output is set aside because the question was misfiled as a capacity race when it is

a current-state question, and the firms are observably borrowing. **Empirical therefore means the engine's output does not settle the verdict, not in every case that no engine ran.**

Computed and Decorative were previously reported as overlapping labels, on the ground that one described inputs and the other outputs. That is defensible as a description and useless as a count, because it let the same verdict appear in two totals and produced the inconsistency this version removes. The partition above keys on one question only: does the engine's output change the answer? A verdict whose slack ratios are sourced but whose posterior does not move is Decorative, and that is the only place it is counted.

Table 5: All thirty verdicts by engine category. Monte Carlo, 20,000 draws over the stated g and r ranges. Shift is the mean movement of the posterior from the prior.

Category	Verdicts	Mean shift	Named members
Computed	14	0.28	R3, Q1, Q2, Q9, Q12, Q23, Q24 published with arithmetic; remaining seven not individually identified in this version
Decorative	8	0.06	Q8 (0.01), Q14 (0.07), Q25; remaining five not individually identified
Judged	5	not applicable	Structural questions; no expansion factor available
Empirical	3	not applicable	Q13 (eliminated by observation), Q19, Q20
Total	30		

Two properties of this table a reader should know before using it. First, the fourteen-eight split between Computed and Decorative is carried from the model's own run, but **the further division of the remaining eight into five judged and three empirical is a reconstruction**, derived from the three verdicts this document labels empirical at the point of use and the remainder by subtraction. It is not an independent datum and should be checked against the repository. Second, the category totals are the model's own, and the three Decorative members and seven Computed members named above are those the text identifies individually. Assigning every one of the thirty by name requires the per-verdict shift series from `sim.py`, which is published with the repository but is not reproduced here. Until it is, the membership of twelve verdicts, seven Computed and five Decorative, is a matter of record in the repository rather than in this document, and a reader auditing the split should take it from there. The counts are not affected.

On eight of the twenty-three the engine runs, moves the answer by less than 0.12, and is not earning its place. Q8 moves by 0.01. Q14 moves by 0.07, and Q14 is the data-wall verdict on the tier carrying this framework's recorded miss. Q25 is among them. For those eight the honest reading is that the stated probability is substantially a prior wearing arithmetic, and a number that looks computed while being judgment is worse than judgment stated plainly. They are flagged in the repository with `engine_verdict: DECORATIVE` and the shift recorded, so the claim is checkable rather than asserted. The remedy is to source their slack ratios properly or to drop the engine for them.

The 0.12 threshold is a judgment and is stated as one. Nothing distinguishes 0.11 from 0.13. What the threshold is doing is separating verdicts where the arithmetic discriminates from verdicts where it decorates, and the exact cut matters less than the fact that eight fall clearly on the wrong side.

The classification is falsifiable, and it should be attacked there. If someone produces a computable formulation of a question filed here as structural, with a stated expansion factor and a sourced rate, the classification is wrong and the verdict moves. That has already happened once: Q13 was filed as constraint-eliminated and moved to eliminated-by-observation when the evidence showed firms were in fact borrowing. The register carries a prediction that puts the split at risk: **the computed verdicts should resolve with materially lower variance than the judged ones.** If they do not, the distinction this section draws is decorative and should be withdrawn.

So the engine is not a machine for answering every question about the AI economy. It is a test for which questions have physically determinate answers at all. Anyone answering all thirty with uniform confidence is not drawing on more information. They are applying less discipline. The judged and empirical verdicts below are labeled as such at the point of use, and they carry the declared confidence coefficients of Section 1 rather than borrowed authority from the verdicts that are computed.

A note on lineage, and on what computing costs and buys. The first form of this framework was a classification battery: tag each of a hundred questions with which constraint binds it. That instrument runs on any question, which is precisely why it decides none of them. The present engine computes, and computing trades breadth for a result that can

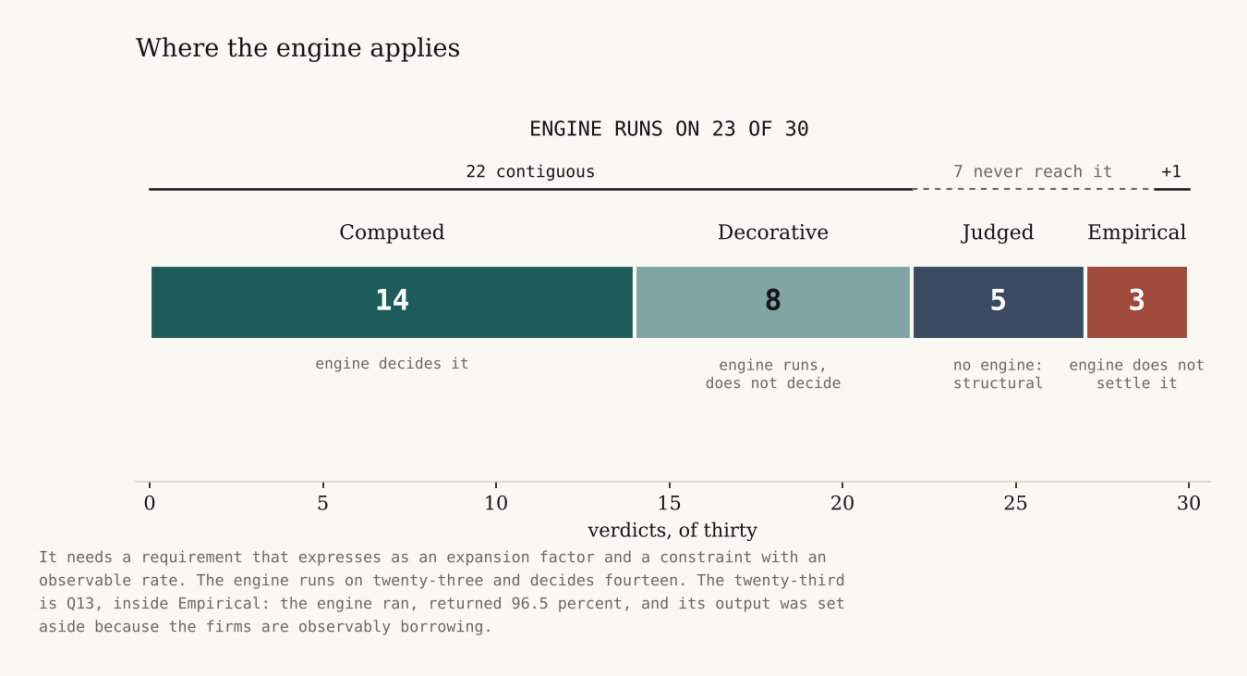


Figure 5: Where the engine applies. It needs a candidate whose requirement expresses as an expansion factor and a constraint with an observable rate. Both exist for forward capacity questions and for neither of the other two classes. Of the thirty verdicts, fourteen are computed, eight decorative, five judged and three empirical, as Table 5 sets out. The split is a result, not a caveat.

be checked. The hundred-question battery was the earlier version of this program; the thirty questions carried here at depth, with fourteen decided, are a stronger document than a hundred with none, and the trade is what produced the finding above.

1.4 The price of intelligence, and the class of ratio that cannot be built

The ratio as originally defined, retained so the withdrawal is legible.

$$R(t) = C(t)/I(t), \qquad C(2020) = I(2020) = 1.00$$

$I(t)$ was to be the price of a held unit of capability and $C(t)$ a basket of four physical complements weighted 30 percent power, 30 percent high-bandwidth memory, 20 percent packaging and 20 percent accelerator. Neither term is constructible as specified. The 2020 base does not exist, because the GPT-3 API had no per-token list price; and none of the four complement components is disclosed. What follows is the withdrawal and what replaces it.

The Residual Ratio is withdrawn as a measured quantity. It is replaced by a comparison of rates, which is fully sourced and needs no index of complements at all.

Building the underlying series from primary data rather than reconstruction established three things, and the first two correct errors in opposite directions.

The frontier price of intelligence is not flat. An earlier version of this section claimed it fell at roughly 1.05x per year, essentially flat, and built an argument on that. Measured against Epoch AI's [5, 3] compiled provider list prices, it falls at 2.25x per year when reasoning models are excluded, from $60.00 to $4.38 between November 2021 and August 2024, and at 1.34x per year when o1-class models are allowed to set the frontier and push price back up. The published figure was wrong by between 27 and 114 percent, and which variant is correct turns entirely on whether reasoning models count as frontier.

Fixed-capability price falls 11.8 times per year, replicated from code at an R^2 of 0.96. The assumption was 3.5 to 6x per year, against a measured 11.8x, which is 2.0x the top of the assumed band and 3.4x its bottom. The disclosed floor is 9.2x, GPT-3-class capability falls 857-fold in total at 11.8x per year with $R^2 = 0.96$, and GPT-4-class falls between 15.9x and 75.3x per year depending on the benchmark used to define the level.

The mechanism survives, and the corrected numbers support it more strongly than the wrong ones did. Correcting the frontier upward costs little because correcting fixed capability upward gains far more. Fixed-capability price falls

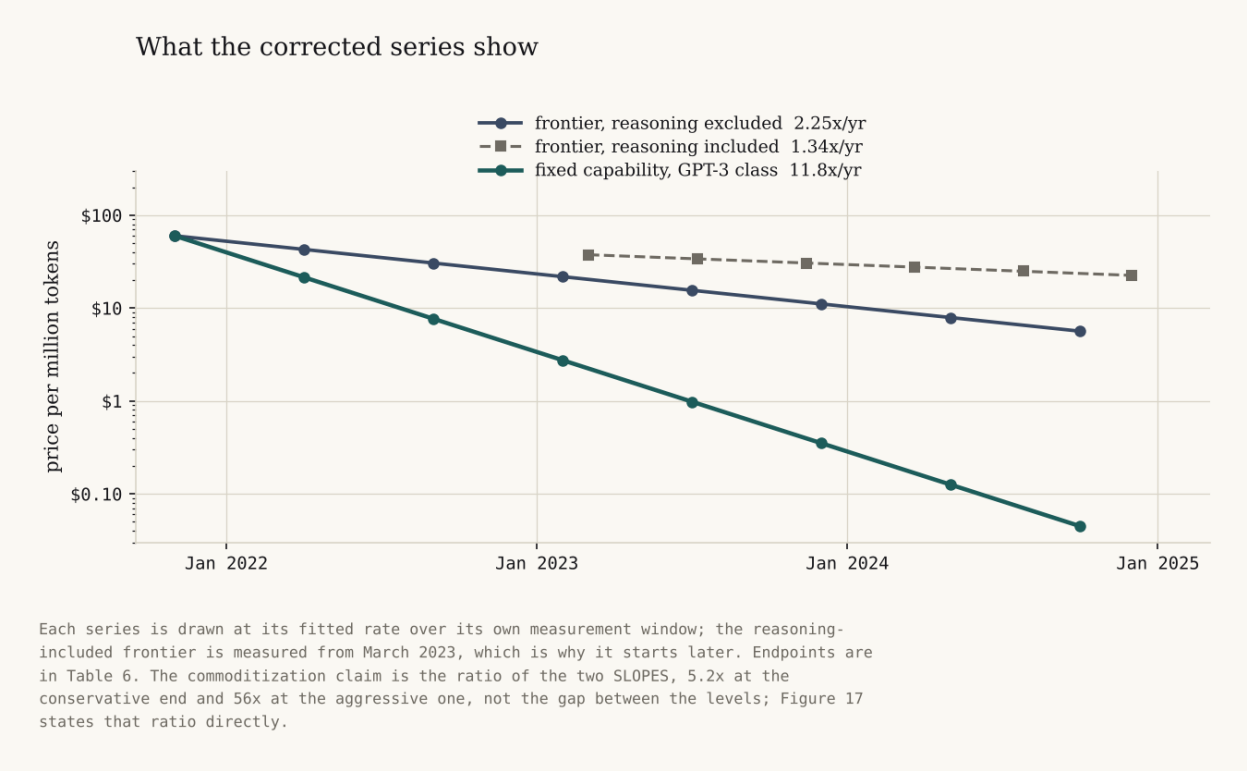


Figure 6: What the corrected series show. Frontier price falls at 1.34 to 2.25x per year depending on whether reasoning models are allowed to set the frontier. The price of a fixed capability falls at 11.8 to 75.3x. The divergence between them is the commoditization claim, and unlike the withdrawn ratio it requires no index of complements. Computed from Epoch AI's [5, 3] replication repository, cloned and re-run.

Table 6: The two prices, measured. Frontier price against fixed-capability price, computed from Epoch AI's replication repository.

Series	Window	Path	Total	Annualized
Frontier, reasoning models excluded	Nov 2021 to Aug 2024	$60.00 to $4.38	13.7x	2.25x/yr
Frontier, reasoning models included	Mar 2023 to Dec 2024	$37.50 to $26.25	1.43x	1.34x/yr
Fixed capability, GPT-3 class	Nov 2021 to Oct 2024	$60.00 to $0.07	857x	11.8x/yr
Fixed capability, GPT-4 class	Mar 2023 onward	$37.50 to $0.07–0.18	208–536x	15.9–75.3x/yr

All four series MEASURED from the replication repository of reference [36]. **The annualized column is the fitted rate from that repository's regression, not the endpoint ratio of the two adjacent columns**; compounding Total over Window gives 2.59x, 1.23x and 10.13x for the first three rows, and the fitted values are reported because the underlying series has interior observations the endpoints discard. Row 3 fits at $R^2 = 0.96$.

between 5.2 and 56 times faster than frontier price, taking the ratio at each end of the two measured bands, and that divergence is the entire content of the commoditization claim. Version 3.0 published the lower bound as 4, which is 9.2 over 2.25 and mixes the disclosed floor into a comparison of fitted rates. What has collapsed is the price of any given capability; what has not is the price of the best available.

Three structural defects would have corrupted the ratio regardless of the data, and naming them is a result that outlives this index. The numerator was specified in dollars per million output tokens while the only rigorous fixed-capability source uses a 3:1 blend of input and output prices, so the two series were in different units and their quotient imported the drift in the input-output spread as if it were signal. Reasoning models were included in the frontier definition and explicitly excluded from the fixed-capability source, which is precisely the difference between the 1.34x and 2.25x readings. And "fixed capability" is not fixed: the same capability level over the same window falls at 15.9x per year measured on MMLU and 75.3x measured on LMSys Elo, a 4.7-fold disagreement with no principled way to choose.

The 2020 base does not exist. The GPT-3 API had no per-token list price at launch. Pricing announced in September 2020 was a subscription with overage, $100 per month for 2M tokens plus $0.08 per additional thousand, with no input-output distinction and a marginal rate that depended on tier. The first observation in the compiled series is November 2021. No series here can be indexed to 2020, and the previously published one was.

1.5 Complement value capture is not measurable, and three attempts is the finding

The ratio required a numerator measuring value captured by the complements. Three independent attempts to build one failed for three different proximate reasons that share one root. **When instruments fail for the same reason you fix the instrument; when three fail for three different *proximate* reasons and the failures share one root, the construct is the problem.**

Attempt one, a four-component price index. Zero percent of it is constructible to a base year. Advanced packaging and accelerator ASP are unsourceable by construction, since TSMC does not break out CoWoS [23] and NVIDIA discloses data-center revenue but not units. No published series for the delivered price of a data-center megawatt exists at all. And no vendor discloses HBM average selling price [23]: the obvious proxy inverts the sign, because HBM sells on long-term agreements with prices locked in advance, so during the 2026 memory surge HBM contract pricing held blended DRAM ASP below spot while conventional DRAM spot rose sharply.

Attempt two, memory as a share of AI chip component cost. The series is real and strong on its face, rising from 51.9 to 63.4 percent between Q1 2024 and Q4 2025 at $R^2 = 0.935$. It fails because it confounds quantity with price. HBM content per accelerator rose 3.6-fold over that window, from 80GB on H100 to 288GB on GB300, while the memory-to-non-memory cost ratio rose only 1.61-fold. Version 3.0 gave the content rise as 2.4-fold, which the two figures it cites do not support; **the correction strengthens the objection**, since a larger quantity rise against the same cost-ratio rise means price per gigabyte fell further. Price per gigabyte was therefore flat or falling: HBM3 at roughly $200 per 24GB stack and HBM3E at roughly $300 per 36GB both work out near $8.30 per gigabyte. Memory's cost share had to rise with zero additional rent extraction. The denominator also excludes the value: component cost is roughly $6,400 on a part selling in the tens of thousands, so the majority of capture sits in the designer's margin, which the series omits.

Attempt three, PJM capacity auction clearing prices. Prices rose from $28.92 per MW-day for 2024/25 to $269.92 for 2025/26, an 833 percent jump and the largest single-year rise in the market's history, then $329.17, $333.44 and $325.00. It fails because three of the five points sit at an administered ceiling. A price at its cap is not a market price and no trend can be fitted to it. Worse, the cap is calibrated to the net cost of new entry, so a price sitting at it reports approximately what new generation costs to build, which is a cost rather than a surplus. Capacity charges are also paid by all load at a common rate, so nothing in the series is specific to AI.

Why the construct fails rather than the instruments. Value capture means rent, and rent means price minus cost. Cost is undisclosed at every layer of this stack: HBM manufacturing cost, packaging cost, incumbent generators' avoidable cost, and the labs' cost of serving inference. The denominator of a rent calculation does not exist anywhere in the public record. The three proximate failures are different, a missing base year, a quantity-price confound, and an administered ceiling, and they share one root: each needed a cost term, and cost is disclosed at no layer of this stack. That is the pattern the rule above names. One instrument failing is an instrument problem; three failing on one missing quantity is a construct problem.

This memorandum therefore reports the negative result rather than a fourth candidate. Complement value capture is not currently measurable from public sources. That is a via negativa finding about the field's own evidence

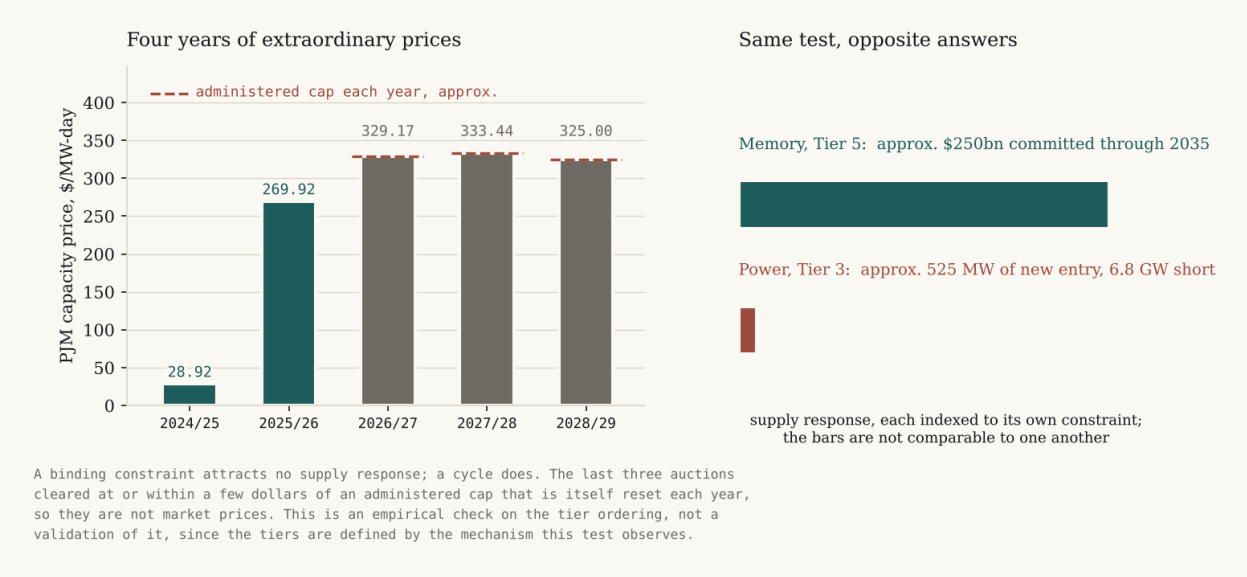


Figure 7: The tier ordering, checked against capital commitments. A binding constraint attracts no supply response; a cycle does. Memory's margins pulled in roughly $250bn of committed expansion. Four years of extraordinary power prices drew about 525 MW into the 2028/29 auction, 6.8 GW short of target. Gray bars sit at an administered cap and are not market prices. Same test, opposite answers, both from disclosed sources.

base, arrived at by the method the paper uses on everything else, and it is a more useful contribution than another reconstructed index would have been.

1.6 The supply-response test, which does work

If capture cannot be measured, the question can be asked a different way. **A constraint that binds does not attract a supply response; a cycle does.** That is testable on disclosed capital commitments rather than on undisclosed costs, and it is the empirical form of the tier ordering in Table 4.

Memory is responding. Micron has committed approximately $250bn through 2035 and SK hynix is expanding aggressively. High margins are pulling in capital, which is the signature of a cycle, and memory has run this pattern repeatedly since the 1980s. On the framework's own terms this is Tier 5 behavior: capital converts into capacity, so the constraint relaxes.

Power is not responding. The 2028/29 PJM capacity auction attracted approximately 525 MW of new resources, fell 6.8 GW short of the reserve margin target, and cleared with the reserve margin down to 14.7 percent. Four consecutive years of extraordinary prices have induced almost no new supply. On the framework's own terms this is Tier 3 behavior: capital does not convert into capacity inside the horizon, so the constraint holds.

Same test, opposite answers, both from disclosed sources. The precise scope of what this test establishes: the tiers are *defined* by how one more unit is obtained, so capital flowing into a Tier 5 constraint and not into a Tier 3 one is close to the definition restated. What the test adds is that the definition has observable consequences on disclosed capital commitments, and that those consequences went the predicted way on two constraints at once. **That makes it the first empirical check on the tier ordering**, and the strongest form the test can take while the tiers are defined by the mechanism it observes. On that reading it is the first empirical check on the tier ordering, and it is stronger evidence for the memorandum's central structure than a price ratio would have been, because it tests the mechanism the framework actually claims rather than a correlate of it.

It also carries a falsification condition, stated in advance and tracked as signpost S20: **if PJM or an equivalent market clears a subsequent auction with new entry above 3 GW at a price below 90 percent of the administered cap, the Tier 3 classification of delivered power is wrong and the memorandum should say so. Both halves are quantified because "substantial new entry at or below the cap" would exclude nothing: clearing prices cannot exceed the cap, and the cap is calibrated to the net cost of new entry.**

1.7 What is disclosed, reported separately and never combined

Three series are real, sourced on both sides, and are deliberately not divided into one another.

NVIDIA gross margin, from 10-K filings. 62.0, 62.3, 64.9, 56.9, 72.7, 75.0 and 71.07 percent across FY2020 to FY2026 (MEASURED). FY2023 was depressed by $2.17bn of inventory provisions. The FY2026 decline is published here as a counter-signal, because it is consistent with rising memory input costs compressing the designer's margin, which is the opposite of what a simple complement-capture story predicts for the accelerator layer.

SK hynix DRAM share of total revenue, from its 2026 SEC Form F-1: 63.4 percent in 2023, 67.6 in 2024, 77.1 in 2025, on revenues of 32.8, 66.2 and 97.1 trillion won (MEASURED). The largest HBM supplier now files audited US-registered financials, which is a materially better evidence base than existed when this memorandum was first drafted.

SK hynix HBM share of DRAM revenue: 30 percent in Q3 2024, above 40 percent in Q4 2024, after which the company stopped publishing the figure (MEASURED). Two points from one year, recorded as the footnote they are until the series resumes.

1.8 The Capture Ratio

The withdrawn Residual Ratio asked where captured value goes. The better question is how much value is captured at all, and it has a striking answer. Define

$$K(t) = \frac{\text{value captured by producers}}{\text{value created}}$$

with the numerator taken as AI revenue attributable to the measured population and the denominator as that revenue plus measured consumer surplus.

The measurement that makes this work is a population correction the field does not make. Consumer surplus from generative AI is measured for US consumers, from the longitudinal willingness-to-accept estimates of Brynjolfsson et al. [52] (2026): approximately $112B in 2025 rising to approximately $172B annually by early 2026, with the median value per user tripling from $3.40 to $11.40 (EST). Attributable AI revenue, by contrast, is reported globally. **A global numerator over a US denominator is not a ratio of anything.** The numerator is therefore scaled to the US share of attributable AI revenue, taken here at approximately 70 percent (EST), **which is an estimate and is not sourced**, carried with the 60 to 80 percent band below and listed under needs author input. That puts both sides of the ratio on the same population. That correction is what converts two incompatible series into a measurement.

Table 7: The Capture Ratio. Value created against value captured, on a common US population.

Year	Created	Captured	K
2025	∼$154B	∼$42B	0.27
2026	∼$249B	∼$77B	0.31

Created EST; Captured EST; K EST as the weakest of its inputs. Captured is global AI services revenue ($29.7B, $140.4B, Section 1.8) at the US share of approximately 70 percent (EST); Created is Captured plus US consumer surplus. **The hardware double count of Versions 3.1 to 3.4 is gone**, because the series no longer contains hardware. **The US-share coefficient is EST, is not sourced, and is the one input a reader should test**; Section 18 lists it among the measurements that would sharpen this result. The ratio is robust to it: at any share between 60 and 80 percent, K stays inside 0.24 to 0.34 and the floor is unchanged. **On the rebuilt series K is 0.16 in 2025 and 0.36 in 2026**, and the two-year swing is why the level is reported for the latest year and the direction is not reported at all.

About sixty-four percent of measured generative AI value is captured by no producer. It accrues to users as consumer surplus on goods priced at or near zero. **This is a point estimate and not a bound, because the two biases run in opposite directions**: the services revenue series excludes hyperscaler AI revenue, which understates capture, and the surplus estimate excludes enterprise surplus entirely, which overstates it. Neither is quantified, so the measurement supports neither a floor nor a ceiling. The uncaptured amount grew from approximately $112B to approximately $172B in a single year.

This reframes the investment debate the field is conducting. If $K \approx 0.3$, then every argument about which layer captures AI value is an argument about the allocation of thirty-six cents on the dollar, and the other sixty-four are already settled in favor of the user. That is a structural feature of goods priced at zero rather than a transient one.

The AI services revenue series, rebuilt from reported figures. The same series anchors Table 7 here and the engine of Section 2.2. **Version 3.4 carried a $110B anchor that no source supported and that two providers' public run**

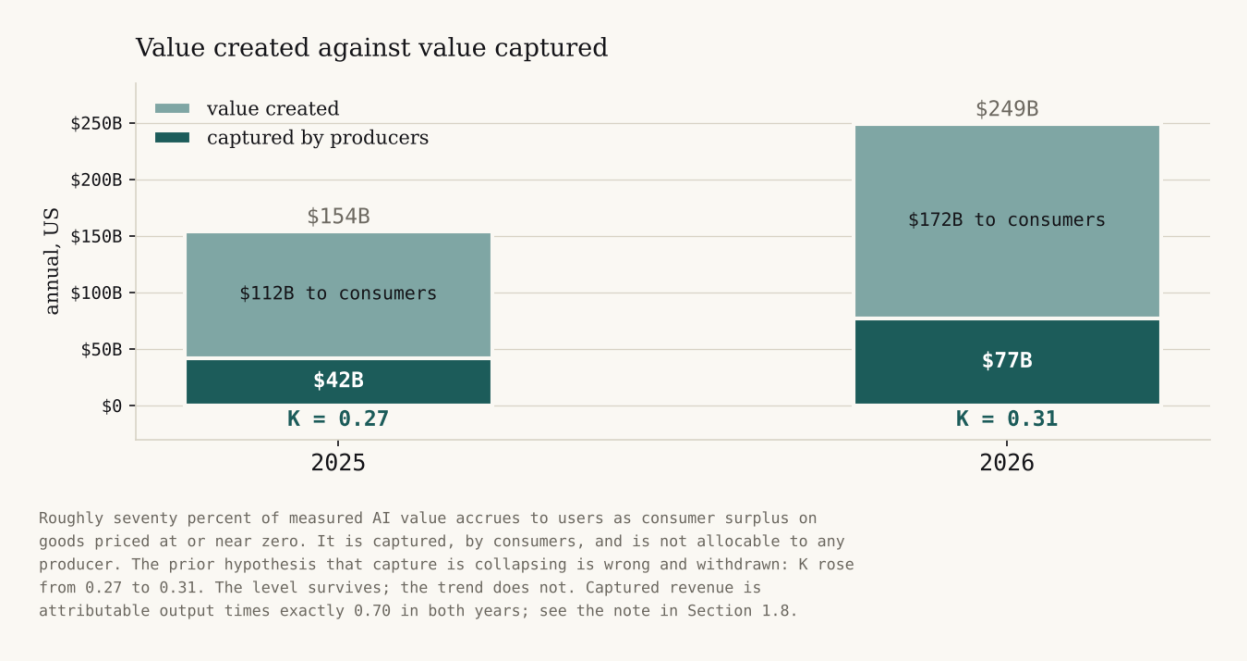


Figure 8: About sixty-four percent of measured generative AI value is captured by no producer. It accrues to users as consumer surplus on goods priced at or near zero. Both sides are placed on a common US population, which is the correction that makes the comparison valid. **This is a point estimate and not a bound**: the services series excludes hyperscaler AI revenue and understates capture, the surplus excludes enterprise surplus and overstates it, and neither is quantified. All quantities EST.

rates exceeded on their own. It is replaced here, and the replacement changes the central result of Section 2, which is recorded in Appendix B and in Section 2.4.

Table 8: The AI services revenue series, 2023 to 2026, rebuilt from reported figures. Convention: annualized run rate at the last reported date in each year. The series is global; Table 7 applies the US-share correction to it, and Section 2.2 uses the global figure as its output anchor O_0.

Provider	2023	2024	2025	2026
OpenAI	1.6	5.5	19.0	70.0
Anthropic	0.1	1.0 (EST)	9.0	65.0
DeepSeek, xAI, Mistral, Cohere, Z.ai, Moonshot, MiniMax	–	0.6	1.7 (EST)	5.4
Series total, $B	**1.7**	**7.0**	**29.7**	**140.4**
Year-on-year multiple		**4.15x**	**4.21x**	**4.73x**

Points are compiled by Epoch AI [3, 5] from company disclosure and reporting by The Information, Bloomberg, CNBC, the Financial Times and Reuters; each carries its original source there. MEASURED where a figure is reported, EST where interpolated between reported points. **The series counts AI services revenue and excludes accelerator and data-center hardware entirely**, which is the definition Section 2.6 requires of the measurement it specifies, and which Version 3.3 did not apply to its own anchor. **Hyperscaler AI services revenue is also excluded**: Microsoft alone disclosed a $37B AI run rate in its third fiscal quarter of 2026, but an unknown share of it is OpenAI revenue already counted here, and separating the two is the attribution problem Section 2.6 names as the binding difficulty. **Every level in this table is therefore a lower bound.** The 2026 column uses the latest reported point for each provider, between July and September 2026; a December figure would be higher, which makes the 2026 multiple conservative. **This series replaces the $110B anchor of Versions 3.1 to 3.4**, which was carried forward with a citation to a debt-issuance source that could not support it, and which was below the reported run rate of either of the two largest providers alone.

The series is a coarser instrument than the consumer-surplus estimate anchoring the denominator, which is why the finding is stated at the level and not the decimal. **The level is the finding, and the level is robust across every plausible value of its one free input.** Section 18 names the filing-grade revenue series that would tighten it, and the two largest frontier labs moving to public registration will supply it.

One prior is withdrawn, and the withdrawal strengthens the result. The hypothesis behind this measure was that K is falling, that competition and open weights compete surplus away faster than any layer can capture it. K rose, from 0.16 to 0.36 on the rebuilt series. **That rise is not evidence about capture.** It reflects a numerator growing from a small base faster than the surplus estimate, over two observations, and two points do not establish a direction. **The level is the durable finding; the direction is not established by this measurement and the Version 3.3 claim that it was is withdrawn.** The complementary claim, that the share accruing to complements is rising, is carried independently by the divergence of Section 1.4, where fixed-capability price falls 5.2 to 56 times faster than frontier price.

Falsification conditions for the law. The condition published in Version 3.0 compared the complement basket against the intelligence index. Section 1.5 establishes that the basket cannot be constructed, so that condition was never evaluable, and the two that replace it are. It is replaced by two conditions built from series this memorandum already uses:

1. **The law is abandoned if, for two consecutive annual measurements, the price of a fixed capability falls no faster than the price of the frontier.** The divergence of Section 1.4, currently 5.2x to 56x, is the entire content of the commoditization claim. If it closes to 1.0 or inverts and stays there for two annual readings, intelligence is not commoditizing faster than it is advancing and the migration has no engine.
2. **The law is abandoned if a named Tier 3 complement attracts new entry above 3 GW at a price below 90 percent of its administered cap.** This is the Section 1.6 test, tracked as signpost S20. A binding constraint that draws new supply at an unexceptional price was never binding, and the tier ordering that the whole method rests on would be wrong.

Both are evaluable on published data at a stated cadence, which the withdrawn condition was not.

2 The Elasticity Gap

The method has so far been used to answer questions. It can also be used on the answers, and turning it on the published forecasts of this field produces the memorandum's most consequential result. It is presented before any verdict because it bounds what every verdict that follows can claim.

2.1 The parameter every output forecast is betting on

Every forecast of AI's contribution to output assumes a relationship between deployed compute and attributable output. Write it:

$$O = k \cdot C^{\alpha}$$

where O is AI-attributable output, C is deployed AI compute measured in efficiency-adjusted terms, and α is the elasticity of the first with respect to the second. A forecast asserts an output figure; the constraint set permits a quantity of compute; α is what connects them. This is a reduced form and is used as one: it is the minimum structure needed to convert a published output claim into a compute requirement, which is the only operation this section performs on it.

Theory bounds this parameter before any data is consulted, and the bound is the first result of this section. Under competitive factor pricing the output elasticity of an input equals that input's share of the output it helps produce. Compute dominates the cost structure of AI services, which puts its share of AI-attributable output on the order of 0.4 to 0.8 and α below one under constant returns. **An α materially above one is not a large number; it is a claim about nonrivalry**, that a model trained once serves many users at near-zero marginal cost, which is the Romer and Jones machinery [79] and requires the increasing-returns apparatus that comes with it. The practical consequence is sharp: **any series that reports α comfortably above one is either measuring nonrivalry or is mismeasured**, and Section 2.3 shows which of the two is happening in the only series available.

What is new here is narrower than the slogan and more useful. Elasticities of output with respect to information capital have been estimated for decades: Jorgenson, Ho and Stiroh [81] estimate them as factor shares, Nordhaus [77] estimates the output elasticity of information capital with standard errors in the course of testing for a computing-driven singularity, and Aghion, Jones and Jones [78] make the substitution elasticity between automated and non-automated tasks the entire content of their growth result. Epoch's GATE model [46] states compute-output parameters explicitly, and is the one published AI forecast in this memorandum's bibliography that does. **What no published forecast of AI's aggregate output contribution reports is the elasticity of AI-attributable output with respect to deployed physical compute capacity**, which is the quantity the physical constraint set actually governs. Six major forecasts are run against that quantity below. None of them states it, none states the compute convention, output anchor or efficiency channel that Table 11 shows decide the verdict, and on the rebuilt series of Section 1.8 the range it occupies clears all six, which is the reversal Section 2.4 reports.

2.2 Every major forecast, run through the engine

Anchoring on 2026 at approximately $140B of AI services output, the series Section 1.8 publishes provider by provider from reported run rates and which excludes hardware entirely, on 29.6 GW of [3, 54] AI data-center capacity, taking the permitted incremental capacity to 2030 as approximately 200 GW [29], and applying an energy-efficiency correction of roughly 40 percent per year [3], which carries a stated substitution: the engine deflates gigawatts and therefore requires FLOP per watt, while the cited source reports price-performance in FLOP per dollar, and the two diverge through fab cost, yield and amortization, the slack ratio for each published forecast is a function of α alone.

The engine is published here in closed form, so that every cell below can be checked with a calculator:

$$\sigma_i(\alpha) = \frac{g}{(R_i/O_0)^{1/\alpha}}, \qquad g = \frac{200 \times 1.4^4}{29.6} = 25.96$$

where $O_0 = \$140.4$B is the 2026 services output anchor of Section 1.8, R_i is forecast i's required 2030 AI-attributable output, and g is the permitted growth in efficiency-adjusted compute: 200 GW of permitted new capacity, each gigawatt delivering $1.4^4 = 3.84$ times the compute of a 2026 gigawatt, against a 29.6 GW base. **A forecast survives where $\sigma \geq 1$.** Setting $\sigma = 1$ and solving gives each forecast's threshold directly, $\alpha_i^* = \ln(R_i/O_0)/\ln g$, which is the last column of Table 9 and the most useful number in this section.

Table 9: Published forecasts of AI's output contribution against the Tier 3 constraint, as a function of the elasticity α. Cells give σ, permitted over required, on the services anchor $O_0 = \$140.4$B of Section 1.8. The last column is the elasticity each forecast requires.

Forecast	α=0.5	0.7	1.0	1.2	1.4	1.7	α^*
PwC 2017, +$15.7T by 2030 [41]	0.00	0.03	0.23	0.51	0.89	1.62	1.45
Goldman 2023, +7% global GDP [38]	0.00	0.05	0.34	0.70	1.17	2.02	1.33
McKinsey 2023, $4.4T annual [40]	0.03	0.19	0.83	1.47	2.22	3.42	1.06
This memorandum, upper $10.5T	0.00	0.05	0.35	0.71	1.19	2.05	1.32
This memorandum, lower $7.0T	0.01	0.10	0.52	1.00	1.60	2.61	1.20
Acemoglu 2024, 0.66% TFP over ten years [37] (adjusted: under 0.53%)	0.18	0.74	2.16	3.26	4.39	6.01	0.76

Every cell EST, inheriting the weakest label among the output anchor, the capacity anchor and the efficiency correction. Every cell is computed from the engine of Section 2.2 on the services anchor, so the table and the engine cannot disagree. **The four published claims are different objects**: PwC's is a 2030 annual contribution to global GDP, McKinsey's an annual flow of value added, Goldman's a ten-year level shift on a 2030 global nominal GDP of $154T, Acemoglu's a ten-year TFP increment. Acemoglu's conversion to $1.69T is the one not reproduced here and is listed as needing author input. **A forecast's α^* is what it requires; none of the six states an α of its own**, and the thresholds are arithmetic on their published claims rather than a parameter any of them reports.

Every σ rises monotonically in α, so every forecast is a bet that the elasticity clears a floor, and the forecasts differ only in how high a floor they need. The six floors span 0.76 to 1.45, an interval of 0.69. **The two available series span 0.77 to 1.26 and 1.11 to 1.99 and give opposite verdicts**, which is the structural fact this section exists to establish. Table 12's bias ledger adds the sharper form of it: three of the four uncorrected biases on the services series run in the forecasts' favor and the one that runs against them is unmeasured.

2.3 Measuring it, and the convention that decides the answer

The elasticity is computed from AI-attributable output against deployed capacity. **The convention used to measure compute changes the answer, and no forecast in this literature states which convention it is using.** Compute can be counted as physical capacity in gigawatts, or as efficiency-adjusted compute, where a 2026 gigawatt delivers roughly 1.4 times the compute of a 2025 one. The engine above permits capacity and then credits it with four years of efficiency gains, so **it runs on efficiency-adjusted compute, and the elasticity applied to it must be measured the same way**.

Version 3.1 and Version 3.2 published a raw-capacity column and ran the engine on effective compute, measuring the parameter in one unit and applying it in another. That correction was made in Version 3.3. **Version 4.0 makes a larger one**: the series itself is replaced, because the anchor every reading was built on was unsourced and too low, as Section 1.8 records.

The correction removes a unit inconsistency; it does not bring the result into line with theory, and this section no longer claims it does. Section 2.1's factor-share bound of 0.4 to 0.8 applies to the output elasticity of one input in a

Table 10: Observed values of the elasticity α on the rebuilt services series of Section 1.8. The engine of Section 2.2 deflates compute; the right two columns deflate output as well, by the frontier price series of Section 1.4.

			output	output deflated	
Period	Output	Capacity	nominal	at 1.34x	at 2.25x
2023 to 2024	4.15x	2.57x	1.11	1.34	1.74
2024 to 2025	4.21x	2.11x	1.33	1.60	2.08
2025 to 2026	4.73x	1.56x	1.99	2.36	3.03
Range			**1.11 to 1.99**	1.34 to 2.36	1.74 to 3.03

All values EST, inheriting the weakest label in the series. Effective compute applies the 40 percent annual efficiency correction [3] the engine applies forward; the deflated columns apply Section 1.4's frontier price series to the numerator as well, which is the symmetric treatment. **The three readings rise monotonically in every column**, by a factor of 1.79 on the nominal convention and by similar factors on both deflated ones. **That is the finding, and it is not a band.** Three readings that rise monotonically are not noisy draws from a constant; the reduced form $O = k \cdot C^{\alpha}$ assumes a constant α, and this series declines to supply one.

multi-input production function, while every column of Table 10 is the reduced form $O = k \cdot C^{\alpha}$ with no other input in it, which Section 2.6 states recovers scale rather than an output elasticity. **The two objects are not comparable.**

2.4 What the table says

The two series available give opposite verdicts, and that is the result of this section.

The hardware-inclusive series fails four of six forecasts, and it is an accounting identity. Its numerator is built largely from accelerator revenue, so $O \equiv p_C \cdot C$ and $\alpha \equiv 1 + \mathrm{d}\ln p_C / \mathrm{d}\ln C$ exactly, and every deviation of the measured elasticity from one is a movement in the price of compute rather than a fact about production. Versions 3.1 to 3.4 of this memorandum ran on a series described that way and published the four eliminations.

The services series fails none, and it divides lab revenue by all AI data-center capacity. The numerator is the revenue of frontier labs and AI-native providers; the denominator is every gigawatt of AI data-center capacity, most of which serves hyperscalers, enterprises and internal workloads whose output is not in the numerator. **A rising lab share of that capacity raises the measured elasticity with no change in production whatever.**

Neither series adjudicates a forecast. One is a price index wearing the numerator's clothes; the other is a mix shift. **The thresholds are the durable result**: 0.76 to 1.45, arithmetic on published claims, invariant in order to both anchors by construction, and **what every forecast in this literature is silently assuming while none of them reports it**. **The dataset of Section 2.6 decides**, and nothing available now does.

The reduced form of this memorandum's own engine is what the rising readings refuse. On the services series the elasticity reads 1.11, 1.33 and 1.99, rising monotonically by a factor of 1.79, and on both deflated conventions too. **That is a statement about $O = k \cdot C^{\alpha}$ with a fixed exponent, which is this engine's reduced form**, and not about the six forecasts, which state no α at all. A parameter that moves like that cannot be represented by a constant, and this engine represents it with one.

A reductio that settles what the services series is measuring. At $\alpha = 1.99$, the top of the observed range, the engine permits $O_0 \cdot g^{1.99} = \$140.4\text{B} \times 652 \approx \$92\text{T}$ of 2030 AI-attributable output. **The claim that AI output exceeds a majority of 2030 global GDP requires $\alpha = 1.94$, which sits inside the observed range.** Section 4.1 rejects that claim on independent grounds by a factor of 4.8. **A series that cannot eliminate a claim this memorandum rejects elsewhere by 4.8 times is not measuring a production elasticity.** It is measuring diffusion and mix: the rise from a \$1.7B base in 2023 to \$140.4B in 2026 is adoption arriving at capacity already built, which is the signature of Tier 1 absorption rather than of returns to compute.

2.5 What the result survives, and what would overturn it

The conclusion is robust to every modeling choice this section makes, and the sensitivities are published as results. In Version 4.0 the conclusion that is robust is a negative one: the capacity constraint does not bind.

Three inputs remain unsourced, and all are named so a reader can attack them. The first is the reading of the capacity figure. The 200 GW is Bain's [29] projection of compute added to 2030, and the engine above reads it as total permitted 2030 capacity. **Read strictly as an increment it implies 229.6 GW and $g = 29.80$, which is a row**

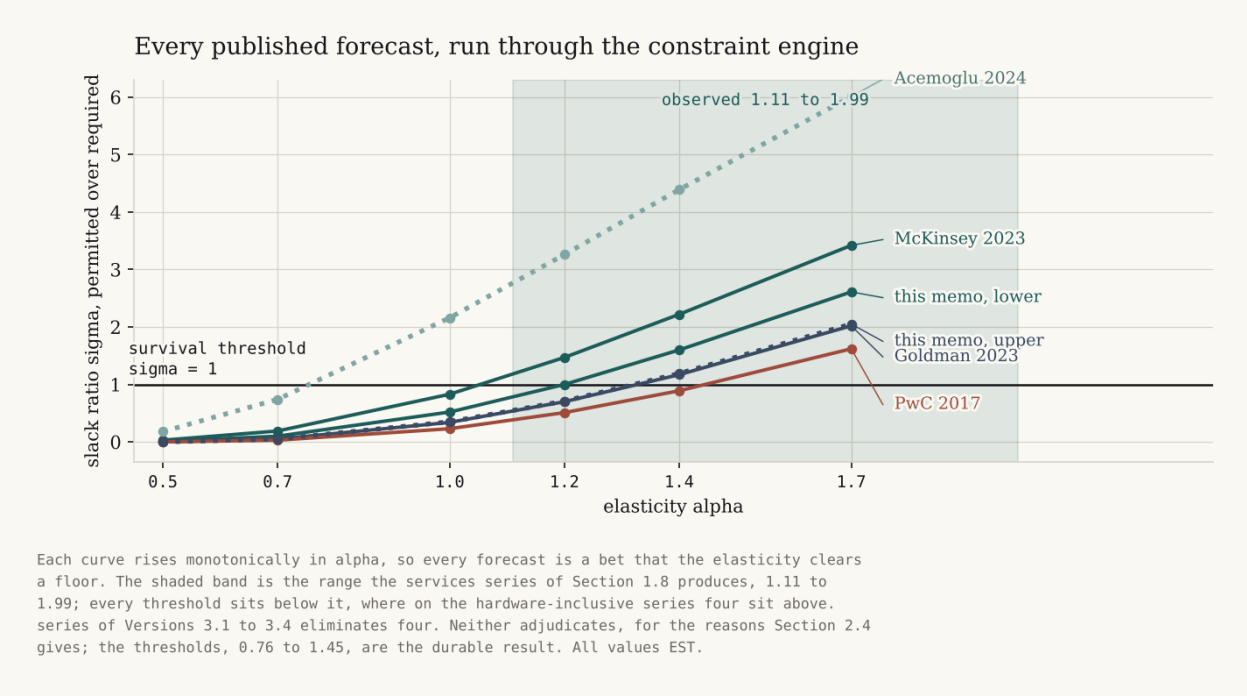


Figure 9: Every published forecast run through the constraint engine as a function of the elasticity, on the services anchor of Section 1.8. The shaded band is the observed range that series produces, 1.11 to 1.99. Every threshold sits below it, where on the hardware-inclusive series of Versions 3.1 to 3.4 four sit above. **The two series give opposite verdicts and neither adjudicates, for the reasons Section 2.4 gives.** All quantities EST. Reproduce from the formula in Section 2.2.

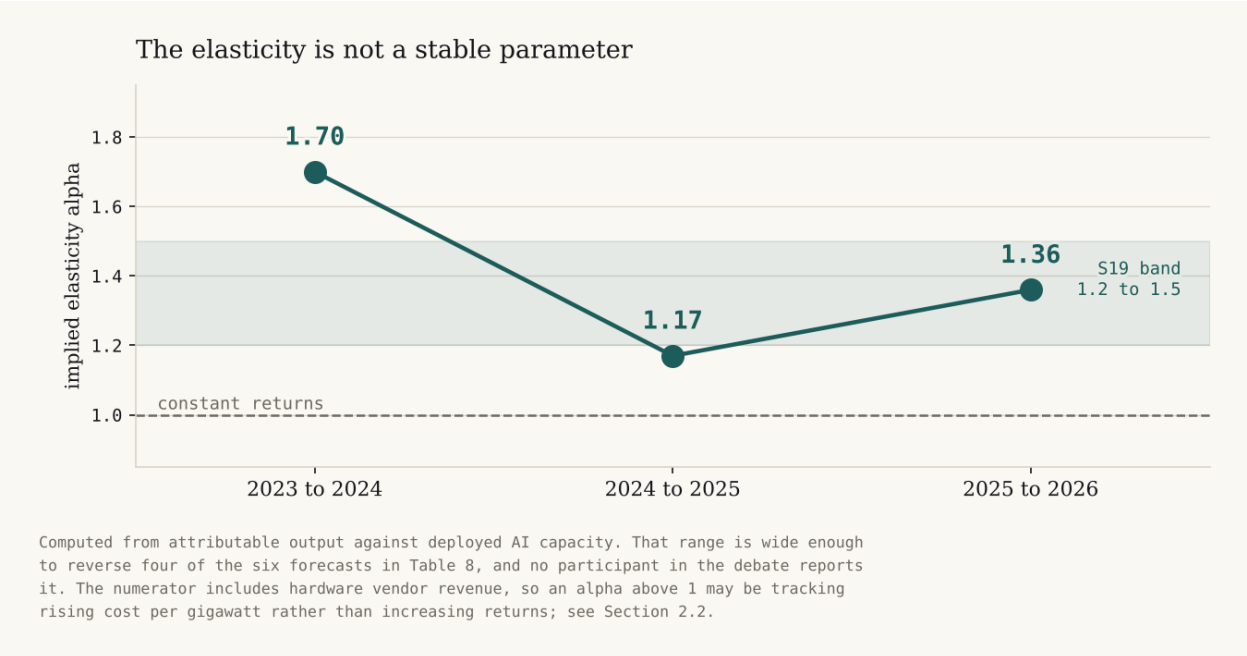


Figure 10: Each mark is the elasticity a published forecast requires to survive the Tier 3 capacity constraint, from the final column of Table 9. The six thresholds span 0.76 to 1.45 and are invariant in order to both anchors by construction. **They are the durable result of this section**, where the two available series give opposite verdicts and neither adjudicates. All values EST.

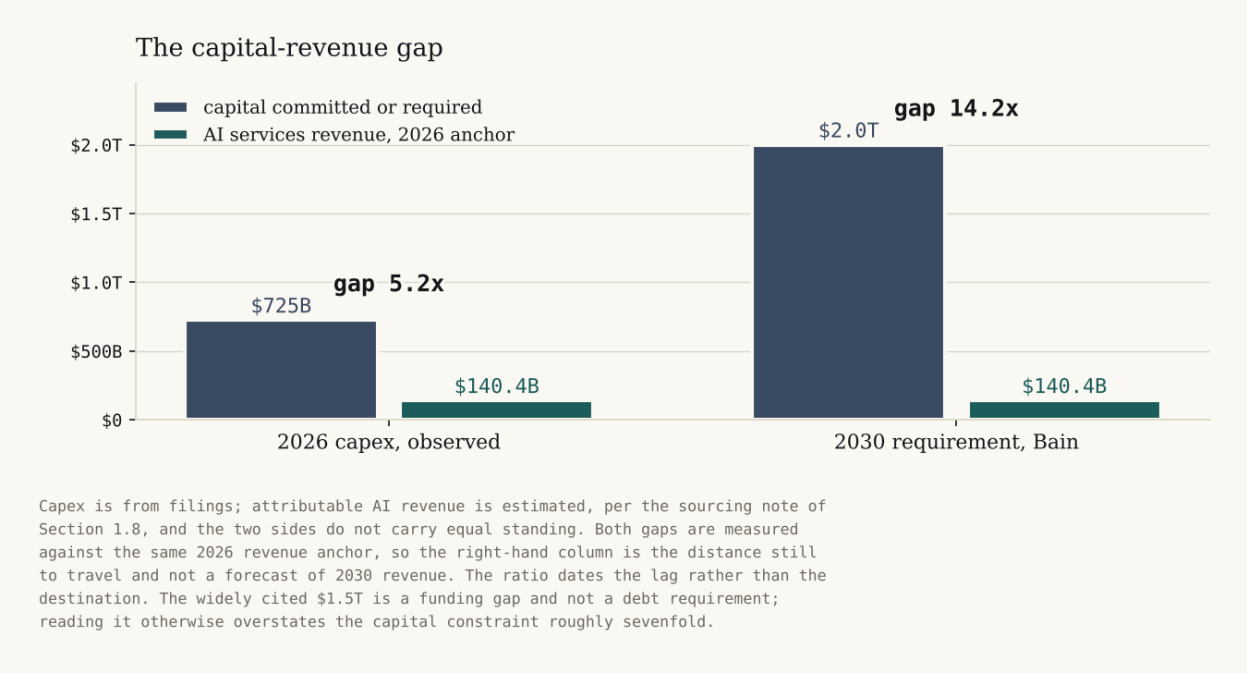


Figure 11: The capital-revenue gap, which is what the elasticity question is ultimately about. Capex is from filings; attributable AI revenue is estimated, per the sourcing note of Section 1.8, and the two sides do not carry equal standing. The ratio dates the lag rather than the destination. The widely cited $1.5T [27] is a funding gap and not a debt requirement; reading it otherwise overstates the capital constraint roughly sevenfold, on the one constraint least likely to bind. All quantities EST.

Table 11: Every compute-side convention that could reasonably be chosen differently, each applied to both sides. **Each row carries its own band**, because a convention that changes the forward multiple changes the measured elasticity through the same channel.

Choice	g	Band	Fails throughout	Survives somewhere
Adopted: effective compute, 200 GW total, $C_0 = 29.6$ GW	25.96	1.11 to 1.99	**none**	all six
Efficiency factor 1.3 rather than 1.4	19.30	1.18 to 2.20	none	all six
Efficiency factor 1.5 rather than 1.4	34.21	1.05 to 1.83	none	all six
Permitted capacity 150 GW	19.47	1.11 to 1.99	none	all six
Permitted capacity 250 GW	32.45	1.11 to 1.99	none	all six
$C_0 = 25$ GW	30.73	1.11 to 1.99	none	all six
$C_0 = 35$ GW	21.95	1.11 to 1.99	none	all six
200 GW read as incremental, total 229.6 GW	29.80	1.11 to 1.99	none	all six
Output deflated at 1.34x per year	25.96	1.34 to 2.36	none	all six
Output deflated at 2.25x per year	25.96	1.74 to 3.03	none	all six

All rows EST. **No forecast fails in any row.** That is the result of this table, and it is a stronger statement than the eliminations it replaces, because it holds across the full range of every input the engine runs on and not at one setting of them. **The invariance has a reason.** The ordering of the thresholds is fixed by construction: each is $\ln(R_i/O_0)/\ln g$ with a common g and O_0, so no choice of either can reorder them. And when a convention is applied to both sides, the efficiency factor moves the permitted multiple and the measured band through the same channel, so they travel together. **The levels are sensitive and the comparison is not**, which is why this section reports the comparison. **Two inputs remain unsourced and are listed under needs author input**: the 40 percent efficiency correction is cited to a price-performance series in FLOP per dollar where the engine requires energy efficiency in FLOP per watt, and the 200 GW is a demand projection used as a supply ceiling. The first three rows and the eighth are the disclosure of the first; the fourth and fifth are the disclosure of the second.

of Table 11; on the services series no forecast fails on either reading, and on the hardware-inclusive series the incremental reading is what lets this memorandum's own lower bound survive. The strict reading is the more literal one and is carried to the needs-author-input list pending the underlying Bain definition; the adopted reading is retained because it is the conservative one against this memorandum's own conclusion. The second is the level. Section 5.2 of this memorandum separately finds that the Tier 3c transformer and switchgear constraint cannot deliver capacity at that pace, which means 200 GW is generous to every forecast in Table 9, including this memorandum's own. **Using a number that favors the claims being eliminated is the correct direction of conservatism**, and halving it to 100 GW raises g to 12.98 and every threshold by roughly 0.4, on which no forecast fails on the services series either; the two series still disagree, which is the point of Section 2.4.

This memorandum's own band is no longer eliminated, and the reversal is reported first because the method was run against its own conclusion before anyone else's. Versions 3.1 to 3.4 reported its lower bound failing. On the rebuilt series its threshold is 1.20 against an observed range starting at 1.11, and it survives, as does every other forecast in Table 9. **The elimination this memorandum published against itself for four versions was an artifact of its own unsourced anchor**, and the correction removes it along with the others. **The result that would damage the author's disclosed positions is an estimated elasticity at or above 1.4**, which clears most of Table 9 and retires the constraint thesis, and it is named so the pre-registration runs in both directions. **Disclosure is not mitigation, so four commitments come with the ask**: the analysis plan, the attribution rule and the falsification thresholds registered with a timestamp before any estimation; the attribution rule applied by the data holder, blind to the thresholds; all estimation code and constraint series released whatever the result; and a replication run by a party with no position in it.

The bias ledger. Four biases on the services series are uncorrected, and three of the four run in the forecasts' favor. They are listed with direction rather than magnitude, because none of them is quantified.

Table 12: Uncorrected biases on the services series, and the direction each one pushes the verdict for the six forecasts of Table 9.

Bias	Direction	Mechanism
Output is nominal and AI prices fall at 1.34 to 2.25x per year (Section 1.4)	**favors**	Deflating the numerator raises the measured elasticity, as Table 10's right columns show
The 2026 column is a partial-year run rate, taken between July and September	**favors**	A December figure would be higher, raising the 2025 to 2026 multiple and the elasticity with it
A GDP contribution exceeds the revenue of the firms producing it	**favors**	R_i measured against a revenue anchor overstates the required growth, so every threshold is an upper bound
The lab share of AI data-center capacity is rising	**against**	The denominator includes capacity serving output not in the numerator; a rising lab share raises the measured elasticity with no change in production. **Unmeasured**

All EST. **The three that favor the forecasts are directional and the one that runs against them is unmeasured**, which is why Section 2.4 reports that neither series adjudicates rather than reporting the services verdict as a finding. The fourth is the one the dataset of Section 2.6 is specified to measure.

2.6 The measurement that closes the question, specified to build

The question resolves the moment one dataset exists, and this section specifies it precisely enough to construct. A panel of AI services output against deployed compute, at firm or sector level, with hardware vendors excluded from the numerator.

1. **The numerator counts output attributable to AI services**, not revenue from selling the means of production. **Realized revenue is the estimable form and a surplus-inclusive numerator is the correct object**, and the two are not interchangeable: surplus estimates [52] are annual, survey-based and not firm-attributable, so they do not exist at firm-quarter frequency, and adding a usage-driven surplus term to the left side would mechanically pull the estimate toward one. **Estimate on revenue, report the surplus-inclusive anchor as a stated scaling, and treat the gap between them as the bound it is**, which is the same gap the surplus row of Table 11 opens. This excludes the accelerator and memory vendors whose revenue dominates the only series now available. **Without this the specification is an identity, not a production relationship**: if output is measured as accelerator revenue, then $O \equiv p_C \cdot C$ and $\alpha \equiv 1 + \mathrm{d}\ln p_C / \mathrm{d}\ln C$ exactly, so every deviation of the measured elasticity from one is a

movement in the price of compute. **That is the real defect in the available series, and it is a specification defect rather than a data-availability one**, which is why no amount of additional revenue data repairs it.

2. **The denominator is deployed compute, not shipped or announced compute, and capacity is not compute.** Utilization, cooling overhead and the accelerator share of facility power all sit between gigawatts and delivered compute, and every forecast reviewed in Table 9 sets them to one by silence. The gap between deployed, shipped and announced is large, moving, and is itself one of the constraints this memorandum tracks. Three proxies are available and each carries a stated bias: utility interconnection filings, biased toward new build; facility-level power draw, biased by cooling overhead and non-AI load; and chip shipment registrations, biased by a shipped-to-deployed lag that is correlated with the growth rate and so is not classical measurement error, which no instrument repairs.
3. **The unit of observation is the firm or the sector, not the economy, and the series must be deflated.** An aggregate series cannot separate a rising elasticity from a changing mix of firms, and the mix is changing fast enough to dominate. **Price per token at fixed capability has fallen by more than an order of magnitude per year**, so a nominal numerator understates real output growth and biases the estimate down, and the same bias is present in the three readings of Table 10.

Identification, and the instrument this memorandum already maintains. The estimating equation is $\ln O_{it} = \ln k_i + \alpha \ln C_{it} + \beta' \ln X_{it} + \delta_t + \varepsilon_{it}$, with X carrying labor and non-compute capital, since a single-input specification recovers scale rather than an output elasticity. Ordinary least squares on it is biased, because firms buy compute in anticipation of output. **This is the simultaneity problem of production-function estimation and it has a literature** [82] which any serious attempt must satisfy. **The exclusion restriction is available and this memorandum already publishes the series that supply it.** Deployed compute at the firm level is shifted by physical delivery constraints that are orthogonal to that firm's own demand: advanced-packaging allocation at Tier 3a, HBM supply allocation, interconnection queue position at Tier 3d, and transformer and switchgear lead times by utility territory at Tier 3c. **A shift-share design on firm exposure to those constraints, interacted with national shocks to each input's availability, gives variation in deployed compute driven by what could physically be delivered rather than by what the firm expected to earn**, and the constraint series this memorandum maintains are its first-stage inputs. **Two difficulties come with that design and are stated rather than assumed away.** First, it must pick a branch: exogenous shares after Goldsmith-Pinkham, Sorkin and Swift [83], under which vendor allocation must be argued unrelated to each firm's own expected demand, which is least plausible for packaging and HBM, where suppliers allocate partly on contracted volume, and most plausible for utility-territory lead times; or quasi-random shocks after Borusyak, Hull and Jaravel [84], which needs many shocks, where four categories is not many, and under which inference over-rejects severely [85]. Second, the time effects this panel requires absorb the national shocks, so identification falls back on cross-firm exposure and therefore on the first branch. **What the design can detect, derived rather than asserted.** $\mathrm{SE}(\hat{\alpha}) \approx \sigma_\varepsilon \sqrt{\mathrm{DE}} / (\sqrt{NT}\, \sigma_{\tilde{C}}\, \rho)$. At $N = 25$ firms, $T = 14$ quarters, $\sigma_\varepsilon = 0.25$ for within-firm log AI services revenue, $\sigma_{\tilde{C}} = 0.35$ for log deployed compute net of firm and time effects, a first-stage partial $\rho^2 = 0.15$, and a design effect of 2 for clustering on 25 firms with serial correlation, $\mathrm{SE} \approx 0.14$ and the minimum detectable difference at 80 percent power and 5 percent size is $2.8\,\mathrm{SE} \approx 0.39$. **That separates 0.8 from 1.4. It does not separate 1.00 from this memorandum's own threshold of 1.20, nor 1.4 from McKinsey's 1.06, so it adjudicates neither this memorandum's band nor the disclosure threshold below.** Every input is a stated assumption and each is the first thing to argue about; that limit belongs in any collaboration agreement before the work starts.

What this costs and what it settles. The universe is on the order of fifteen to forty firms over twelve to sixteen quarters, which is enough for the design above and not enough for anything that discards the instrument. **The binding difficulty is the attribution rule**, the share of a diversified firm's revenue counted as AI-attributable, which must be stated and defended before anything is estimated. **This memorandum stakes a falsification on the result**: an estimated α below 1.20 obtained on this design eliminates this memorandum's own \$7 to \$10.5T band, which is the band Section 4.1 defends, and the register will record it. Signpost S19 tracks the series and is marked dormant until it exists.

This is the most valuable open measurement in the economics of AI, it is an engineering problem with a known solution, and the organization that builds it holds the number on which every forecast in Table 9 turns.

3 The Five Roots

Five root uncertainties sit beneath the thirty questions carried here, and every branch question inherits its verdict from one or more of them. Absorption at Tier 1 binds two of the five, R1 and R2, which is the empirical discovery that put it in the constraint set.

3.1 R1. Will AI create more value than it destroys?

Candidates. (i) net destruction; (ii) frictionless broadly-shared abundance; (iii) net-positive but uneven and lagged; (iv) neutral reallocation of profit among firms; (v) net-positive in value but net-negative socially; (vi, taboo) a financing cascade destroys more capital than the technology creates.

Table 13: R1 test matrix: candidates against the constraint set. Will AI create more value than it destroys?

Candidate	Binding constraint	Tier	Verdict
(i) net destruction	Capital: every prior general-purpose technology delivered positive contribution once diffused	5	Repriced to ∼0
(ii) frictionless abundance	Absorption: decade diffusion; ∼5% of integrated pilots P&L-positive	1	Repriced to ∼0
(iii) net-positive, lagged	none	–	Survives
(iv) neutral reallocation	Data: additive new categories exist	2	Repriced to ∼0
(v) value positive, social negative	Law and Legitimacy: live if safety nets fail	4	Live tail
(vi, taboo) financing cascade	Capital: plausible tail, not central	5	Live tail

Computed. (i) requires the aggregate productivity contribution to be negative. Electricity and IT each added on the order of 1.5 percentage points per year at peak. Requires contribution below zero; permits above zero on all precedent. Sign-level gap. (ii) requires near-zero adoption lag. Requires diffusion under approximately 2 years; Absorption permits a decade at $r = 0.02$ to 0.05. $T_{\text{req}} \approx 5$ to 10 years against 2 required. This is the verdict that most clearly needed Tier 1: it was previously attributed to "Time," which named the symptom rather than the constraint. (vi) is not eliminated. Morgan Stanley's approximately $1.5T [27] is a funding gap through 2028 closable by bonds, private credit, asset-backed structures or equity, not a debt requirement; the corresponding corporate debt issuance estimate is approximately $200B, and 2026 hyperscaler borrowing is expected above $400B. **Reading that figure as a debt requirement, as most commentary does, overstates this constraint roughly sevenfold.** Read correctly, the cascade tail sits at 0.07.

Isolate. (iii) survives with (v) and (vi) as tails. A J-curve: destruction front-loads while creation lags. The capital gap dates the lag: approximately $725B [28] of 2026 capex against approximately $140.4B of AI services revenue (Section 1.8) is a ∼5.2x current gap, and ∼14.2x against Bain [29]'s ∼$2T maturity requirement.

Scenario. By 2027 the gap persists and (vi) risk is elevated but unrealized. By 2029 revenue narrows the gap. By 2030 net-positive value is realized unevenly. Probability: (iii) ∼0.73; (v) ∼0.15; (vi) ∼0.07.

Bottom line. Net-positive, with a J-curve shape and a ∼5.2x present capital gap that dates the lag rather than the destination. The binding constraint on the shape is Absorption at Tier 1, not capital.

3.2 R2. How fast do organizations become AI-native?

Candidates. (i) overnight, two to three years; (ii) never; (iii) a decade-long grind gated by trust, data, integration; (iv) bimodal; (v) already largely done; (vi, taboo) a disillusionment retreat.

Table 14: R2 test matrix: candidates against the constraint set. How fast do organizations become AI-native?

Candidate	Binding constraint	Tier	Verdict
(i) overnight	Absorption: ∼5% of integrated pilots value-positive; Data: integration debt	1, 2	Repriced to ∼0
(ii) never	Capital: ROI forces adoption; purchased tools succeed ∼67%	5	Repriced to ∼0
(iii) decade grind	none	–	Survives
(iv) bimodal	none	–	Survives
(v) already done	Absorption: only ∼5% value-positive	1	Repriced to ∼0
(vi, taboo) retreat	Absorption: live transient, not permanent	1	Live tail

Computed. (i) requires broad P&L-positive deployment now. Requires majority value-positive; permits approximately 5%. Gap approximately 20x on the success base. At $r_{\text{absorption}} = 0.02$ to 0.05, closing a 20x gap requires $T_{\text{req}} = \ln 20 / \ln 1.035 \approx 87$ years at the midpoint, which overstates the case because the gap is not purely a stock expansion;

the honest reading is that Absorption cannot deliver a 20x improvement in integration success within the horizon by any observed mechanism. (ii) requires zero adoption despite returns; a shadow-AI economy across more than 90% of firms falsifies it.

Isolate. (iii) and (iv) survive together. This verdict is the empirical discovery of the Absorption constraint. The bottleneck here is organizational rather than technical, which is precisely what Tier 1 records. A constraint set without Absorption would reach this finding and have nowhere to put it.

Scenario. By 2027 a disillusionment trough coincides with rising capability. By 2029 the value-positive base widens. By 2030 AI-native is normal for entrants and a retrofit problem for incumbents. Probability: (iii/iv) ∼0.70; (vi) ∼0.20.

Bottom line. A decade-long uneven grind. The binding constraint is Absorption at Tier 1, and the 95%-no-return rate is the number that dates it.

3.3 R3. Who controls compute, energy, data, and models?

Candidates. (i) a single monopoly; (ii) full decentralization; (iii) a US-led oligopoly with a parallel Chinese stack; (iv) state seizure; (v) China overtakes; (vi, taboo) a packaging or foundry leapfrog.

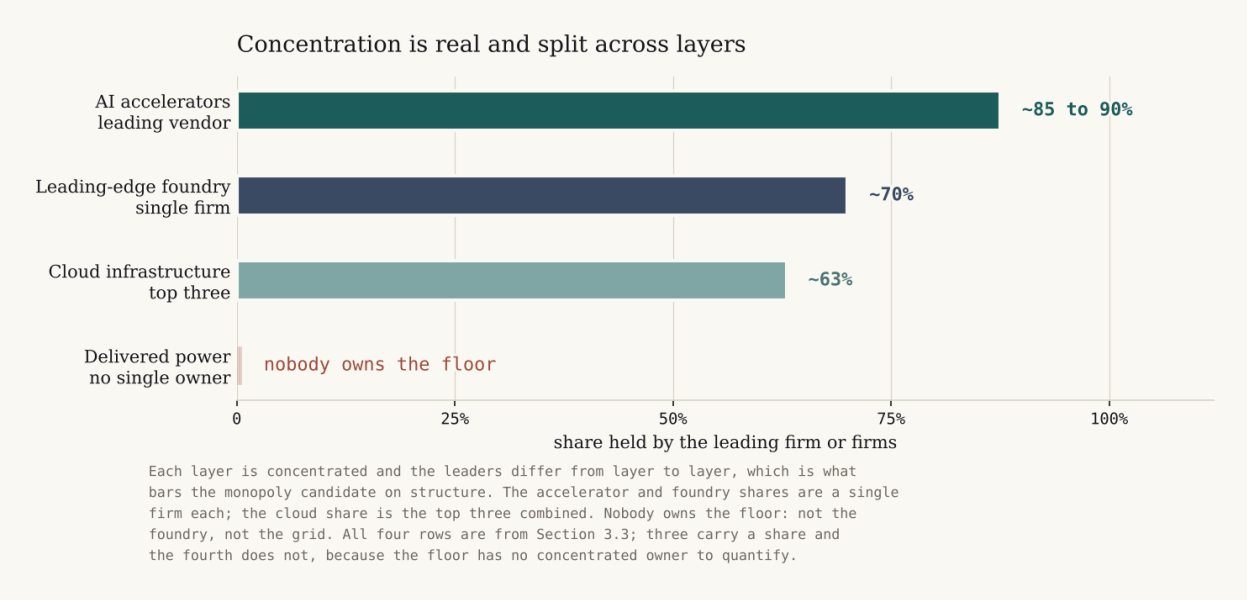


Figure 12: Concentration is real and split across layers. Each layer is concentrated and the leaders differ from layer to layer, which is what bars the monopoly candidate on structure. The accelerator and foundry shares are a single firm each; the cloud share is the top three combined. Nobody owns the floor: not the foundry, not the grid.

Table 15: R3 test matrix: candidates against the constraint set. Who controls compute, energy, data, and models?

Candidate	Binding constraint	Tier	Verdict
(i) monopoly	Matter: control split across layers; Law: antitrust and export controls	3, 4	Repriced to ∼0
(ii) decentralization	Matter 3c/3d: gigawatt-class delivered power; Capital: ∼$725B/yr	3, 5	Repriced to ∼0
(iii) US-led oligopoly plus walled China	none	–	Survives
(iv) state seizure	Capital: the state is not funding the buildout	5	Repriced to ∼0
(v) China overtakes near-term	Matter 3b: ∼5x lead widening toward ∼17x, HBM-gated; Talent	3, 1	Repriced to ∼0
(vi, taboo) foundry leapfrog	Matter: live long tail	3	Live tail

Computed. (i) requires one actor to own accelerators, foundry, cloud, and models. NVIDIA holds ∼85 to 90% of data-center accelerators, TSMC ∼70% of foundry, the top three clouds ∼63%, and several labs hold frontier models.

Barred on structure. (ii) requires distributed compute to match hyperscale training, needing gigawatt-class delivered power. **Re-attributed from Energy to Matter 3d:** the constraint is not physics but the interconnection queue, whose median exceeded 5 years for projects reaching commercial operation in 2025. (v) requires closing a $\sim$5x chip-capability gap projected to widen toward $\sim$17x by 2027. Talent is now named as a co-constraint: approximately 47% of top-tier researchers were trained in China and roughly 72% of them work in the United States, which is a Tier 1 asymmetry that no capital deployment reverses inside a decade.

A measured decoupling of compute scale from frontier quality narrows the verdict, and the measurement is the finding. The elimination of near-term parity holds for deployed compute, not for model capability. US and Chinese models traded the lead repeatedly through 2025 and 2026, and as of March 2026 the leading US model led the leading Chinese model by approximately 2.7 percent [45, 54] on aggregate benchmark performance (AI Index 2026). A compute advantage of five to seventeen times has therefore not produced a proportional capability advantage. This is substitution defeating a stock argument at Tier 3, the same failure mode Section 15 documents at Tier 2, and the measurement is published because it cuts against the consensus reading. What survives is that the compute gap constrains scale of training and deployment; what does not survive is any claim that it constrains frontier model quality.

Isolate. (iii) survives. Control is layered and US-led, with power as the deepest floor. Probability: $\sim$0.75. **Computed: candidate (ii) requires 13.3 years of Matter 3d expansion against 4.4 available, repriced to 0.1% [< 0.05 to 0.4].**

Scenario. By 2027 the US compute lead widens even as the Chinese stack matures domestically. By 2029 delivered power, not chips, is the binding control variable. By 2030 control still sits with the layered oligopoly, gated by the grid.

Bottom line. A US-led oligopoly plus a walled Chinese stack. Concentration is real but split across layers, and the true chokepoint is grid equipment and interconnection at Tier 3, not physics and not capital.

3.4 R4. Can institutions adapt without stifling innovation?

Candidates. (i) a unified global regime; (ii) a durable regulatory vacuum; (iii) fragmented, lagging, sectoral governance; (iv) an innovation-killing clampdown; (v, taboo) a race to the bottom.

Table 16: R4 test matrix: candidates against the constraint set. Can institutions adapt without stifling innovation?

Candidate	Binding constraint	Tier	Verdict
(i) unified regime	Law: EU, US and China diverging; no treaty on the horizon	4	Repriced to $\sim$0
(ii) vacuum	Law: EU Act binding, penalties to 7% of turnover	4	Repriced to $\sim$0
(iii) fragmented, sectoral	none	–	Survives
(iv) clampdown	Capital: ceding the field forbidden by competition	5	Repriced to $\sim$0
(v, taboo) arbitrage	Law: live texture	4	Live tail

Computed and updated. (i) requires convergence. The timeline has moved and the movement is itself the evidence for Tier 4. The EU AI Act's high-risk obligations, treated as fixed at August 2, 2026 for over a year, were deferred by the Digital Omnibus [58] to December 2, 2027 for Annex III systems and August 2, 2028 for Annex I, with final Council approval on June 29, 2026. Article 50 transparency obligations still apply from August 2, 2026, and Product Liability Directive transposition remains fixed at December 9, 2026 [59]. Meanwhile the United States has no federal statute, Executive Order 14365 [60] of December 11, 2025 directs an AI Litigation Task Force against state law, and states enacted 109 AI laws and 28 data-center laws in the first half of 2026 alone. **A flagship compliance deadline moved sixteen months in seven months of legislative process. No Tier 3 constraint can move like that and no Tier 0 constraint can move at all.**

Isolate. (iii) survives with (v) as live texture. Probability: $\sim$0.78. The Omnibus is direct confirmation of the fragmentation verdict.

Scenario. By 2026 Article 50 transparency binds while the high-risk regime does not. By 2027 the US still has no federal statute and the preemption litigation is unresolved. By 2029 the deferred Annex III duties are in force and sectoral divergence is the standing risk.

Bottom line. Fragmented, lagging, sectoral governance. The operative risk is divergence and arbitrage, not stringency, and the 2026 deferral is the cleanest available demonstration that law occupies its own tier.

3.5 R5. What are the real limits of intelligence, and how close are we?

Candidates. (i) imminent unlimited AGI from scale; (ii) a permanent plateau; (iii) continued rise via test-time and post-training compute [46]; (iv) a discontinuous architectural jump; (v, taboo) a jagged plateau, benchmarks climb while deployment reliability stalls.

Table 17: R5 test matrix: candidates against the constraint set. What are the real limits of intelligence, and how close are we?

Candidate	Binding constraint	Tier	Verdict
(i) scale to AGI soon	Matter 3b and 3c: another 100x of compute and power; Capital	3, 5	Repriced to $\sim$0
(ii) permanent plateau	empirical: capability still rising via test-time compute	state	Repriced to $\sim$0
(iii) rise via new axes	none	–	Survives
(iv) architecture jump	unforecastable	–	Live tail
(v, taboo) jagged plateau	none	–	Live and strengthening

Computed. (i) previously rested substantially on the data wall at Tier 2. That attribution is withdrawn. Section 15 documents the framework predicting, at maximum confidence, that the public-text wall would bind by 2026, and being wrong: Epoch's own median moved from 2024 to 2028, and multimodal and verified synthetic sources substituted. With r_{data} revised to 0.30 to 0.60, Data no longer eliminates (i). The elimination survives on Matter and Capital: another 100x of training compute requires fab and grid expansion that Tier 3 cannot deliver on the horizon at $r = 0.15$ to 0.25.

(v) has strengthened materially and is now the near-term texture rather than a tail. METR [9]'s 50 percent horizon [10] rose several-fold through 2025 and 2026, while the 80 percent horizon [9] has been approximately flat: Claude Opus 4.5 [9] measured approximately 27 minutes in December 2025, and the series had been flat at 27 to 32 minutes [9] across the preceding frontier releases. The ratio between the 50 and 80 percent horizons [9] has widened from approximately 5x to somewhere between 10x and 25x. **Capability gains are arriving as intercept rather than slope: models succeed more often on long tasks without succeeding more reliably.** This is the jagged plateau, measured.

Isolate. (iii) survives, with (iv) a live tail and (v) the near-term texture. Probability: (iii) $\sim$0.65; (v) as near-term reality $\sim$0.30.

Scenario. By 2027 the 50 percent horizon [10] extrapolates toward month-long tasks while the 80 percent horizon [9] remains in hours. By 2028 pre-training saturates against fabrication and power rather than tokens. By 2029 test-time compute is the dominant lever.

Bottom line. Capability keeps rising via test-time and post-training compute [46]. The pre-training data wall is no longer load-bearing in this verdict, and the elimination of scale-to-AGI rests on fabrication and grid, not on tokens. Reliable long-horizon autonomy remains years out, and the widening gap between the 50 and 80 percent horizons is now the strongest single datum in the memorandum.

4 Domain I. Economic Transformation

AI adds roughly 1.0 to 1.5 percentage points of annual growth, not a majority of GDP: the majority claim fails by 4.8-fold against the most bullish credible ceiling. Realized gains run at 0.1 to 0.2 [38] points against a 1.5 point [62, 38] electricity precedent. The binding constraint is organizational absorption, which relaxes on a twenty to forty year clock.

4.1 Q1. How much of global GDP will AI create?

Candidates. (i) most of GDP, above 50%; (ii) transformative, 10 to 25%; (iii) material, single-digit percent, compounding; (iv) negligible; (v, taboo) negative in a correction window.

Computed. (i) requires AI-attributable output above 50% of a $\sim$\$154T [42] 2030 economy, roughly \$77T of new output in about four years. The most bullish credible estimate is PwC's \$15.7T [41] as a 2030 annual contribution. Requires $\sim$\$77T; permits $\sim$\$16T. **Gap 4.8x, which is 0.68 orders of magnitude.** Two earlier figures are withdrawn. A draft cited $\sim$1.5 orders. Version 3.0 cited $\sim$9x and $\sim$0.94 orders, and $10^{0.94} = 8.71$: that is the ratio of \$77T to this memorandum's own band midpoint of roughly \$8.75T, not to the \$16T ceiling the sentence names. **The elimination holds comfortably on either denominator**, and the published arithmetic now matches the sentence it supports. The Data leg of this elimination is withdrawn per Section 15; the elimination holds comfortably on Capital and Matter without it.

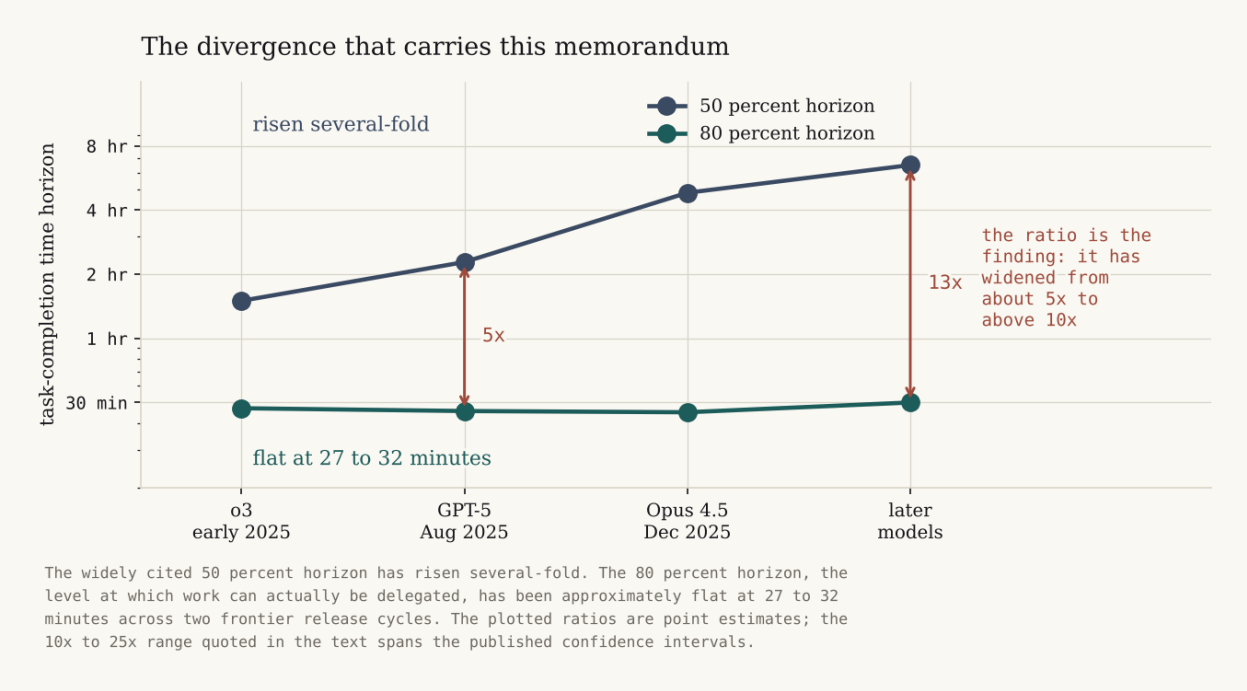


Figure 13: The divergence that carries this memorandum. The widely cited 50 percent horizon [10] has risen several-fold. The 80 percent horizon, the level at which work can actually be delegated, has been approximately flat at 27 to 32 minutes across two frontier release cycles. Capability is arriving as intercept rather than slope. Source: METR [9].

Table 18: Q1 test matrix: candidates against the constraint set. How much of global GDP will AI create?

Candidate	Binding constraint	Tier	Verdict
(i) most of GDP	Capital: needs ∼$77T new output; Matter 3b/3c/3d; Absorption	5, 3, 1	Repriced to ∼0
(ii) transformative 10-25%	Capital: exceeds the most bullish credible estimate; Absorption	5, 1	Repriced to ∼0
(iii) material single-digit	none	–	Survives
(iv) negligible	Capital: savings already measurable	5	Repriced to ∼0
(v, taboo) negative	Capital: possible in a correction window	5	Live tail

Isolate. (iii) survives. The band is 1.0 to 1.5 percentage points per year of GDP growth, squarely within the electricity and IT precedents, roughly $7 to $10.5T as a 2030 annual contribution.

This verdict is α-exposed, and this memorandum publishes the exposure against itself. Section 2 establishes that the band defended here requires an elasticity of at least 1.20 at its lower bound and 1.32 at its upper, against a measured range of 1.11 to 1.99 on the rebuilt services series of Section 1.8. **On that range the band survives, and it survives on every compute-side convention Section 2.5 tests. Versions 3.1 to 3.4 reported the opposite**, that this memorandum's own band failed, and that reversal is reported here rather than quietly dropped: the elimination was an artifact of an output anchor no source supported, and a check of the anchor against NVIDIA's fiscal 2025 10-K found it. The exposure is still real, because three readings that rise monotonically do not establish a level, so the range bounds the verdict without settling it. **The confidence coefficient prices the exposure explicitly**: $c = 0.45$ carries a 0.40 deduction for dependence on exactly this parameter, which is the largest single deduction anywhere in the rubric and cuts the stated probability from $P_{\text{engine}} = 0.93$ to 0.53. **Section 2.6 publishes the falsification**: an estimated α below 1.20 on the specified panel eliminates this band, and the register will record it. The probability below is stated at $\alpha = 1.4$, the value this band requires, so a reader can see precisely what it is betting on.

What this verdict measures, and what sits outside it by construction. This verdict forecasts measured GDP. Value delivered as consumer surplus on goods priced at zero is outside that quantity by construction, and Section 1.8 puts US consumer surplus at approximately $172B [52] annually and growing at 54 percent. A reader concluding that AI

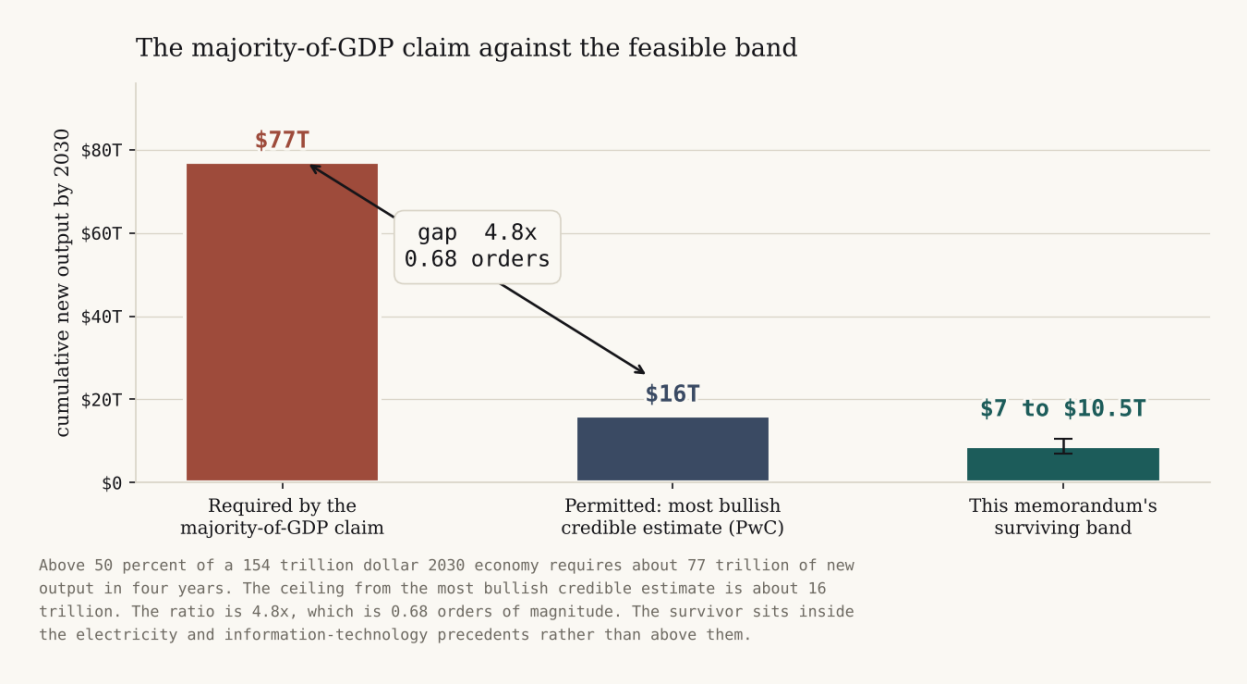


Figure 14: The majority-of-GDP claim against the feasible band. It requires $77T of new output in four years against a $16T ceiling from the most bullish credible estimate in the field. The survivor sits inside the electricity and information-technology precedents rather than above them.

creates little value from a GDP verdict of this size has misread it: the verdict is about capture and measurement, not about welfare.

Scenario. By 2027 realized contribution stays near ∼0.1 to 0.2 percentage points [38]. By 2029 diffusion lifts it toward ∼1pp. **Probability (rubric), survivor:** (iii) ∼0.53 at $\alpha \approx 1.4$, provisional, $c = 0.45$. **Computed (engine), tails:** (i) ∼0.19 [computed, p10-p90 7.1 to 31.6]; ∼0.12 lower-realization; ∼0.07 transient-negative; (ii) not stated, and would be forced by subtraction. The two lines are on different scales, as Section 1.3 requires, and are not summed.

How the survivor probability is derived, since three different values were in circulation. Version 3.0 carried 0.62 in the headline results, 0.53 in this section and 0.55 in Figure 19. Under the rubric, $P_{\text{stated}} = c \cdot P_{\text{engine}} + (1 - c)/n$ with $n = 5$ candidates. Q1 carries two rubric defects: dependence on a parameter nobody has measured whose observed range reverses the verdict, deducting 0.40, and a load-bearing cell that is not computable and is bounded on a base rate, deducting 0.15. From a base of 1.00 that gives $c = 0.45$ exactly, which is the coefficient this section has always stated. With $c = 0.45$ and $n = 5$ the floor term $(1 - c)/n$ is 0.11, so a stated probability of 0.62 would require $P_{\text{engine}} = 1.13$, which is not a probability. **0.62 is arithmetically impossible under this document's own rubric and is withdrawn.** 0.53 implies $P_{\text{engine}} = 0.93$ (EST), which is a valid probability, and 0.55 implies 0.98, which is valid but is not the value this section derives. **The rubric alone settles it.** Version 3.1 also cited the candidate distribution summing to 1.00 as corroboration. That argument is withdrawn and does not appear here: candidate (ii) is obtained by subtraction, so the distribution sums to 1.00 by construction and carries no information about the survivor. Section 1 is also explicit that rubric-scale survivor probabilities and engine-scale tail probabilities are separate scales that are not summed, which the corroborating argument quietly violated. **0.53 is used everywhere in this version**, including Figure 19 and the headline results.

Under $\alpha = 1.0$ the survivor is repriced below 0.30 and the verdict is not resolved. **The engine declines to reprice the majority-of-GDP candidate to zero where a purely qualitative reading would**, which is the discipline it exists to impose.

Bottom line. Material but not civilizational: ∼1.0 to 1.5 percentage points of annual GDP growth, ∼$7 to $10.5T as a 2030 annual contribution. The majority-of-GDP claim is repriced to near zero at Tier 5 and Tier 3 by 4.8x against the stated ceiling.

4.2 Q2. Will AI raise productivity as much as electricity or the internet?

Candidates. (i) an instant surge; (ii) no measurable gain; (iii) electricity-scale magnitude on an electricity-length timeline; (iv) internet-scale and faster; (v, taboo) a net-negative drag during reorganization.

Table 19: Q2 test matrix: candidates against the constraint set. Will AI raise productivity as much as electricity or the internet?

Candidate	Binding constraint	Tier	Verdict
(i) instant surge	Absorption: realized gain ∼0.1-0.2pp against a ∼1.5pp peak precedent	1	Repriced to ∼0
(ii) no gain	Data: delivered gains exist where measurable	2	Repriced to ∼0
(iii) electricity-scale, electricity-slow	none	–	Survives
(iv) internet-speed	Absorption and Matter: physical and organizational deployment gates it	1, 3	Repriced to ∼0
(v, taboo) transition drag	Absorption: live transient	1	Live tail

Computed. (i) requires realized gains now. Requires ∼1.5pp; permits ∼0.1 to 0.2pp. Gap ∼10x. (iv) requires software-style distribution, but the binding constraint is deployment, and the precedent is a two-to-four-decade adoption lag: electric motors [62] were under 5% of factory mechanical drive in 1900 with gains arriving in the 1920s.

The strongest contrary datum, stated. US labor productivity growth reached 2.7 percent [45, 54] in 2025, nearly double the 1.4 percent average of the preceding decade, which Brynjolfsson [52] (2026) reads as the early phase of a J-curve (AI Index 2026). If that is AI-attributable and sustained, this verdict is wrong. Three considerations bear on it and none is decisive: aggregate productivity growth is not AI-attributable by default, and the AI Index itself calls the macro evidence earlier and less conclusive than the micro; the J-curve reading is this memorandum's own thesis, since its early phase is absorption cost preceding measured gain; and one year against a decade average is within the historical variance of the series. **Falsification threshold: US productivity growth holding above 2.5 percent for three consecutive years with a credible AI attribution falsifies this verdict**, and the memorandum says so in advance. Tracked as S18.

Isolate. (iii) survives. **Probability (rubric), survivor:** (iii) ∼0.67, $c = 0.75$. **Computed (engine), eliminated candidate:** (i) requires 69.9 years of Absorption expansion against 4.4 available, and is repriced below 0.05% [p10-p90 both $< 0.05\%$]. The two are on different scales and are not summed. It is written this way rather than as 0.0 because only Tier 0 bars, and Absorption is Tier 1; the value is small, not zero. The 69.9 years is the engine's run over the full rate band; the midpoint shorthand $\ln 10 / \ln 1.035$ gives 66.9, and the two differ because the engine integrates over the band rather than evaluating at its center. This verdict is α-exposed on its magnitude claim and inherits the numerator caveat of Section 2.6.

Scenario. By 2027 realized gains reach approximately 0.2 to 0.4 percentage points. By 2029 approximately 0.5 to 1. By 2030 approaching 1 to 1.5 if adoption and delivered power both progress. **One apparent contradiction, resolved.** Candidate (i) is eliminated for requiring ∼1.5pp *now*, and this scenario reaches 1.0 to 1.5pp by 2030. Those are consistent: the candidate is an instant surge, and the scenario is a four-year ramp. The elimination is of the timing, not of the magnitude, which is the entire content of "electricity-scale in magnitude, electricity-slow in arrival."

Bottom line. Electricity-scale in magnitude, electricity-slow in arrival. This is the Absorption constraint in its purest form, and the ∼10x gap between realized and peak is the measurement of it.

4.3 Q3. Will AI create more value than it destroys, economy-wide?

Inherits R1 and localizes it to first-order accounting.

Computed. (i) requires negative first-order value. Bain [29] sizes the 2030 shortfall at ∼$800B per year, which means the industry does not yet cover its cost of capital, not that aggregate value is negative once diffused. Sign-level elimination. (ii) is forbidden by the ∼5% P&L-positive rate.

Isolate. (iii) survives. Probability: (iii) ∼0.73; ∼0.15 on social-negative; ∼0.07 on the cascade, revised with the corrected financing figure.

Scenario. By 2027 the industry remains cash-negative in aggregate. By 2029 leading operators approach self-funding. By 2030 net-positive value is realized, unevenly.

Table 20: Q3 test matrix: candidates against the constraint set. Will AI create more value than it destroys, economy-wide?

Candidate	Binding constraint	Tier	Verdict
(i) destroys more	Capital: industry-level savings already real	5	Repriced to $\sim$0
(ii) frictionless	Absorption: decade diffusion; Law: distributional friction	1, 4	Repriced to $\sim$0
(iii) net-positive, lagged	none	–	Survives
(iv) neutral reallocation	Data: additive categories exist	2	Repriced to $\sim$0
(v, taboo) cascade	Capital: live tail	5	Live tail

Bottom line. Net-positive, but the industry is not yet self-funding at scale. Front-loaded destruction, lagged creation, and a financing-cascade tail that is real but smaller than a debt-requirement reading implies.

4.4 Q4. How much new wealth will AI generate globally?

Table 21: Q4 test matrix: candidates against the constraint set. How much new wealth will AI generate globally?

Candidate	Binding constraint	Tier	Verdict
(i) evenly-shared abundance	Capital: returns concentrate; Law: no redistribution in place	5, 4	Repriced to $\sim$0
(ii) none	Capital: value is measurable	5	Repriced to $\sim$0
(iii) trillions, concentrated	none	–	Survives
(iv, taboo) negative	Capital: transient in a correction	5	Live tail

Computed. (i) requires a distribution mechanism. Returns concentrate to capital, skill, and geography, and the complements are not cheap where intelligence is.

Isolate. (iii) survives. Carry the full range rather than a point: Acemoglu's 0.66 percent of TFP over ten years, roughly 1 percent of GDP on his own adjustment, against McKinsey's $4.4T per year, about 3.5 percent of a 2026 world economy. The spread is roughly 3.5-fold on a common basis, and the TFP-to-GDP conversion is itself contested, which is why the range is carried rather than a point. Probability: $\sim$0.80. This verdict is α-exposed and inherits the numerator caveat of Section 2.6.

Scenario. By 2027 gains concentrate among compute and verified-data owners. By 2029 the base widens modestly. By 2030 the distribution remains skewed absent redistribution.

Bottom line. Trillions, but concentrated, and accruing to the complements rather than to intelligence itself.

4.5 Q5. Which industries see the largest productivity gains?

Table 22: Q5 test matrix: candidates against the constraint set. Which industries see the largest productivity gains?

Candidate	Binding constraint	Tier	Verdict
(i) uniform	Data: gains track measurability	2	Repriced to $\sim$0
(ii) none	Data: gains visible where verification is cheap	2	Repriced to $\sim$0
(iii) data-dense first	none	–	Survives
(iv) physical first	Matter: embodied AI immature; hardware lead times	3	Repriced to $\sim$0
(v, taboo) too narrow	Absorption: live if adoption stalls	1	Live tail

Isolate. (iii) survives. Software engineering, customer operations, back-office and finance, drug discovery, logistics, and marketing lead, wherever work is digital, measurable and cheaply verified. Probability: $\sim$0.80.

Scenario. By 2027 software engineering and finance operations show the clearest gains. By 2029 discovery and logistics follow. By 2030 physical sectors begin to participate as embodied systems mature.

Bottom line. Digital, measurable, verification-cheap sectors first. Gains are sectorally jagged, not uniform.

Table 23: Q6 test matrix: candidates against the constraint set. Which industries become structurally obsolete?

Candidate	Binding constraint	Tier	Verdict
(i) entire industries vanish	Law: licensure, installed base; Absorption: switching costs	4, 1	Repriced to $\sim$0
(ii) none change	empirical: exposed roles already compressing	state	Repriced to $\sim$0
(iii) middle layers die	none	–	Survives
(iv, taboo) a sub-industry vanishes	Absorption: live in one or two niches	1	Live tail

4.6 Q6. Which industries become structurally obsolete?

Isolate. (iii) survives. What dies is the routine middle: commodity business-process outsourcing, generic content production, tier-1 support, basic data processing, some SaaS middleware. Probability: (iii) $\sim$0.80; (iv) $\sim$0.15.

Scenario. By 2027 routine clerical roles compress fastest. By 2029 mid-tier analytical tasks follow. By 2030 the middle is thinner but no sector has vanished.

Bottom line. No whole industry vanishes this horizon. Routine middle layers compress, and the WEF [47] steepest-decline roles are the map.

4.7 Q7. Will AI increase or decrease economic inequality?

Table 24: Q7 test matrix: candidates against the constraint set. Will AI increase or decrease economic inequality?

Candidate	Binding constraint	Tier	Verdict
(i) auto-democratizing	Capital: complements are not cheap; Law: no redistribution	5, 4	Repriced to $\sim$0
(ii) neutral	Capital: returns concentrate	5	Repriced to $\sim$0
(iii) rises absent redistribution	none	–	Survives
(iv, taboo) backlash reversal	Law and Legitimacy: live political tail	4	Live tail

Isolate. (iii) survives with (iv) as a political tail linked to R1(v). Probability: (iii) $\sim$0.75; (iv) $\sim$0.15.

Scenario. By 2027 wage polarization sharpens in exposed occupations. By 2029 the capital-labor split widens absent policy. By 2030 the distributional question is the operative political variable.

Bottom line. Inequality increases through the transition absent deliberate redistribution. Cheap intelligence does not democratize the scarce complements.

4.8 Q8. Will AI create more jobs than it eliminates?

Table 25: Q8 test matrix: candidates against the constraint set. Will AI create more jobs than it eliminates?

Candidate	Binding constraint	Tier	Verdict
(i) mass unemployment	Capital: $\sim$23% of vision-task wages [53] cost-effective now; Absorption	5, 1	Repriced to $\sim$0
(ii) no disruption	empirical: 22% structural churn [47]	state	Repriced to $\sim$0
(iii) flat-to-positive, churn [47]	none	–	Survives
(iv, taboo) net-negative	state: live if autonomy inflects	–	Live tail

Computed. (i) requires displacement to swamp creation. MIT finds only $\sim$23% of vision-task wages [53] cost-effective to automate today; the WEF projects +78M net [47] to 2030 against 170M created and 92M destroyed. **Note the attribution: the $\sim$23% figure is a cost line, which is Capital at Tier 5, not adoption speed. It is easily misfiled.**

Isolate. (iii) survives. Probability: (iii) $\sim$0.70; (iv) $\sim$0.15. The engine runs on this verdict and moves it by 0.01, so it is classified Decorative under Section 1.3 and the stated probability is substantially a prior.

Scenario. By 2027 clerical displacement runs ahead of new-role creation. By 2029 reskilling determines the net. By 2030 the count is net-positive but the churn is large.

Bottom line. Roughly flat-to-positive by count with 22% churn. The pain is in the transition, and the destroyed jobs are not the created ones.

4.9 Q9. How long will the productivity transition take?

Table 26: Q9 test matrix: candidates against the constraint set. How long will the productivity transition take?

Candidate	Binding constraint	Tier	Verdict
(i) two to three years	Matter 3d: interconnection median above 5 years; Absorption: 95% of pilots no P&L	3, 1	Repriced to $\sim$0
(ii) never	Capital: competitive returns force adoption	5	Repriced to $\sim$0
(iii) decade-plus	none	–	Survives
(iv, taboo) faster	state: live if reliability inflects	–	Live tail

Computed. (i) requires diffusion in approximately 2 years. A figure of four to eight years for power queues circulates widely and is not one LBNL [14] publishes. **The correct datum is LBNL [14]'s median from interconnection request to commercial operation [14], which exceeded 5 years for projects built in 2025**, up from over 4 years for 2018 to 2024 and under 2 years for 2000 to 2007. Only 13% of capacity entering queues from 2000 to 2019 had reached commercial operation by end-2024. The Data leg is withdrawn; the elimination is stronger without it, because the interconnection figure has worsened while the data figure has softened.

Isolate. (iii) survives. Probability: (iii) $\sim$0.72; (iv) $\sim$0.18. **Computed: candidate (i) requires 46.5 years against 2.0 available, repriced to 0.2% [< 0.05 to 0.4].**

Scenario. By 2027 diffusion is early and uneven. By 2029 integration tooling matures. By 2030 the transition is well under way but incomplete.

Bottom line. A decade-plus, clocked by deployment friction at Tier 1 and grid interconnection at Tier 3, not by model capability. Only an agent-reliability inflection resets it.

4.10 Q10. Sustained growth, or merely shifted profits among firms?

Table 27: Q10 test matrix: candidates against the constraint set. Sustained growth, or merely shifted profits among firms?

Candidate	Binding constraint	Tier	Verdict
(i) regime-change	Capital and Matter: needs output no constraint permits; Absorption	5, 3, 1	Repriced to $\sim$0
(ii) pure reallocation	Data: additive categories exist	2	Repriced to $\sim$0
(iii) reallocation then growth	none	–	Survives
(iv, taboo) investment incineration	Capital: live tail	5	Live tail

Computed. (iv) is not eliminated. High-fixed, low-marginal-cost GPU infrastructure sees prices competed toward marginal cost absent pricing power, and semiconductors depreciate on a Moore's-Law curve unlike railroads. Bain's $\sim$\$800B shortfall [29] is the same signal.

Isolate. (iii) survives with (iv) as a live tail. Probability: (iii) $\sim$0.65; (iv) $\sim$0.25. This verdict is α-exposed and inherits the numerator caveat of Section 2.6.

Scenario. By 2027 the effect is reallocation and margin. By 2029 additive categories begin to register as growth. By 2030 growth is real and inside the historical precedents.

Bottom line. Reallocation and margin first, real growth later, with a live investment-incineration tail.

Table 28: Q11 test matrix: candidates against the constraint set. Who controls compute?

Candidate	Binding constraint	Tier	Verdict
(i) monopoly	Matter: no actor owns all layers; Law: antitrust and export controls	3, 4	Repriced to ∼0
(ii) decentralization	Matter 3c/3d: gigawatt delivered power; Capital	3, 5	Repriced to ∼0
(iii) layered oligopoly plus walled China	none	–	Survives
(iv) state control	Capital: the state is not funding the buildout	5	Repriced to ∼0
(v) China parity near-term	Matter 3b: ∼5 to 17x lead, HBM-gated; Talent	3, 1	Repriced to ∼0
(vi, taboo) ASIC revolt	Matter: live erosion tail	3	Live tail

5 Compute and Its Financing

5.1 Q11. Who controls compute?

Computed. (vi) is not eliminated. Hyperscaler custom silicon is scaling and the incumbent's accelerator share is projected to ease from ∼85 to 90% toward ∼75% in 2026.

Isolate. (iii) survives. Probability: (iii) ∼0.75; (vi) ∼0.15.

Scenario. By 2027 the US lead widens even as custom silicon trims the accelerator share toward approximately 75 percent. By 2029 power, not chips, is the binding control point. By 2030 control still sits with the layered oligopoly.

Bottom line. A layered, US-led oligopoly with a live custom-silicon erosion tail. Power and grid equipment, at Tier 3, are the 2027-and-beyond chokepoint.

5.2 Q12. How much compute will AI actually need?

Table 29: Q12 test matrix: candidates against the constraint set. How much compute will AI actually need?

Candidate	Binding constraint	Tier	Verdict
(i) unlimited	Matter 3c/3d: transformer and interconnection lead times; Capital	3, 5	Repriced to ∼0
(ii) plateau soon	empirical: test-time compute is exploding demand	state	Repriced to ∼0
(iii) bounded growth	none	–	Survives
(iv, taboo) demand collapse	Capital: Jevons [69], cheaper compute raises use	5	Live tail, weak

Computed. (i) requires power, capital and fabrication without bound. Bain projects ∼200 GW [29] of incremental compute by 2030. **The specific mechanism is now dated and precise: US large-power transformer lead times average 128 weeks [19] and reach 3 to 5 years for the largest units; medium-voltage switchgear [19] is effectively sold out through 2028; GE Vernova [20]'s gas-turbine backlog reached 100 GW in Q1 2026 with roughly 10 GW of slots remaining across 2029 and 2030 combined.** Of roughly 12 to 16 GW of announced US 2026 data-center capacity, approximately 5 GW is under construction. At $r_{3c} = 0.10$ to 0.20, closing a 2.8x gap requires $T_{\text{req}} \approx 7.4$ years at the midpoint, and 5.6 to 10.8 across the band. **Version 3.0 published 5.2 years, which is shorter than the 5.6 years the fastest rate in its own stated range would need, and is withdrawn.** Against the 1.4 years available to 2026 delivery, the elimination holds at every rate in the band.

Note the direction of travel at Tier 3a. Advanced packaging, the binding constraint of 2024, is loosening: CoWoS [23] capacity is moving from 13,000 wafers per month at end-2023 toward a 120,000 to 130,000 target at end-2026, and the supply-demand gap is reported narrowing from approximately 20% toward approximately 10%. **The binding constraint has migrated from packaging to grid equipment, and the sub-constraint expansion rates in Table 4 predict exactly that migration without additional assumptions.** The ratio is 6.0 at the low ends of the two bands and 5.5 at the high ends, so five to six fold; Version 3.0 called it an order of magnitude; **the measured five to six fold is the number the migration argument actually rests on**.

Isolate. (iii) survives. **Probability (rubric), survivor:** (iii) ∼0.79, $c = 0.85$. **Computed (engine), eliminated candidate:** (i) requires 7.3 years against 4.4 available, repriced to 10.5% [1.0 to 24.5], which the arithmetic states and the prose now matches. Note that this is a different computation from the 7.4 above: that one is the 2026 delivery gap

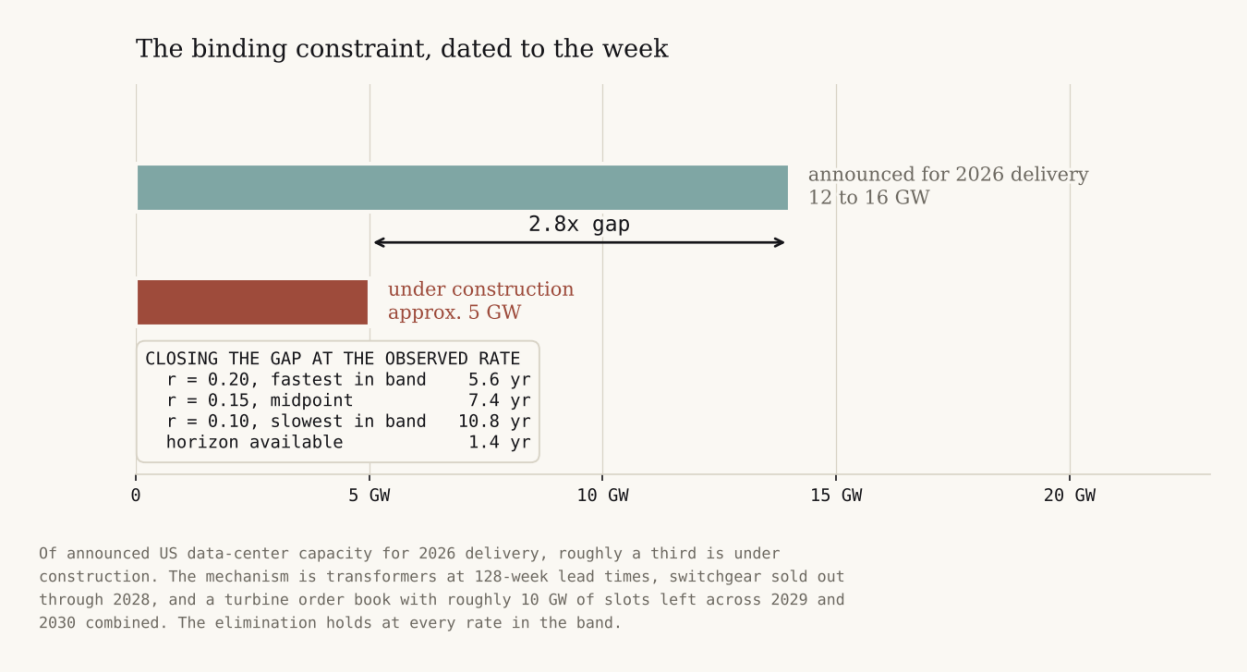


Figure 15: The binding constraint, dated to the week. Of announced US data-center capacity for 2026 delivery, roughly a third is under construction. Closing that gap requires 7.4 years of expansion at the midpoint observed rate against 1.4 available, and 5.6 even at the fastest rate in the band. The mechanism is transformers at 128-week lead times, switchgear [19] sold out through 2028, and a turbine order book with roughly 10 GW of slots left across 2029 and 2030 combined.

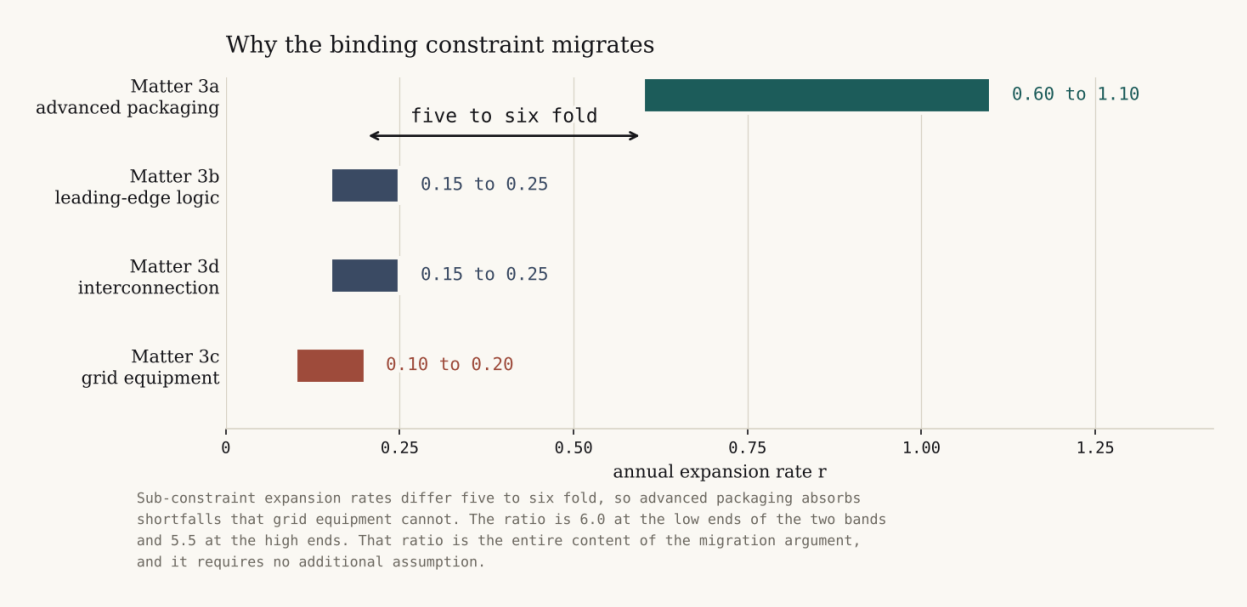


Figure 16: Why the binding constraint migrates. Sub-constraint expansion rates differ five to six fold, 6.0 at the low ends of the two bands and 5.5 at the high, so advanced packaging absorbs shortfalls that grid equipment cannot. The ratio between those two rates is the entire content of the migration argument, and it requires no additional assumption.

against a 1.4-year horizon, this one is unbounded demand against the 2030 horizon. The 10.5% is raw engine output and is not rubric-adjusted; under the rubric with $c = 0.85$ and $n = 4$ it would state as 0.127. Section 1 sets out why the two scales are reported separately.

Scenario. By 2027 delivered power, not chip supply, sets the pace. By 2029 inference dominates the compute bill. By 2030 growth continues but decelerates against the grid.

Bottom line. Large but constraint-bounded. The ceiling is grid equipment and interconnection at Tier 3, not appetite, not capital, and not physics.

5.3 Q13. Who finances the next trillion?

Table 30: Q13 test matrix: candidates against the constraint set. Who finances the next trillion?

Candidate	Binding constraint	Tier	Verdict
(i) cash flows alone	observation, not constraint: firms are in fact borrowing above $400B in 2026	state	Repriced to ∼0 by observation
(ii) it stalls	Capital: competition and demand force the spend	5	Repriced to ∼0
(iii) debt, private credit, SPVs	none	–	Survives
(iv, taboo) cascade	Capital: live tail	5	Live tail

A reclassification. Section 1.3 records that the engine returns 96.5% survival for this candidate, because a capex-to-free-cash-flow gap of 1.3 to 1.8x closes in 1.4 years at the observed capital expansion rate. The engine is right and the question was misfiled: this is a current-state question, not a capacity race. The candidate is eliminated by observation, since the firms are borrowing, not by any constraint, and the verdict is classified Empirical under Section 1.3. (i) requires internal funding to cover approximately $690B [28] to $725B of 2026 capex. **The widely cited figure of approximately $1.5T of external financing needed through 2028 misstates its source.** In Morgan Stanley's analysis the $1.5T is a funding gap closable by bonds, private credit, asset-backed structures or equity, and the corresponding corporate debt issuance estimate is approximately $200B through 2028. Actual issuance is large but smaller than the misquote implies: hyperscaler borrowing is expected above $400B in 2026 against approximately $165B in 2025, and global AI-related debt issuance is forecast near $570B for 2026, with approximately $236B [28] priced by May 31, four times the year-earlier pace.

Stress is idiosyncratic before it is systemic. One major issuer's five-year credit default swap widened from approximately 40 basis points at the start of 2025 to approximately 200 by March 2026, implying roughly a 16% default probability, while sector spreads remained near cycle lows. Bond order coverage for hyperscaler issues fell from approximately 5x in February 2026 to below 2x by July. Those two, together, are the earliest available systemic signal.

Isolate. (iii) survives. Probability: (iii) ∼0.78; (iv) ∼0.10, on the corrected financing figures.

Scenario. By 2027 private credit and bond issuance dominate the funding mix. By 2029 the revenue-versus-financing gap is the market's central question. By 2030 the buildout is financed but carries elevated refinancing risk.

Bottom line. Debt and private credit, not cash flow, but the gap is a funding gap and not a debt requirement. The buildout is leveraged and SPV-heavy, and the cascade tail is live but smaller than previously stated.

6 Data

6.1 Q14. Does AI run out of data?

The confidence on this verdict is bounded by the backtest of Section 15.

Candidates. (i) a hard wall stops scaling outright; (ii) synthetic and multimodal data solve it entirely; (iii) public text exhausts around 2028 while scaling continues on other axes; (iv, taboo) contamination degrades the stock faster than exhaustion approaches.

Computed. A reading that reprices (ii) to near zero, on the ground that synthetic data risks model collapse [6, 7] in open domains, is too strong, and Section 15 shows the framework getting the closely related question wrong at maximum confidence. The evidence as of mid-2026: Gerstgrasser [7] et al. established that accumulating synthetic data while retaining the original real data bounds test error and breaks collapse; the consensus position in the 2026

Table 31: Q14 test matrix: candidates against the constraint set. Does AI run out of data?

Candidate	Binding constraint	Tier	Verdict
(i) hard wall stops all	Data: ignores test-time compute and verifiable synthetic	2	Repriced to ∼0
(ii) fully solved	Data: model collapse [6, 7] in unverifiable domains	2	Live tail, upgraded
(iii) exhausts ∼2028, scaling shifts	none	–	Survives
(iv, taboo) contamination	Data: live, strengthens the wall	2	Live tail

International AI Safety Report [8] is that synthetic data works where answers are machine-verifiable and risks collapse where verification is hard [8, 7]. Epoch's own estimate of total available training data across modalities by 2030, after quality correction, spans 400 trillion to 20 quadrillion tokens. **The width of that interval is the finding.** With r_{data} revised from 0.02 to 0.05 up to 0.30 to 0.60, candidate (ii) is no longer eliminated; it is a live tail at approximately 0.28, the value carried in the distribution below. Version 3.0 stated 0.20 here and 0.28 there; 0.28 is the one consistent with the stated distribution and is used throughout.

The ">74% of new web pages AI-generated" figure supporting candidate (iv) could not be traced to a stable primary in the current research pass and is carried here with that caveat, not as an anchor (EST). The nearest traceable source is a commercial vendor's detector study of approximately 900,000 newly created English pages in April 2025, released alongside the detector itself, reporting 74.2 percent containing some AI-generated content with only 2.5 percent assessed as pure AI. Lower independent figures exist for adjacent constructs. The figure is retained with its label and is not load-bearing on the verdict.

Isolate. (iii) survives, with (ii) and (iv) as live tails pulling in opposite directions. The pre-training lever is closing, and the scarce complement migrates from raw data to verification apparatus. **Probability (rubric), survivor:** (iii) ∼0.41 at $c = 0.60$. **Computed (engine), tails:** (ii) ∼0.28; (iv) ∼0.21. These do not sum to 1.00 and are not meant to: the survivor is rubric-adjusted and the tails are engine posteriors, which Section 1 sets out as separate scales. **Lowered twice, and the second correction overrules the first.** An earlier revision moved this from 0.60 to 0.52 on the ground that the engine returned 0.52. That reasoning was circular: the engine's output was itself a consequence of slack ratios chosen for this verdict, so the number was being matched to a construction rather than to evidence. Under the rubric of Section 1 the verdict carries two independent defects, a binding constraint on the tier holding this framework's recorded backtest miss and a primary source that has already revised its own estimate once, giving $c = 0.60$ and a stated probability of 0.41. The full history is therefore 0.60 to 0.52 to 0.41. **This is the least confident verdict in the memorandum, and the confidence is low because the backtest earned it that way** on the one tier where this method has a measured failure. One precision note: the two defects are correlated, since both concern the reliability of the same estimate, so the deduction may be double-counting. Under-confidence is the right direction of error on a tier with a measured miss. The engine runs on this verdict and moves it by 0.07, so it is classified Decorative under Section 1.3.

Scenario. By 2027 frontier datasets approach the effective public-text ceiling and verified synthetic share rises. By 2028 the pre-training wall binds on the median for text alone. By 2029 multimodal and verified synthetic sources are the marginal inputs and verification apparatus is the priced scarcity.

Bottom line. Public text exhausts around 2028 on the median and scaling migrates to verified synthetic, multimodal, and private sources. The confidence on this verdict is materially lower than the surface evidence suggests, because the framework has been measured getting an adjacent question wrong, and Tier 2 is the tier where substitution defeats stock-based reasoning.

7 Models and Scale

7.1 Q15. Do models commoditize?

Computed. Inference cost for a fixed capability level falls roughly ten times per year, from approximately $30 to under $0.50 [31] per million tokens for GPT-4-class output on the a16z series, against Table 6's $37.50 to $0.07–0.18 on the Epoch series, the two differing because they define the capability level on different benchmarks; Stanford HAI [54] measures a 99.7 percent fall in cost per token between November 2022 and December 2024, a 333-fold fall over twenty-five months, or approximately 16.3 times per year. **The frontier price, the best model at launch, falls more slowly, at 1.34 to 2.25 times per year depending on whether reasoning models are allowed to set the frontier.** Section 1.4 sets out why the two must not be interchanged. Open-weight and low-cost entrants, one priced approximately 90% under incumbents [54], have compressed the gap to months.

Table 32: Q15 test matrix: candidates against the constraint set. Do models commoditize?

Candidate	Binding constraint	Tier	Verdict
(i) one model wins	empirical: several labs at the frontier, no durable lead	state	Repriced to ∼0
(ii) zero margin everywhere	Capital: the frontier still commands a premium	5	Repriced to ∼0
(iii) floor commoditizes, frontier premium	none	–	Survives
(iv, taboo) open overtakes	live tail	–	Live tail

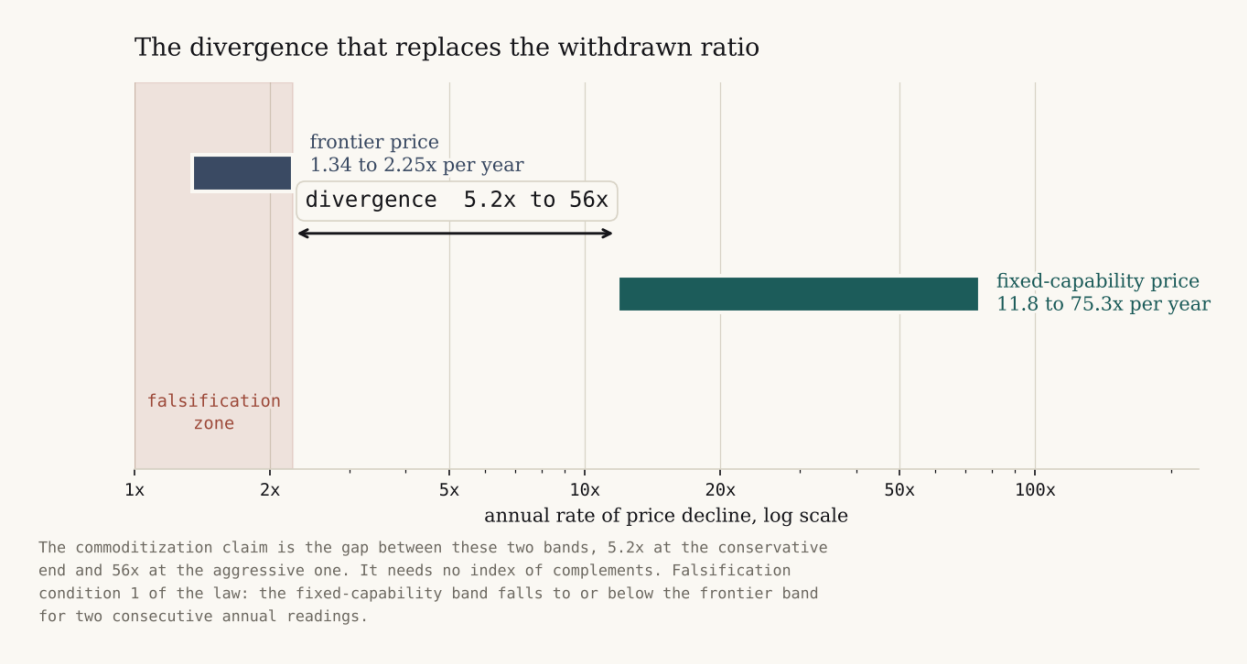


Figure 17: The divergence that replaces the withdrawn ratio. The two measured bands, frontier at 1.34 to 2.25x per year and fixed capability at 11.8 to 75.3x, and the gap between them: 5.2x taking the conservative end of each and 56x taking the aggressive end. That gap is the commoditization claim stated as one quantity, and unlike the withdrawn ratio it requires no index of complements. The shaded region is falsification condition 1 of the law: the fixed-capability band falling to or below the frontier band for two consecutive annual readings.

Isolate. (iii) survives. The model becomes a commodity and the moat is everything around it. This is the commoditization mechanism stated in one market: the floor falls at 11.8 to 75.3x per year while the frontier falls at 1.34 to 2.25x. Probability: (iii) ∼0.75; (iv) ∼0.15.

Scenario. By 2027 last year's frontier is this year's commodity floor. By 2029 open-weight [54] parity at the floor is routine. By 2030 differentiation lives in data, workflow and distribution, not the model.

Bottom line. The capability floor commoditizes fast while the frontier holds a brief, eroding premium, so the divergence between the two prices widens: 5.2 to 56 times, and it is the migration measured in one market.

7.2 Q16. Does scale still work?

Computed. (ii) previously rested on three legs: the 2028 data wall, prohibitive power and capital, and diminishing gain per doubling. The data-wall leg is withdrawn per Section 15. The elimination holds on the remaining two: another hundredfold of pre-training compute requires fab and grid expansion that Tier 3 cannot deliver on the horizon.

Isolate. (iii) survives. Probability: (iii) ∼0.68; (iv) ∼0.17. Unchanged, because the surviving legs were always sufficient.

Table 33: Q16 test matrix: candidates against the constraint set. Does scale still work?

Candidate	Binding constraint	Tier	Verdict
(i) scale is dead	empirical: test-time scaling is delivering gains	state	Repriced to $\sim$0
(ii) as before	Matter 3b/3c: another 100x of fab and grid; Capital: cost per gain rising	3, 5	Repriced to $\sim$0
(iii) new axes sustain gains	none	–	Survives
(iv, taboo) architecture break	unforecastable	–	Live tail

Scenario. By 2027 test-time compute is the headline scaling axis. By 2028 pre-training saturates against fabrication and grid. By 2029 the question is how much thinking to buy at inference, not how large to pre-train.

Bottom line. Scale still works but the lever moved, and the reason it moved is fabrication and power rather than tokens. The question flips from how big to pre-train to how much thinking to buy at inference.

8 Labor and Value

8.1 Q17. How much knowledge work is actually automatable?

Table 34: Q17 test matrix: candidates against the constraint set. How much knowledge work is actually automatable?

Candidate	Binding constraint	Tier	Verdict
(i) nearly all	Capital: $\sim$23% of vision-task wages cost-effective today [53]; empirical: reliability caps multi-step	5, state	Repriced to $\sim$0
(ii) almost none	Capital: two-thirds of jobs are exposed [38]	5	Repriced to $\sim$0
(iii) task-level, cost-gated	none	–	Survives
(iv, taboo) larger core, reliability-capped	live as costs fall	–	Live tail

Computed, re-attributed. (i) requires broad cost-effective automation now. Only $\sim$23% of *vision-task* wages are cost-effective to automate today [53]. **The scope matters and Version 3.0 dropped it here**: reference [53] measures computer-vision tasks, not knowledge work, so it bounds this question by analogy rather than by measurement. No equivalent cost-effectiveness study for knowledge work exists, which is why this elimination is weak on two counts rather than one. **This figure is easily misfiled under adoption speed. It is a cost line, which is Capital at Tier 5, and the distinction matters because it means this elimination is weak: Tier 5 relaxes in weeks.** As fixed-capability inference cost falls [31] approximately tenfold per year, which is the a16z series rather than the 11.8 to 75.3x band of Table 6, the cost-effective share rises, and the realized-automation frontier is a moving cost line rather than a fixed capability line. Talent enters here as a co-constraint on the deployment side: the skilled workforce to integrate these systems is Tier 1 and does not track the cost curve.

Isolate. (iii) survives. Probability: $\sim$0.75.

Scenario. By 2027 the cost-effective share rises above the current approximately 23 percent of vision-task wages [53] as inference cheapens. By 2029 reliability, not cost, is the binding limit. By 2030 task-level automation is widespread but job-level replacement is not.

Bottom line. Task-level, not job-level. Exposure is broad but cost-effective automation is partial and rising, and reliability, not capability, caps what is realized.

8.2 Q18. What gains value when intelligence is cheap?

Isolate. (iii) and (iv) survive together, and this question is the thesis in its purest form. Value migrates to what stays scarce when tokens are free: energy, verified data, fabrication capacity, trust, distribution, and human accountability for consequential decisions. Probability: $\sim$0.80.

Scenario. By 2027 verified proprietary data and delivered power are the priced scarcities. By 2029 distribution and trust are the durable moats. By 2030 accountable human judgment is a premium complement, not a commodity.

Table 35: Q18 test matrix: candidates against the constraint set. What gains value when intelligence is cheap?

Candidate	Binding constraint	Tier	Verdict
(i) nothing	Data: additive categories and complements exist	2	Repriced to $\sim$0
(ii) intelligence stays scarce	Capital: fixed-capability inference 10 to 75x/yr cheaper commoditizes it	5	Repriced to $\sim$0
(iii) the complements gain	none	–	Survives
(iv, taboo) human judgment gains	none	–	Survives

Bottom line. Value migrates to the complements intelligence cannot make abundant. This is the migration in its purest form: the layer beneath is where scarcity lands.

9 Agents: Reliability and Liability

A methodological note governs this section. Q19 and Q20 are not constraint eliminations. They are empirical arguments from measured reliability, and both are classified Empirical under Section 1.3. Under the stratified logic of Section 1.1, they operate at the abductive layer, not the deductive or inductive layers, and this memorandum labels them as such, where the field routinely dresses observation as constraint. The verdicts are strong; their strength comes from measurement, not from a wall.

9.1 Q19. Can we trust agents to act reliably?

Table 36: Q19 test matrix: candidates against measured reliability. Can we trust agents to act reliably?

Candidate	Basis	Verdict
(i) already reliable	empirical: $\sim$56% production single-task success [13]; $\sim$88% of pilots never reach production [48]	Repriced to $\sim$0
(ii) never	empirical: the 50% task horizon keeps lengthening	Repriced to $\sim$0
(iii) reliable in bounded, verifiable domains	none	Survives
(iv, taboo) sub-threshold plateau	none	Live and strengthening

Measured, and the verdict strengthens. The figures in general circulation, a 50 percent horizon of approximately 2h17m with the 80 percent horizon four to eight times shorter, are both superseded, and the current measurements strengthen the verdict. Claude Opus 4.5 [9] measured a 50 percent horizon of approximately 4h49m in December 2025 with a 95 percent interval from 1h49m to 20h25m, and later models were reported considerably higher. **Its 80 percent horizon was approximately 27 minutes, and the 80 percent series had been approximately flat at 27 to 32 minutes [9] across the preceding frontier releases.** The ratio between the two horizons has therefore widened from approximately 5x toward 10x to 25x.

The mechanism is the finding. Capability is arriving as intercept rather than slope: models succeed more often on long tasks without succeeding more reliably on them. Trust scales with verifiability, not with raw capability, and the widening ratio is that statement measured on a third-party instrument.

Isolate. (iii) survives, with (iv) upgraded from tail to near-term texture. Probability: (iii) $\sim$0.70; (iv) $\sim$0.25.

Scenario. By 2027 bounded-task reliability is production-grade in narrow domains while the 80 percent horizon stays in hours. By 2029 supervised autonomy widens if and only if that horizon lengthens. By 2030 open long-horizon autonomy remains gated.

Bottom line. Reliable only within bounded, verifiable, short-horizon [12, 13] tasks. The 80 percent horizon, the one that matters for delegation, has been flat at roughly half an hour across two release cycles while the headline number rose several-fold.

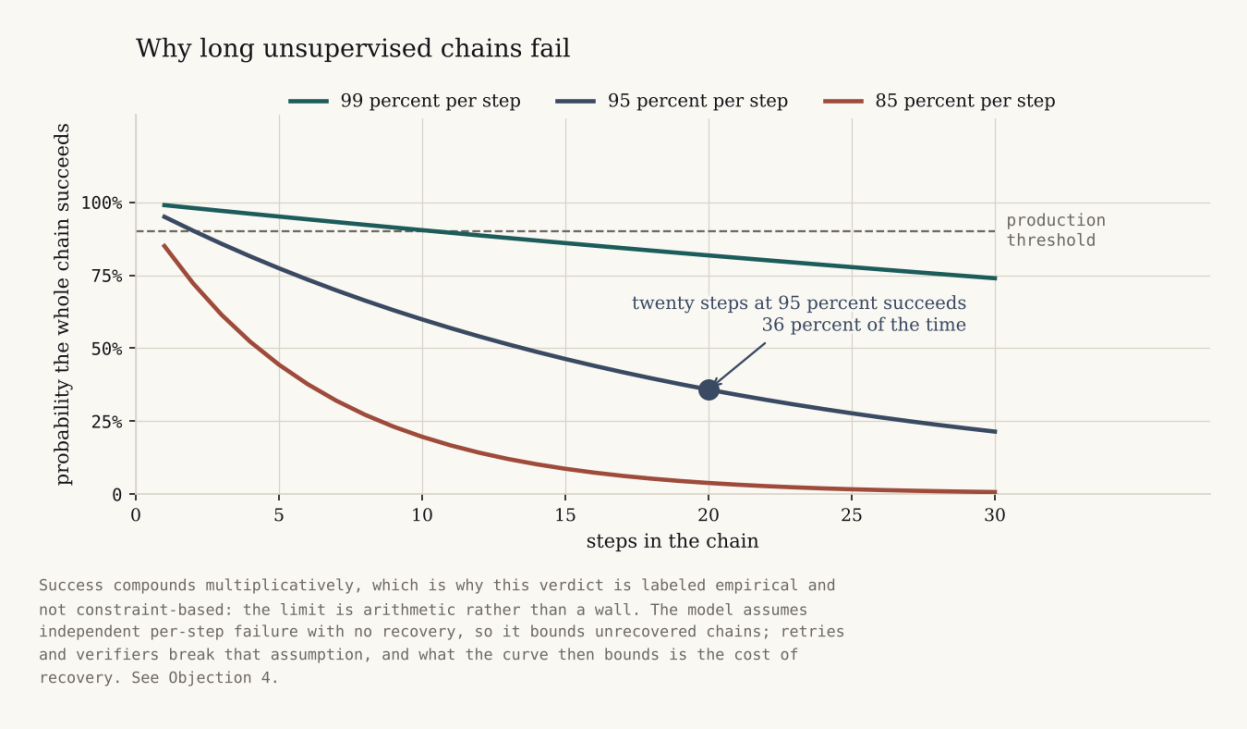


Figure 18: Why long unsupervised chains fail. Success compounds multiplicatively, so at 95 percent per step a twenty-step workflow succeeds about 36 percent of the time. The limit is arithmetic rather than sentiment, which is why this verdict is labeled empirical and not constraint-based.

Table 37: Q20 test matrix: candidates against measured reliability. Can agents run whole workflows unsupervised?

Candidate	Basis	Verdict
(i) full autonomy soon	arithmetic: $0.95^{20} \approx 36\%$; production single-task $\sim$56%	Repriced to $\sim$0
(ii) never	empirical: short verifiable chains already work	Repriced to $\sim$0
(iii) supervised at handoff boundaries	none	Survives
(iv, taboo) permanent compounding wall	none	Live tail

9.2 Q20. Can agents run whole workflows unsupervised?

Measured. Success compounds multiplicatively. At 95% per step a twenty-step workflow succeeds about 36% of the time; $\sim$99%+ per step is required for long chains, and production single-task success [13] is $\sim$56%. Compounding is arithmetic, not sentiment.

Isolate. (iii) survives. Probability: (iii) $\sim$0.72; (iv) $\sim$0.18.

Scenario. By 2027 short-chain autonomy is common and long-chain autonomy is not. By 2029 per-step reliability improvements extend the economic chain length. By 2030 whole-workflow autonomy remains bounded to verifiable domains.

Bottom line. Agents run unsupervised only in short, verifiable chains. The binding limit is compounding reliability, and it is arithmetic rather than a constraint in this framework's sense.

9.3 Q21. Who is liable for an agent's decisions?

Computed. Agency law makes a principal liable for an agent's authorized acts and for negligent selection or supervision. The Product Liability Directive 2024/2853, treating software as a product under strict liability, has a transposition deadline of December 9, 2026 that the Digital Omnibus [58] did not move. The Omnibus deferred the AI Act's high-risk

Table 38: Q21 test matrix: candidates against the constraint set. Who is liable for an agent's decisions?

Candidate	Binding constraint	Tier	Verdict
(i) no one	Law: agency law and product liability assign it to humans	4	Repriced to ∼0
(ii) developer primarily	Law: developer liable for a defect, not for use	4	Repriced to ∼0
(iii) the deploying principal	none	–	Survives
(iv, taboo) new developer regime	Law: AI Liability Directive withdrawn October 2025	4	Live tail

obligations to December 2, 2027 and August 2, 2028, so the operative near-term instrument for AI harm is product liability and existing sectoral law, not the AI Act.

Isolate. (iii) survives. Probability: (iii) ∼0.75; (iv) ∼0.15.

Scenario. By 2027 deployer liability is the working default and contracts allocate it explicitly. By 2029 strict product liability for defective AI is litigated under the transposed Directive. By 2030 a bespoke agent-liability regime is still emerging, not settled.

Bottom line. The deploying principal, by default. The residual responsibility gap for genuinely autonomous acts makes accountable human oversight a scarce, valuable complement: the same migration appearing in law.

10 Software

10.1 Q22. Does SaaS survive AI?

Table 39: Q22 test matrix: candidates against the constraint set. Does SaaS survive AI?

Candidate	Binding constraint	Tier	Verdict
(i) SaaS dies	Data: incumbents keep system-of-record [40] and data moat; Capital	2, 5	Repriced to ∼0
(ii) unchanged	Capital: seat contraction and multiple compression under way	5	Repriced to ∼0
(iii) survives, reprices	none	–	Survives
(iv, taboo) incumbents hollowed	live tail	–	Live tail

Isolate. (iii) survives. SaaS reprices from per-seat to consumption and outcome models, with margins bifurcating: AI-native software commands roughly two to four times the multiple of legacy peers. Probability: (iii) ∼0.70; (iv) ∼0.18.

Scenario. By 2027 outcome and consumption pricing is standard and seat counts soften. By 2029 the AI-native application layer is a durable tier, not a wrapper. By 2030 the system of record endures while the seat model is gone.

Bottom line. SaaS survives but reprices. Value migrates from seats to outcomes [40] and data.

11 The Physical World

11.1 Q23. When are humanoid robots economically viable?

Table 40: Q23 test matrix: candidates against the constraint set. When are humanoid robots economically viable?

Candidate	Binding constraint	Tier	Verdict
(i) now, mass	Matter: last-10% dexterity [54] gap; Absorption: pilot-to-scale gap	3, 1	Repriced to ∼0
(ii) never economic	Capital: payback under 18 months in two-shift roles	5	Repriced to ∼0
(iii) structured niches first	none	–	Survives
(iv, taboo) cost-led early inflection	Matter: live if dexterity [54] lags	3	Live tail

Computed. (i) requires production-grade dexterity and reliability at scale. Roughly 16k units shipped in 2025 against a mass-deployment requirement one to one and a half orders of magnitude higher. At the observed unit-cost and shipment expansion rate of 40 to 80 percent annually, $T_{\text{req}} = 6.3$ years against 4.4 available, repriced to 22.9 percent [p10 to p90: 7.0 to 40.9]. **The computed elimination is weaker than the qualitative reading suggests, and the survivor probability below follows the arithmetic rather than the prose.** (ii) is falsified by payback: at a ∼$50k build the payback against a two-shift loaded wage is under eighteen months, and entry units are already near $16k.

Isolate. (iii) survives. Payback runs twelve to eighteen months at ∼$50k and under six at the ∼$20-30k target, with roughly 16k units shipped in 2025. **Probability (rubric), survivor:** (iii) ∼0.67 at $c = 0.90$. **Computed (engine), eliminated candidate and tail:** (i) ∼0.23 [computed 22.9%, p10-p90 7.0 to 40.9]; (iv) ∼0.13, placed on the engine scale under the rule stated at Q14 that tails are engine posteriors. These do not sum to 1.00, because the survivor is rubric-adjusted and the others are raw engine output; Section 1 sets out why the two scales are reported separately rather than reconciled by adjusting one to fit. Unit cost is falling fast enough that the engine will not reprice mass deployment to zero on this horizon.

Scenario. By 2027 structured-niche deployments scale from thousands to tens of thousands of units. By 2029 costs approach the target and payback shortens. By 2030 general-purpose dexterity is still the gating problem.

Bottom line. Viable first in structured, high-wage, multi-shift niches. The last-10% dexterity and reliability gap keeps general deployment out.

11.2 Q24. When do autonomous vehicles go mainstream?

Table 41: Q24 test matrix: candidates against the constraint set. When do autonomous vehicles go mainstream?

Candidate	Binding constraint	Tier	Verdict
(i) already mainstream	empirical: ∼11 metros [54], geofenced, under 1% of trips	state	Repriced to ∼0
(ii) never	empirical: millions of autonomous miles per week, safety at scale	state	Repriced to ∼0
(iii) geofenced, scaling city by city	none	–	Survives
(iv, taboo) vision-only generalizes	live wildcard	–	Live tail

Computed. (i) requires ubiquity. The leading service runs approximately 500k paid rides [54] per week across approximately 11 mapped metros, a fraction of a percent of trips, against a requirement of 15 to 40 times that footprint. At the observed metro-expansion rate of 30 to 60 percent annually, $T_{\text{req}} = 8.8$ years against 4.4 available, repriced to 6.5 percent [0.8 to 14.6]. (ii) is falsified by scale and safety: millions of autonomous miles per week at a safety record roughly an order of magnitude better than human.

Note the constraint attribution. The binding constraint on (iii)'s pace is Law and Legitimacy at Tier 4, permission to deploy per jurisdiction, combined with the mapping and validation work per market. Neither is a Matter constraint, and an attribution to Matter would be imprecise.

Isolate. (iii) survives. Probability: (iii) ∼0.72; (iv) ∼0.16. **Computed: candidate (i) requires 8.8 years against 4.4 available, repriced to 6.5% [0.8 to 14.6].**

Scenario. By 2027 robotaxi service is normal in tens of metros but still geofenced. By 2029 coverage broadens and cost per mile falls below rideshare. By 2030 mainstream-everywhere is approaching but unsolved for personal ownership in all conditions.

Bottom line. Robotaxis scale city by city, gated by per-jurisdiction permission and validation. Mainstream everywhere is a decade-plus away.

12 The Strategic Capstone

12.1 Q25. Does intelligence become an abundant utility?

Computed. (i) requires unbounded supply. Even as fixed-capability price per token falls approximately tenfold annually on the a16z series, against the 11.8 to 75.3x of Table 6, delivered intelligence is capped by fabrication and grid. Meeting the unbounded requirement demands a five to tenfold expansion of the physical layer; at the Tier 3 rates of 15 to 25 percent annually, $T_{\text{req}} = 11.0$ years against 4.4 available, repriced to 0.9 percent [< 0.05 to 2.5]. (ii) is falsified by the

Table 42: Q25 test matrix: candidates against the constraint set. Does intelligence become an abundant utility?

Candidate	Binding constraint	Tier	Verdict
(i) free and infinite	Matter 3b/3c/3d: fabrication and grid cap supply; Capital	3, 5	Repriced to $\sim$0
(ii) stays scarce	Capital: inference $\sim$10x/yr cheaper	5	Repriced to $\sim$0
(iii) tokens commoditize, physical layer scarce	none	–	Survives
(iv, taboo) the grid beneath is the real utility	none	–	Survives

annual collapse in fixed-capability price toward near-free tokens, tenfold on the a16z series and 11.8 to 75.3x on the Epoch series of Table 6.

Isolate. (iii) and (iv) survive together, and this is the capstone. Probability: $\sim$0.78. **Computed: candidate (i) requires 11.0 years against 4.4 available, repriced to 0.9% [< 0.05 to 2.5].** The engine's output moves this verdict by less than 0.12, so it is classified Decorative under Section 1.3: its slack ratios are sourced but the candidate set is lopsided enough that the arithmetic does not discriminate.

Scenario. By 2027 token pricing approaches utility economics for commodity capability. By 2029 delivered power and fabrication are the explicitly priced scarcities. By 2030 the frontier premium persists atop a near-free floor.

Bottom line. Abundance at the token layer makes the physical complements the binding scarcity, and the divergence between the two prices is at its widest here. The mechanism is Tier 3, and it is now dated to the week: transformers at 128 weeks [19], switchgear sold out through 2028, interconnection at a median above five years.

13 Interim Residual

Intelligence commoditizes, and scarcity and value migrate to its complements. The same migration appears in every one of the thirty questions. The law now carries two falsification conditions and the thesis is bounded to the horizon where physical constraints dominate.

Across the five roots, the ten economic-transformation questions, and the fifteen branch questions, the same constraints reprice the same class of claim to near zero. Wherever a candidate imagines intelligence itself as the prize, a constraint removes it, and what stands in its place is a complement.

Stated as the memorandum's coined law: **intelligence commoditizes; scarcity and value migrate to its complements.**

Two qualifications carry throughout.

The law now has two falsification conditions (Section 1.8). A law with no counterexample in thirty applications invites the charge of tautology or selection effect. **Both are now answerable, because the law carries conditions that would end it.** The condition published in Version 3.0 could not be evaluated at all, because it required the complement basket that Section 1.5 shows cannot be built. The two that replace it are evaluable on series this memorandum already publishes: a closing of the fixed-capability against frontier price divergence across two annual readings, and new entry above 3 GW into a Tier 3 complement at a price below 90 percent of its administered cap.

The thesis is bounded to this horizon. Section 16 grants the absorption objection: if organizational absorption is the terminal binding constraint, the residual accrues to whoever solves adoption rather than to owners of physical complements. The claim that value migrates to physical complements holds on the 2026 to 2030 horizon, where physical constraints dominate and absorption has not yet had time to bind. Beyond that horizon this memorandum does not make the claim.

14 The Call

One dated bet, publicly adjudicated, with five rules and a void condition. The 80 percent reliability horizon has been flat at 27 to 32 minutes across two frontier release cycles while the widely cited 50 percent series rose several-fold. Falsification requires four doublings in under two years.

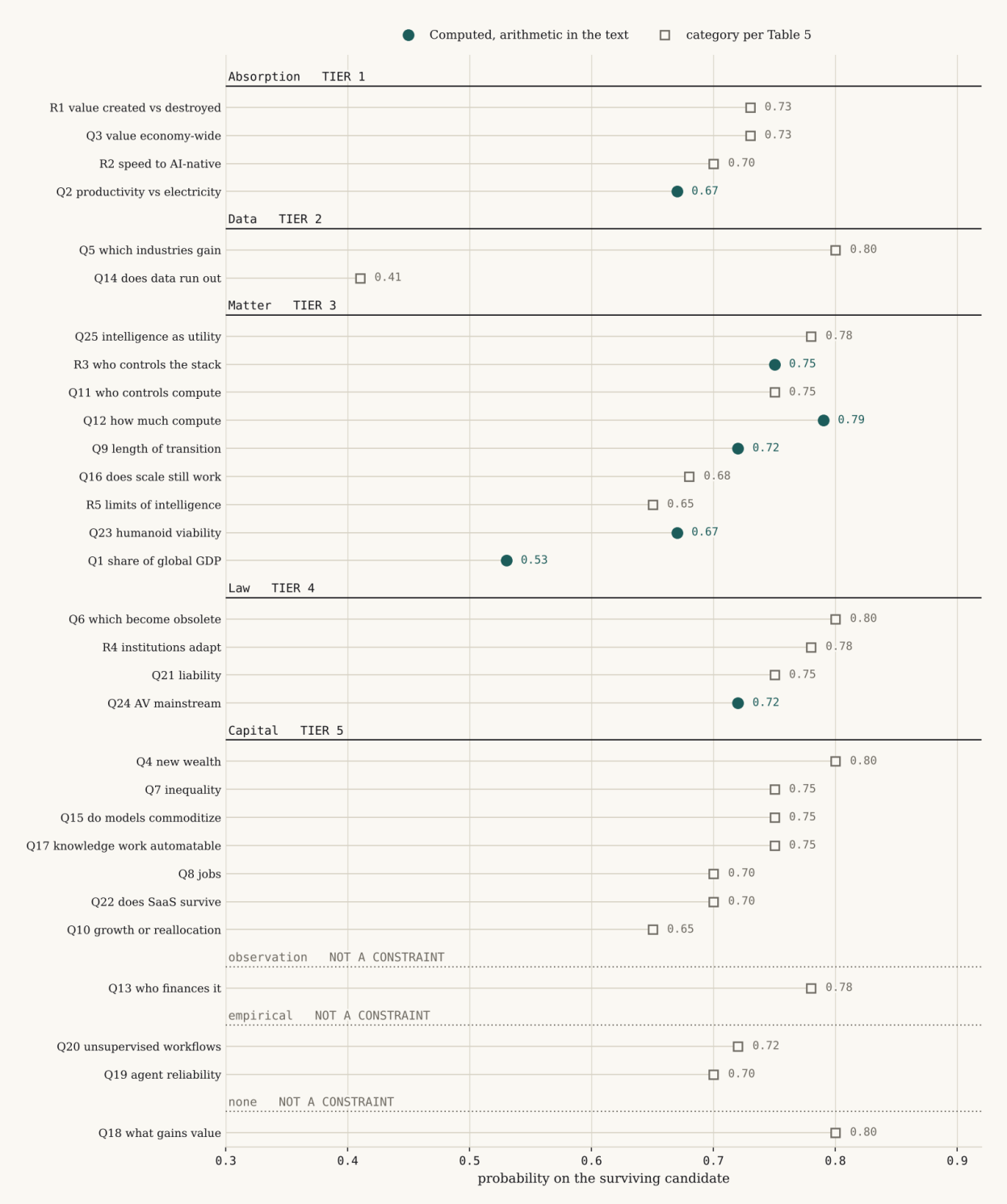


Figure 19: Every verdict, its binding constraint, and its tier. Matter binds nine and Absorption four, thirteen of the thirty between them. Of the thirty, fourteen are computed, eight decorative, five judged and three empirical, as Table 5 sets out; the seven marked here are the Computed verdicts whose arithmetic is published in the text, which is the scope Table 5 gives that phrase. The probabilities run 0.41 to 0.80 and most sit in the high sixties and seventies, which is itself a finding: a forecaster whose every answer is *probably* is barely forecasting, and widening this distribution is the next increment this program will deliver.

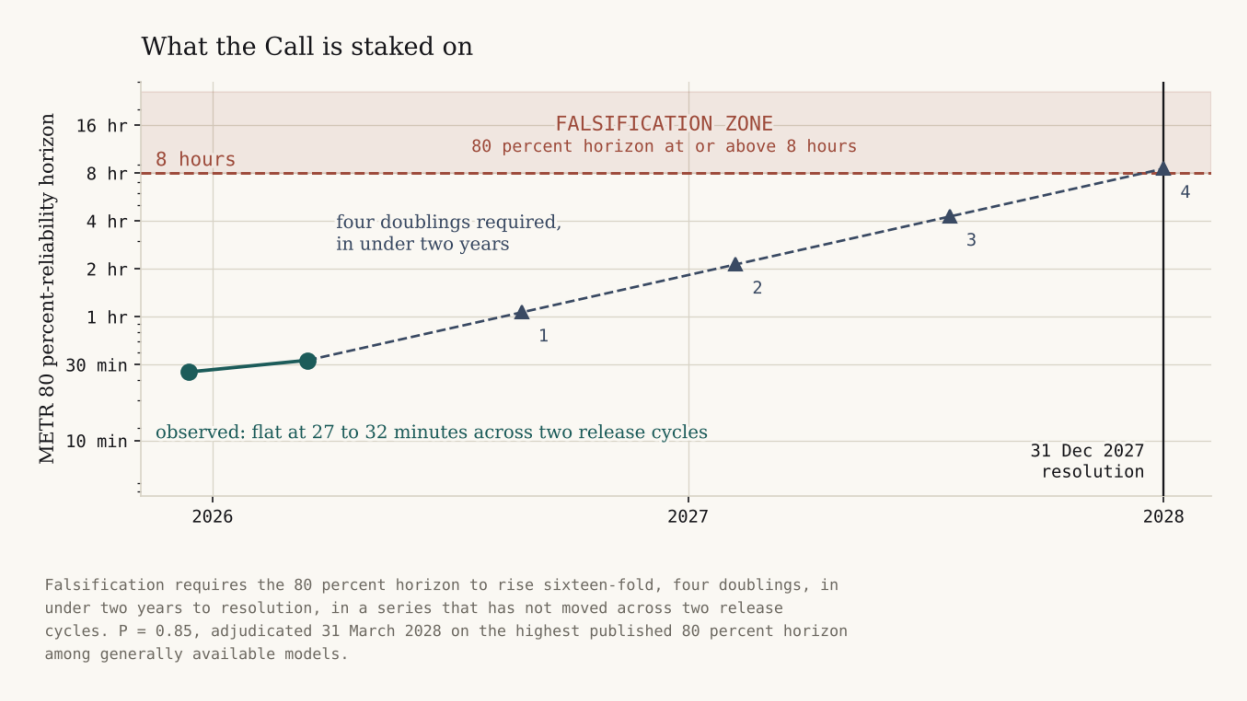


Figure 20: What the Call is staked on. Falsification requires the 80 percent horizon to rise sixteen-fold, four doublings, in under two years, in a series that has not moved across two release cycles. The shaded region is the falsification zone.

THE HEADLINE CALL

Reliability, not capability, remains the binding constraint on AI agents through 2027. On December 31, 2027, the best generally available AI model's METR [9] 80 percent-reliability task-completion time horizon is under eight hours.

Probability 0.85 · Adjudication date March 31, 2028 · Signpost S13

Basis. The 50 percent horizon has risen several-fold since late 2025. The 80 percent horizon has not. Claude Opus 4.5 measured approximately 27 minutes in December 2025, and the 80 percent series had been approximately flat at 27 to 32 minutes across the preceding frontier releases (SPEC). Falsification therefore requires roughly a sixteen-fold rise, four doublings, in under two years, in a series that has not moved across at least two release cycles.

The basis was re-checked for this version and is unchanged. METR's published time-horizon page carries a last-updated date of May 8, 2026, the same state of the source on which Version 3.0 relied; no revision has been issued in the interval. The newest model named on that page is not generally available under rule 3 below, so it does not qualify under the test this Call sets. The 27 to 32 minute band therefore stands, labeled SPEC as publisher-reported rather than independently recomputed here, and the threshold, the probability and the adjudication rules are carried forward without change. One precision note travels with the Call: the per-model 80 percent values are published in the source's machine-readable data files rather than in its text, and those files were not reachable in the current pass, so the band is carried on the publisher's page as read for Version 3.0.

Adjudication rules. A bet that cannot be settled is not a bet.

1. **Series.** Adjudicated on METR Time Horizon 1.1 [9] or its successor, as published at `metr.org/time-horizons`.
2. **Methodology change.** If METR replaces the series before resolution, the Call adjudicates on the successor. If the successor reports no 80 percent horizon, the Call resolves void rather than in either party's favor.
3. **Generally available.** A model is generally available if any member of the public can obtain access by payment without individual negotiation, waitlist approval, or preview enrollment, as of December 31, 2027.

4. **Unmeasured models.** Adjudicated on the highest published 80 percent horizon among generally available models. If the strongest generally available model has no published horizon, the Call resolves on those that do, and the omission is recorded.
5. **Adjudication date.** March 31, 2028, to allow for publication lag.

Two hazards, both published in advance, and one of them cuts the author's way. METR has noted that measurements above 16 hours [9] are unreliable with the current task suite, which makes the Call harder to settle. But rule 4 cuts the other way, and Version 3.0 claimed otherwise. **Resolving on the measured subset when the strongest generally available model is unpublished lowers the adjudicated maximum, which makes "under eight hours" more likely to resolve in the author's favor.** Coverage was not comprehensive as of May 2026: several frontier models had no published horizon while a non-generally-available preview model did. The rule is kept, because the alternative is a Call that cannot be settled at all, and the direction of that bias is published with the Call. If the strongest generally available model is unmeasured at resolution, the register records the Call as resolved *on a restricted field*, and that notation travels with the result.

The signature bet, restated. A market-value formulation of this bet, comparing the capitalization of the three leading high-bandwidth-memory makers against the two leading frontier labs, is tempting and unsound. The memory makers are diversified conglomerates whose capitalization is mostly not HBM, so such a bet could be won by a handset business, and frontier-lab valuations are private marks with no closing price. The bet is therefore stated on revenue, disclosed on both sides:

> **Signature bet.** Combined high-bandwidth-memory revenue at SK hynix, Samsung Semiconductor and Micron for fiscal 2027 exceeds combined total revenue at the two leading frontier laboratories for calendar 2027. Adjudicated on company disclosure and TrendForce estimates, at prevailing exchange rates on December 31, 2027. **Probability 0.55.**

Both sides are reported figures. The claim tested is the memorandum's thesis rather than a proxy for it. And the bet is genuinely hard, because frontier-lab revenue is growing extremely fast.

15 Retrospective Validation

Five backtest questions with parameters fixed at their historical as-of dates and constructed in 2026, four correct. **Brier is reported by tier and the aggregate only beside it: 1.000 on Tier 2 with one observation, 0.035 on Tiers 3 and 5 with four, against an aggregate of 0.228 and 0.25 for a coin flip.** The single miss occurred at maximum confidence on the tier this framework had already flagged as most exposed to substitution, and it forced a parameter revision that moves three verdicts in this document.

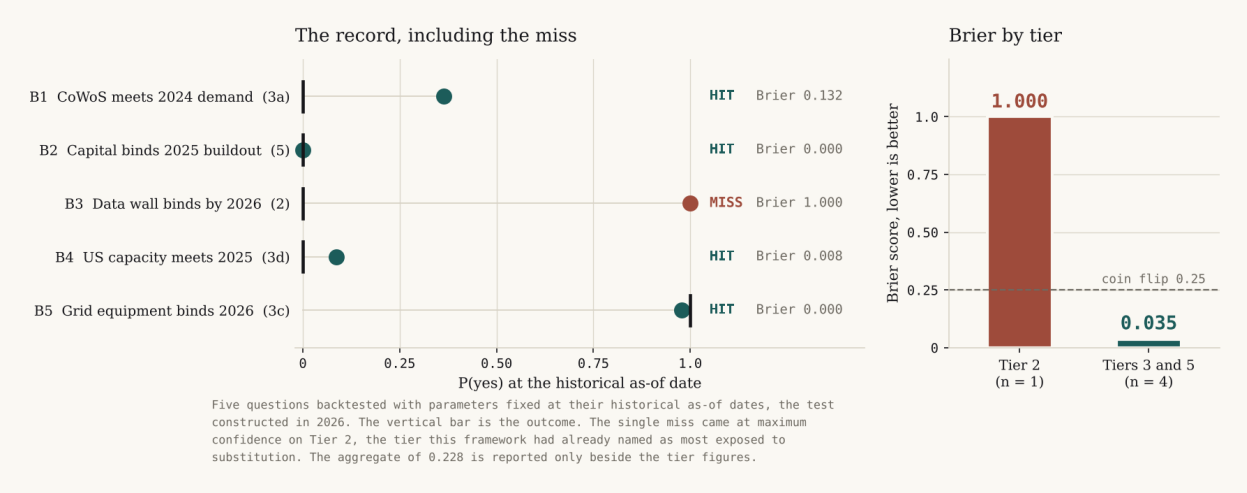


Figure 21: The record, including the miss. Five questions backtested with parameters fixed at their historical as-of dates and the test constructed in 2026. The vertical bar is the outcome. The single miss came at maximum confidence on Tier 2, the tier this framework had already identified as most exposed to substitution.

What these five are, stated precisely. They are a backtest, not a pre-registration. The parameters were fixed at historical as-of dates, but the exercise itself was constructed in 2026, after the outcomes existed. That is a different instrument from a pre-registered forecast, and this version marks the distinction wherever the five are described, which is more than the field's standard practice. The only pre-registered forecasts in this document are the Call and the signature bet of Section 14, whose outcomes do not yet exist, and the signposts of Section 17.

A framework that has never been scored is falsifiable in principle and not in practice. B3 was chosen specifically because the framework predicts its own failure there.

Table 43: Retrospective Validation. Brier reported by tier, never as an aggregate alone.

Question	As-of	Tier	P(yes)	Predicted	Actual	Brier
B1 CoWoS packaging meets 2024 AI demand	end-2023	3a	36.4%	NO	NO, HIT	0.132
B2 Capital is the binding constraint on 2025 buildout	end-2023	5	0.0%	NO	NO, HIT	0.000
B3 Public-text data wall binds frontier scaling by 2026	end-2022	2	100.0%	YES	NO, MISS	1.000
B4 US delivered capacity meets 2025 announcements	end-2023	3d	8.7%	NO	NO, HIT	0.008
B5 Grid equipment is binding on 2026 buildout	end-2024	3c	97.8%	YES	YES, HIT	0.000
Tier 2 ($n = 1$)						**1.000**
Tiers 3 and 5 ($n = 4$)						**0.035**
Aggregate, reported only beside the above ($n = 5$)						0.228

Four of five correct. **The framework is excellent on Tiers 3 and 5, where constraints are industrial and expansion rates are directly observable, at a Brier of 0.035, and it failed on Tier 2, where substitution dominates, at 1.000.** The aggregate of 0.228 would mislead in both directions and is reported only alongside the tier figures, never on its own.

The miss landing on the predicted tier is the most important result in this document. The framework predicted, at maximum confidence, that the public-text data wall would bind frontier scaling by 2026. It did not. Epoch's own median exhaustion estimate moved from 2024 to 2028, and multimodal corpora [5, 8] and verified synthetic generation substituted for the stock the model had treated as fixed.

This is the classical failure of resource-limit forecasting, reproduced inside our own model, on the tier we had already flagged as exposed to it. Malthus [67, 68], the Ehrlich side of the Simon-Ehrlich [68] wager, and The Limits to Growth [66, 65] against the Nordhaus critique all failed identically: the analyst held Permitted fixed while the world held it variable. Table 3 classifies Data as a stock that is drawn down and not replenished on the horizon, and Objection 8's reply names that partition the fossil one. The framework named the mechanism and under-parameterized its escape routes, which are substitutions. **Naming the exposure in advance is what made the miss legible as evidence rather than noise.**

Three corrections, all adopted in this version.

First, r_{data} is set at 0.30 to 0.60 rather than 0.02 to 0.05, because the parameter must be the rate at which the effective stock expands including substitution. Table 4 carries the revision, and Q14, Q16 and R5 are re-run against it in Sections 6.1, 7.2 and 3.5.

Second, **the Brier score is reported by tier, with the aggregate only beside it.** Removing B3 gives a mean of 0.035 across the remaining four. An aggregate figure alone would mislead in both directions, and Version 3.0 stated that rule in this section while reporting the aggregate alone in its abstract, contents and headline results. That inconsistency is removed.

Third, any Tier 2 verdict must now state its substitution assumption explicitly. Q14 does so.

The precise scope of what the backtest establishes. Five questions is not a validation set, four of the five are physical-infrastructure questions, the domain where the method is expected to perform best, and the test was constructed after the outcomes were known. **Two of the five are close to definitional.** B2 asks whether Capital binds and B5 whether grid equipment binds; those are the tier ordering restated as questions, so they could not have resolved against the framework without the framework collapsing, and they carry little independent information. The informative ones are B1, B3 and B4, which is three questions and one miss. The method is untested on institutional, labor and behavioral questions, which is most of Domain I, and the register will establish how it performs there. Extension to twelve questions including four institutional cases is in progress, and every miss will be published.

16 Objections and Replies

Objection 1. The complements you call scarce are scarce only for now. Cheap intelligence induces demand, and a Jevons dynamic pulls capital into the constrained inputs, so supply expands and prices ease.

Reply, with the instance named and measured. Advanced packaging is the confirmed case. CoWoS capacity has gone from approximately 13,000 wafers per month at end-2023 toward a 120,000 to 130,000 target at end-2026, and the supply-demand gap is reported narrowing from approximately 20% toward 10%. That is the migration this memorandum predicted, arriving on schedule, and the register logs it as a confirmed prediction. But the binding complements respond at different rates, and the rates are now parameters rather than claims: packaging expands at 60 to 110 percent annually while grid equipment expands at 10 to 20. The ratio between those two rates is the entire content of the migration thesis. Intelligence deflates by somewhere between tenfold and seventy-five-fold a year depending on the series; grid equipment cannot.

Objection 2. This is old wine. The complement thesis restates picks and shovels, or commoditize-your-complement.

Reply. The lineage is real and worth naming, and the sharpest available answer comes from the most-cited industry compilation in the field. BOND Capital's [72] *Trends: Artificial Intelligence* reaches the premise exactly, observing that training is expensive, serving is getting cheap, pricing power is slipping, and that the economics of general-purpose models look like commodity businesses with venture-scale burn. It then declines to draw the conclusion, calling the question of which business models survive a riddle [72]. The heuristic is common property; arriving at the premise and stopping is the field's normal state. **Five things separate this from the slogans:** a fixed-capability price series replicable from published code at $R^2 = 0.96$; a documented account of why the complement side is not measurable at all; a systematic engine applied across the thirty questions carried here at depth, the hundred-question classification battery having been the earlier version of this program; two falsification conditions for the law itself; and a scored backtest that reports a miss. A heuristic that cannot be wrong is not a competitor to a measure that can.

Objection 3. Your GDP band is anchored to pre-LLM precedents. AI automates cognition itself, so the historical adoption-lag precedent may not bind, and you may be under-forecasting by an order of magnitude.

Reply. The conservatism is a real exposure, it is published, and signpost S18 carries the threshold that would settle it. The defense is that the eliminations producing the band are physical and organizational, not analogical: interconnection medians above five years and a 5 percent P&L-positive rate constrain realization regardless of underlying capability. The band forecasts realized output, gated by deployment. And Section 15 shows the framework has already been caught making the opposite error in the other direction, at Tier 2, so this objection is not hypothetical: **it is the same failure mode with the sign flipped, which is why the register tracks both directions.** If deployment constraints ease, S11 and S15 trigger and the forecast moves up.

Objection 4. Your bound on agent autonomy rests on today's reliability. A jump to 99 percent per step lifts a twenty-step task from 36 to 82 percent, a discontinuous unlock.

Reply, and the evidence has moved. First, the sharpest available form of this objection, raised here against ourselves: **the compounding model of Q20 assumes independent per-step failure with no recovery**, and production agent systems use retries, verifiers and checkpointing, every one of which breaks independence and raises the effective chain success rate above 0.95^{20}. The arithmetic is a floor on unrecovered chains, not a ceiling on engineered ones. What it bounds is the cost of recovery, which rises with chain length, and that is the honest form of the claim. This remains signpost S2 and the register's most consequential trigger. The evidence has moved toward this memorandum's position rather than away from it: the 80 percent horizon has been approximately flat at 27 to 32 minutes across two frontier release cycles while the 50 percent horizon rose several-fold. **The last reliability nines are proving harder to win than the headline trend implies, and the widening ratio between the two horizons is the direct measurement of that.** If S2 triggers, the verdict flips and this memorandum said so in advance.

Objection 5. Your layered-oligopoly verdict fails against vertical integration. One player already owns custom silicon, a cloud, a frontier lab, and models.

Reply. Vertical integration is real and tracked. The elimination holds at the deepest physical layer: even the most integrated player depends on a single external foundry for advanced fabrication and on the public grid for power, and owns neither. Integration compresses the middle of the stack; it does not close the floor.

Objection 6. Your data wall is obsolete on arrival. Synthetic data, reinforcement learning from verifiable rewards, and multimodal corpora [5, 8] are already pushing past it.

Reply. The objection was right, and the backtest is how we know. This objection was correct, and Section 15 is the measurement that proves it, scored and published by the framework it indicts. The verdict has been revised: r_{data} raised

by an order of magnitude, candidate (ii) of Q14 upgraded from eliminated to a live tail, and the survivor's probability cut twice, from 0.60 to 0.52 and then to 0.41, as Section 6.1 sets out in full. What replaces it is narrower, measured and falsifiable: synthetic substitution works where answers are machine-verifiable and risks collapse where verification is hard [8, 7], so the scarce complement migrates from raw data to verification apparatus rather than disappearing.

Objection 7, which bounds the thesis to where it holds. If organizational absorption is the terminal constraint, your investment conclusion is wrong. Every historical analogue this memorandum invokes has the same structure. In electrification, generation and distribution were solved decades before productivity gains arrived, because factory reorganization was the actual bottleneck. In the 1990s telecom buildout the constraint moved from capital to demand, with roughly 5 percent of installed fiber [62] lit by 2001. If the same holds here, the residual accrues to whoever solves organizational adoption, not to owners of physical complements.

This objection is correct, and acting on it makes the thesis sharper. Absorption now sits at Tier 1 in the constraint set, where it is the binding constraint in R1, R2, Q2 and Q3, four of the thirty, with Q6 binding on Law and Q9 on Matter. The thesis is narrowed accordingly: **the claim that value migrates to physical complements holds on the 2026 to 2030 horizon, where physical constraints dominate and absorption has not had time to bind. Beyond that horizon this memorandum does not make the claim.** R2 is the internal evidence: it concluded that the bottleneck is organizational and had nowhere in the old constraint set to record it.

Objection 8. Your constraint set is selected from your conclusion. Energy, matter, data and talent are simultaneously the test battery and the answer, so the method tests candidates against scarce complements and concludes that value accrues to scarce complements.

Reply, with the open half stated precisely. Section 1.2 derives the six from a question that never mentions AI: what does it take to obtain one more unit of a limiting input, partitioning into forbidden, biological, fossil, manufactured, permitted, purchased. That derivation is independent of the thesis. The remaining half is a counterexample, and the conditions that would produce one are now published. Across the thirty applications carried here there is no question where the binding constraint is one of the six and value nonetheless accrues to the intelligence layer, and the two falsification conditions of Section 1.8 are the standing invitation to produce one.

17 Signpost Register

Signposts are four distinct kinds of object, and separating them is what makes the register testable. Constraint signposts test whether a wall moved. Coupling signposts test whether a conversion lag changed. State signposts test where the system sits inside the walls and falsify verdicts without moving any constraint. Meta signposts test the framework rather than the world: S16 watches whether the tier ordering is stable, and S19 watches the parameter on which every output verdict depends. The register runs nine Constraint, five State, four Coupling and two Meta.

Table 44: Signpost Register. Twenty tracked signposts, S1 to S20.

ID	Falsifies	Kind	Constraint	Tier	Metric and trigger	Cadence
S1	Data wall (Q14, Q16, R5)	Constraint	Data	2	Frontier dataset size vs. effective stock including substitution; trigger: verified-synthetic share of frontier training exceeds 50%, or datasets exceed ∼100T tokens	Semiannual
S2	Agent reliability (Q19, Q20)	State	none	–	METR 80% horizon; production single-task success. Trigger: 80% horizon reaches multi-hour AND production success sustainably above 90%	Quarterly
S3	Financing cascade (Q13)	Coupling	Capital to Matter	5 to 3	AI-linked credit spread; hyperscaler bond order coverage; correlated defaults. Trigger: forced refinancing at a top-5 buildout, spread above 150bp, or coverage below 1.5x	Monthly

ID	Falsifies	Kind	Constraint	Tier	Metric and trigger	Cadence
S4	Compute control (Q11)	State	none	–	Incumbent accelerator revenue share below ∼70%	Quarterly
S5	Model commoditization (Q15)	State	none	–	Open-weight parity, or the price decline halting	Quarterly
S6	Liability regime (Q21)	Constraint	Law	4	A bespoke AI-agent liability statute, or a ruling shifting liability to developers	Semiannual
S7	SaaS repricing (Q22)	State	none	–	Incumbent seats fall >15% YoY, or AI-native ARR growth stalls >50%	Quarterly
S8	Humanoid economics (Q23)	Coupling	Capital to Matter	5 to 3	Build cost below ∼$30k AND deployed base above ∼100k units. First half already met: entry units are near $16k (Q23), so the deployed base is the binding half	Semiannual
S9	AV mainstreaming (Q24)	Constraint	Law	4	Above ∼2M weekly rides across 30+ metros, or a vision-only unsupervised launch at scale	Quarterly
S10	Capital gap (R1, Q3, Q13)	Coupling	Capital adequacy	5	Annual AI revenue run-rate above $400B by 2027. This requires a 2.85x rise from the $140.4B 2026 services anchor against an observed 4.73x, so it is demanding by construction; the informative reading is the trajectory against the Bain path	Quarterly
S11	Power ceiling (Q1, Q9, Q12, Q25)	Constraint	Matter 3c, 3d with Law	3, 4	LBNL median interconnection below 3 years; OR US transformer lead time below 80 weeks; OR flexible-load interconnection above 20% of new large-load agreements	Annual to quarterly
S12	Distributional backlash (R1, Q7)	Constraint	Legitimacy	4	Any G7 AI-displacement tax or deployment moratorium enacted	Quarterly
S13	The Call	State	none	–	METR 80% horizon reaches 8 hours by December 31, 2027	Quarterly
S14	Thermodynamic headroom	Constraint	Thermodynamics	0	Joules per million output tokens [3] at fixed quality. **What moves is the distance from the Tier 0 floor, not the floor**, which cannot move. Trigger: annual improvement below 20% across two generations, or above 60%	Semiannual
S15	Talent stock and flow	Constraint	Talent	1	US share of elite AI researcher employment; H-1B [56] and O-1 issuance to AI employers; international STEM PhD enrollment. Trigger: US share below 50%, or two consecutive years of double-digit decline in inbound flow	Annual

ID	Falsifies	Kind	Constraint	Tier	Metric and trigger	Cadence
S16	Tier stability (meta)	Meta	all	–	Observed relaxation time per constraint. Trigger: any r_c moves by a factor above three. **Status: FIRED.** Section 15 revises r_{data} from 0.02–0.05 to 0.30–0.60, a factor of twelve to fifteen, which is the first and largest trigger on record. Second datum: the Digital Omnibus deferring a flagship deadline 16 months in 7 months of process	Semiannual
S17	Correlated stress	Coupling	Capital, Matter, Law jointly	–	Co-movement of AI-linked spreads, interconnection withdrawal rates, equipment order cancellations. Trigger: any two moving adversely beyond one standard deviation in the same quarter	Monthly
S18	Absorption (Q2, Q9, R2)	Constraint	Talent and Absorption	1	US labor productivity growth with AI attribution. Trigger: growth above 2.5% for three consecutive years with credible attribution falsifies the electricity-slow verdict	Annual
S19	The Elasticity Gap (Sec. 2)	Meta	all output verdicts	–	Implied α from attributable output against deployed capacity. **Status: dormant, not currently evaluable.** The hardware-excluded numerator it requires cannot be built from disclosed data (Section 2.6), and on the hardware-inclusive series the trigger is already met, since on the adopted convention the three readings are 0.77, 0.81 and 1.26, and the trigger is not evaluable on a series the document treats as an accounting identity. Retained as the specification of the measurement that would close the question. Trigger, once evaluable: α stabilizing outside 1.2 to 1.5 for two consecutive years	Annual (dormant)

ID	Falsifies	Kind	Constraint	Tier	Metric and trigger	Cadence
S20	Tier 3 classification of power (Sec. 1.6)	Constraint	Matter 3c, 3d	3	New-entry response in capacity auctions. Trigger: PJM or an equivalent market clearing a subsequent auction with new entry above 3 GW at a price below 90 percent of the administered cap falsifies the Tier 3 classification of delivered power, and with it the second falsification condition of the law. **Both halves are now quantified**; Version 3.0 wrote "substantial new entry at or below the cap," which excluded nothing, since clearing prices cannot exceed the cap. The 3 GW threshold is roughly half the 6.8 GW by which the 2028/29 auction fell short of its reserve-margin target, so it is a response large enough to matter and smaller than full closure; 90 percent separates a market clearing on economics from one clearing at the net cost of new entry	Per auction

Five of the twenty are state signposts, including the Call, and the register now says so. By kind the register runs nine Constraint, five State, four Coupling and two Meta. The register carries a check log that records missed checks as well as completed ones. A framework that has never been wrong in public has never been tested in public, and Section 15 is the first entry.

18 Sources, Reconstructions, and Limitations

This memorandum forecasts by elimination, so its force rests on the arithmetic rather than on any single figure. Standing is stated per source, so a reader can weight each one.

Primary-sourced: the interconnection queue (LBNL *Queued Up: 2026 Edition*), the capability horizon (METR), the data stock (Epoch AI), the EU regulatory timeline (Official Journal and Council records), the thermodynamic bound (Ho, Erdil and Besiroglu). **Rebuilt from filings:** the capital-revenue gap. **Estimates reproduced across outlets, attributed as such:** the CoWoS capacity series and most semiconductor allocation figures, which mostly trace to a small number of sell-side houses. **Labeled reconstructions [54]:** the GDP ceiling, concentration, humanoid, robotaxi, and software charts. **Estimated, not rebuilt from filings:** the attributable-revenue side of the capital-revenue gap, whose capex side is filed and whose revenue side is not; Version 3.0 described the whole series as rebuilt from filings and Figure 11's caption repeated it. **Carried with an explicit caveat:** the AI-generated web-page share, which could not be traced to a stable primary in the current pass; the elite-researcher share figures, whose primary tracker was discontinued and which are labeled EST in Table 4 accordingly; the installed AI data-center power capacity of approximately 29.6 GW, which is labeled EST because published values for Q4 2025 span roughly 20 to 31 GW across sources and none is a disclosed figure; and the attributable-output series underlying the elasticity of Section 2, whose numerator includes hardware vendor revenue and which could not be rebuilt from disclosed data in this pass.

Citation discipline, and the two mechanisms that produced every failure. Verifying this bibliography against source produced twelve corrections. **Every substantive figure held up. What failed was the pointer**, and the failures reduce to exactly two mechanisms, both mechanical and both preventable.

Aggregator laundering. A number reported by the AI Index or Epoch AI gets attributed to the origin paper of the benchmark it concerns. The OSWorld case is the clean example: the paper reports 12.24 percent for the best model against 72.36 percent for humans, the widely quoted 66.3 percent is AI Index Figure 2.6.2 sourced to Epoch for Claude

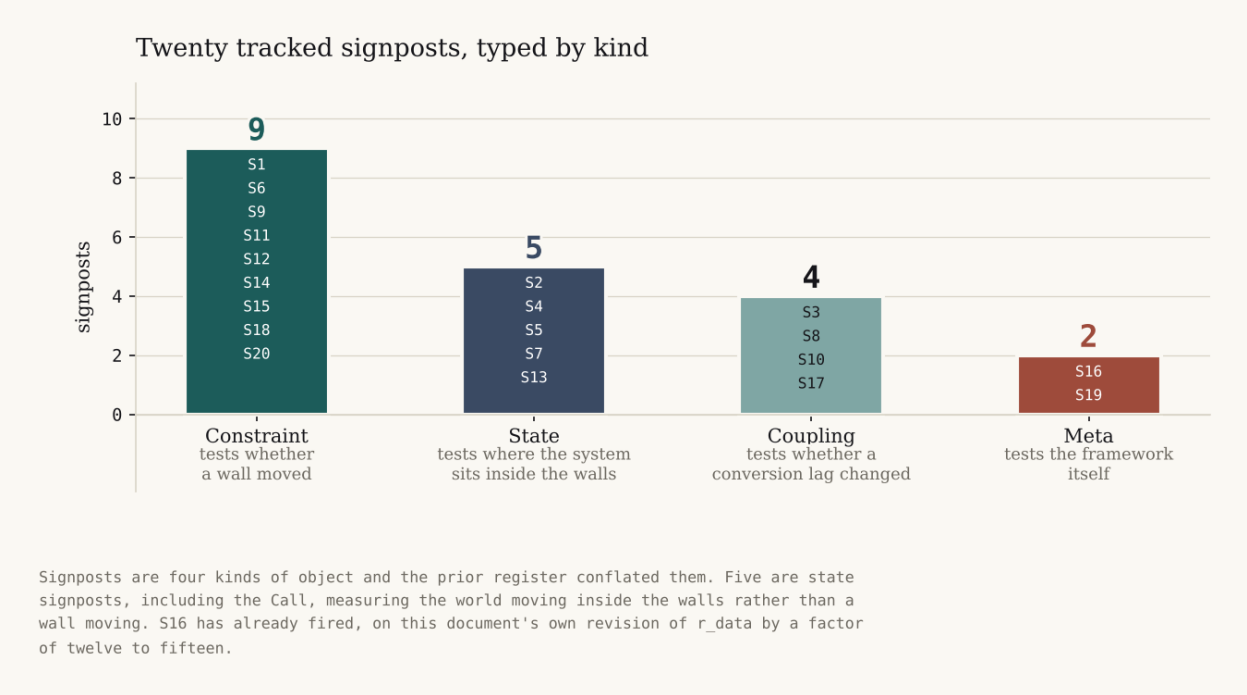


Figure 22: Twenty tracked signposts, typed by kind. Five of them are state signposts, measuring the world moving inside the walls rather than a wall moving, and the Call is one. A register that does not distinguish the two invites over-reading of every trigger.

Opus 4.5, and the live leaderboard stood near 85 percent by July 2026. Three numbers, three sources, and only one of them is the paper. **The rule that prevents it: cite the source of the number, not the source of the concept.**

Version drift. A benchmark paper is cited by its arXiv identifier while its figures move underneath. METR's time-horizon paper was retitled between versions. TheAgentCompany's best model moved from 24.0 and 34.4 percent to 30.3 and 39.3 percent as the leading system changed from Claude 3.5 Sonnet to Gemini 2.5 Pro. The identifier resolved correctly throughout and the numbers were wrong throughout. **The rule that prevents it: pin the version, and treat a living benchmark as a series rather than a citation.**

Both rules are now applied throughout this bibliography. A third, from an earlier pass, governs identifiers themselves: **an identifier is a claim like any other, and one that was not looked up must not be asserted.** A previous version withdrew four JSTOR identifiers as a class on suspicion; checking them individually showed one was correct, one was wrong by a near-miss, and two were unverifiable. Withdrawing a category without a reason is as lazy as asserting one, so the reason is published. Works whose identifiers cannot be confirmed are cited by title, author and venue, which is checkable in one search and honest about what is being claimed.

Two figures in general circulation are corrected here rather than repeated, and one of this memorandum's own is withdrawn. The $1.5T financing figure is a funding gap and not a debt requirement, and reading it as a debt requirement overstates the capital constraint roughly sevenfold. The "4 to 8 year" interconnection range is not a figure LBNL publishes; the published median exceeded 5 years for projects built in 2025. And this memorandum's own Tier 2 parameterization, the one that produced backtest miss B3, has been revised by an order of magnitude, changing three verdicts.

The forecast register is open. It carries eighteen entries: this memorandum's five backtest results, including the maximum-confidence miss; eleven dated forecasts from the wider field, two of which have resolved by revision rather than by outcome; and the two staked forecasts of Section 14. Seven are resolved, the five backtests and the two revisions. The two staked forecasts of Section 14, the Call and the signature bet, are the seventeenth and eighteenth; the Call is also tracked as signpost S13, and the signature bet, which Version 3.0 described as registered without placing it anywhere, is entered here. Each entry states a claim, a source of truth, a resolution date and a status, and every entry carrying a stated probability is Brier-scored on the same basis. The author's own forecasts, the five backtests and the two staked calls, are listed first and last. The two resolved by revision are Epoch AI moving its own data-exhaustion median from 2024 to 2028, and the AI Futures Project revising its timeline outward in December 2025. Both are recorded as resolutions,

because revising a public forecast when the evidence changes is the behavior this register exists to reward. Reviews are scheduled, and a missed review is logged as a miss rather than backfilled. See `REGISTER.md`.

Three things convert this into a field-defining work, and all three require inputs the author must supply: a proprietary dataset no reader can source elsewhere, for which Section 2.6 now gives the specification, namely AI services output against deployed compute by firm or sector with hardware vendors excluded, and which is the most valuable single item on this list; documented adversarial review from people with domain expertise and no stake in the conclusion; and a live tracker carrying a running Brier score across every verdict. The first is buildable now and Section 2.6 specifies it. The second is underway. The third accrues on its own, and the register is already open to receive it.

Three smaller inputs are listed here because the body points to this section for them. A confidence coefficient derived from the rubric for each of the thirty verdicts, where five carry one today. A maintained replacement for the discontinued elite-researcher tracker of reference [55]. And the per-verdict posterior shifts that would name all thirty members of the Table 5 categories, twelve of which currently live only in the repository.

19 Parting Thoughts

A wealth of information creates a poverty of attention. Herbert A. Simon, 1971

Most of the argument about artificial intelligence is an argument about the model: whether it will wake, when it will match or exceed us, whether that arrival is rescue or ruin. The two loud camps disagree about nearly everything except the premise they share, that the model is the thing that matters. This memorandum has taken the opposite premise. It has asked, thirty times over, not what the machine will become but what it will cost, what it will consume, and who will be paid.

What the constraints leave standing is physical, and that is the point. Intelligence is becoming abundant, and abundance is not where value collects. Value collects in whatever the abundance makes scarce by comparison: the power to run the models, the verified data to feed them, the fabrication and memory to build them, the physical capacity to act in the world, and the accountable person who can be trusted and, when wrong, be held to account.

One feature of this memorandum matters more than the rest. **It predicted where it would be wrong, and then was wrong there.** Section 15 shows the method predicting, at maximum confidence, a data wall that did not arrive, and failing in exactly the way it had already named as its own characteristic failure: holding a resource stock fixed while substitution moved the ceiling. The parameter is corrected, three verdicts move, and the register carries the score by tier: 1.000 where the method was exposed, **0.035 where it claims competence**.

That is the first hard evidence that this is a method and not a posture. A forecast that cannot fail is not a forecast. This one has failed once already, in public, on the tier it warned about, and it has told you what it changed as a result.

The central claim is staked on a public number with a public deadline: the best generally available model's eighty-percent-reliability horizon, under eight hours on the last day of 2027, adjudicated on March 31, 2028 under five stated rules. Should it clear eight hours, the claim failed, and the register will say so on schedule, without argument.

Whether the machine wakes is the exciting question, and it will be debated in rooms louder than this. Who it pays is the useful one, and it is being decided now, quietly, in interconnection queues and transformer order books and the slow accrual of trust. Superintelligence may or may not arrive on schedule. The electricity bill, the memory, and the trust will arrive regardless.

Acknowledgments

External review is the next thing this work needs, and the apparatus for it is already built: Section 18 records documented adversarial review as one of three things separating this from a field-defining work. Readers who find an arithmetic error are invited to report it, and every report will be logged publicly whether or not it changes a verdict.

References

1. R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* 5(3), 1961.
2. A. Ho, E. Erdil, and T. Besiroglu. Limits to the Energy Efficiency of CMOS Microprocessors. IEEE ICRC, 2023. arXiv:2312.08595. Maximum CMOS efficiency approximately 4.7e15 FP4/J, roughly 200x current accelerators, about two orders of magnitude of headroom. The claim that accelerators sit six orders of magnitude above the Landauer bound is not supported by this source and has been withdrawn.

3. Epoch AI. Machine learning hardware database and trends in GPU price-performance. epoch.ai/data.
4. J. Koomey et al. Implications of historical trends in the electrical efficiency of computing. *IEEE Annals of the History of Computing* 33(3), 2011.
5. P. Villalobos et al., Epoch AI. Will We Run Out of Data? Limits of LLM Scaling Based on Human-Generated Data, arXiv:2211.04325. Approximately 300T tokens of effective public human text, median exhaustion 2028. The 80% interval of 2026 to 2032 quoted elsewhere is not stated in this source and has been withdrawn from this memorandum.
6. I. Shumailov et al. AI models collapse when trained on recursively generated data. *Nature* 631, 2024.
7. M. Gerstgrasser et al. Is model collapse inevitable? Breaking the curse of recursion by accumulating real and synthetic data, 2024. arXiv:2404.01413.
8. International AI Safety Report 2026. arXiv:2602.21012.
9. METR. Task-completion time horizons of frontier AI models. metr.org/time-horizons, updated May 8, 2026. Re-checked for Version 3.1 on October 1, 2026: no revision had been published and the last-updated date was unchanged.
10. METR. Measuring AI Ability to Complete Long Software Tasks, arXiv:2503.14499v4, July 10, 2026. Version pinned: the paper was retitled and its figures revised after v1; earlier citations of this work in the literature refer to different numbers.
11. J. Becker, N. Rush, E. Barnes, and D. Rein, METR. Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity, July 10, 2025. arXiv:2507.09089. Randomized controlled trial, 16 developers, 246 tasks: developers forecast a 24% speedup, self-reported 20%, and were measured 19% slower. METR's February 2026 follow-up reports the successor experiment gives an unreliable signal because developers increasingly decline to work without AI.
12. T. Xie et al. OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments, arXiv:2404.07972, 2024. The paper reports 12.24% for the best model against 72.36% for humans. The widely quoted 66.3% figure is not from this paper: it is AI Index 2026 Fig. 2.6.2, sourced to Epoch AI, for Claude Opus 4.5. The live leaderboard stood near 85% by July 2026. Three distinct numbers, three distinct sources.
13. F. Xu et al. TheAgentCompany: Benchmarking LLM Agents on Consequential Real-World Tasks, arXiv:2412.14161, current version. Version pinned: best model has moved from 24.0% autonomous and 34.4% with partial credit (Claude 3.5 Sonnet) to 30.3% and 39.3% (Gemini 2.5 Pro).
14. J. Rand, A. Cheyette, C. Talley, S. Zhang, W. Gorman, R. H. Wiser, J. Seel, S. Jeong, and F. Kahrl. Queued Up: 2026 Edition, Characteristics of Power Plants Seeking Transmission Interconnection As of the End of 2025. Lawrence Berkeley National Laboratory, June 2026. emp.lbl.gov/queues. Approximately 8,200 active projects, 1,312 GW of generation and 749 GW of storage.
15. Lawrence Berkeley National Laboratory. 2024 United States Data Center Energy Usage Report, December 2024.
16. International Energy Agency. Energy and AI, April 2025; and Building the Future Transmission Grid, 2025.
17. International Energy Agency. Key Questions on Energy and AI, 2026 update.
18. Electric Power Research Institute. Powering Intelligence: Updated Scenarios of U.S. Data Center Energy Use, February 2026.
19. Wood Mackenzie. T&D Equipment Supply Chain Survey, Q2 2025: power transformers averaging 128 weeks, generator step-up units 144 weeks. Subsequent 2026 reporting (B. Boucher, Wood Mackenzie, via Data Center Knowledge, May 2026) puts substation transformer lead times above 160 weeks, up from roughly 140 in 2023, with US data-center electrical equipment demand rising from about $20B in 2026 to $65B by 2030.
20. GE Vernova. First quarter 2026 results and orders backlog.
21. PJM Interconnection. 2027/2028 Base Residual Auction results, December 2025; Monitoring Analytics, State of the Market.
22. US Federal Energy Regulatory Commission. Large load interconnection and co-location proceedings, 2025–2026.
23. TrendForce. CoWoS capacity and HBM supply estimates, 2024–2026.
24. Taiwan Semiconductor Manufacturing Company. Quarterly results, capital expenditure guidance and Arizona fab disclosures, 2020–2026.
25. SemiAnalysis. AI data-center and advanced packaging analysis, 2024–2026.
26. NVIDIA Corporation. Quarterly financial results, data center segment, 2023–2026.
27. Morgan Stanley. Bridging a $1.5tr Data Center Financing Gap, July 2025. The figure is a funding gap to 2028, not a debt forecast; the associated corporate debt issuance estimate is approximately $200B.
28. CreditSights and J.P. Morgan. Hyperscaler capital structure and AI-linked issuance, 2026.
29. Bain & Company. 6th Global Technology Report, September 2025.
30. D. Cahn, Sequoia Capital. AI's $600B Question, June 2024.
31. a16z. LLMflation: inference cost trends, 2024.
32. SK hynix. Form F-1 / DRS-A registration statement, US Securities and Exchange Commission, 2026. DRAM share of total revenue 63.4% (2023), 67.6% (2024), 77.1% (2025) on revenues of KRW 32,766bn, 66,193bn and 97,147bn. HBM share of DRAM revenue disclosed at 30% in Q3 2024 and above 40% in Q4 2024, after which publication ceased.
33. NVIDIA Corporation. Annual reports on Form 10-K, FY2020–FY2026. Gross margin 62.0, 62.3, 64.9, 56.9, 72.7, 75.0 and 71.07 percent; FY2023 depressed by $2.17bn of inventory provisions.

34. PJM Interconnection. 2028/2029 Base Residual Auction results, July 14, 2026. Cleared at $325.00/MW-day at the FERC-approved cap; approximately 525 MW of new resources; 6.8 GW short of the reserve margin target; cleared reserve margin 14.7%. Preceded by $28.92 (2024/25), $269.92 (2025/26), $329.17 (2026/27) and $333.44 (2027/28).
35. Epoch AI, Somala, 2026. AI chip component cost shares. Memory 51.87% to 63.42% of component cost, Q1 2024 to Q4 2025, $R^2 = 0.935$. Cited here as a series this memorandum tested and rejected, because it confounds quantity with price: HBM content per accelerator rose 3.6x over the same window, from 80GB to 288GB.
36. B. Cottier, B. Snodin, D. Owen, and T. Adamczewski, Epoch AI. LLM inference price trends at fixed benchmark capability, 2025. Replication repository `github.com/epoch-research/llm-benchmark-efficiency`. Prices are a 3:1 weighted blend of input and output tokens, and reasoning models are excluded because per-token price misstates cost per task for models that emit far more tokens. The series in this memorandum are computed by cloning and re-running that repository rather than quoting headline rates.
37. D. Acemoglu. The Simple Macroeconomics of AI. NBER Working Paper 32487, 2024; *Economic Policy*, 2025. TFP effects no more than 0.66% in total over ten years on the headline estimate, and under 0.53% on his adjusted estimate.
38. Goldman Sachs Research. The Potentially Large Effects of Artificial Intelligence on Economic Growth, March 2023; and realized-productivity updates, 2026.
39. J. Covello et al., Goldman Sachs Research. Gen AI: Too Much Spend, Too Little Benefit?, June 2024.
40. McKinsey Global Institute. The Economic Potential of Generative AI, June 2023; and The State of AI, 2025.
41. PwC. Sizing the Prize: What's the Real Value of AI for Your Business, 2017.
42. International Monetary Fund. World Economic Outlook, 2025–2026; and Gen-AI: Artificial Intelligence and the Future of Work, 2024.
43. G. Filippucci et al., OECD. Miracle or myth? Assessing the macroeconomic productivity gains from AI, 2025.
44. Penn Wharton Budget Model. The projected economic impact of artificial intelligence, September 2025.
45. US Bureau of Labor Statistics. Labor productivity and costs, 2025–2026.
46. E. Erdil et al., Epoch AI. GATE: modeling the trajectory of AI and automation, 2025; and Explosive Growth from AI: A Review of the Arguments.
47. World Economic Forum. Future of Jobs Report 2025, January 2025.
48. MIT Project NANDA. The GenAI Divide: State of AI in Business 2025, July 2025.
49. Gartner. Over 40% of agentic AI projects will be canceled by the end of 2027, press release, June 25, 2025.
50. E. Brynjolfsson, D. Li, and L. Raymond. Generative AI at work. *Quarterly Journal of Economics* 140(2), May 2025, 889–942, doi:10.1093/qje/qjae044; originally NBER Working Paper 31161, 2023. The published version reports 15% on 5,172 agents, revised from 14% on 5,179 in the working paper. Crucially, the effect is heterogeneous in a way the headline hides: less experienced workers improve both speed and quality, while the most experienced and highest-skilled see small gains in speed and small declines in quality.
51. Z. K. Cui, M. Demirer, S. Jaffe, L. Musolff, S. Peng, and T. Salz. The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers. SSRN 4945566, 2024. Weekly pull requests +26.08% (SE 10.3%), commits +13.55% (SE 10.0%), builds +38.38% (SE 12.55%). The standard errors are large and the effect differs materially across the three companies.
52. E. Brynjolfsson et al. Longitudinal estimates of consumer surplus from generative AI, 2026, reported in the AI Index 2026.
53. M. Svanberg and N. Thompson, MIT. Beyond AI Exposure: Which Tasks Are Cost-Effective to Automate?, 2024.
54. N. Maslej et al. Artificial Intelligence Index Report 2026. Stanford Institute for Human-Centered AI, April 2026.
55. MacroPolo, Paulson Institute. Global AI Talent Tracker. The primary tracker has been discontinued and provenance of the most recent edition is uncertain. Figures drawn from it are labeled EST throughout this memorandum, and replacement with a maintained source is listed in Section 18 among the inputs the author must supply.
56. US Citizenship and Immigration Services. H-1B registration data, FY2026–FY2027; Presidential Proclamation of September 19, 2025 and subsequent litigation.
57. McKinsey and SEMI. US semiconductor workforce analysis, 2024–2025.
58. Regulation (EU) 2024/1689 (Artificial Intelligence Act), as amended by Regulation (EU) 2026/1744, the citable instrument for the Digital Omnibus deferral, Council final approval June 29, 2026.
59. Directive (EU) 2024/2853 on liability for defective products, transposition deadline December 9, 2026. Extends strict liability to software and AI systems, subject to carve-outs for free and open-source software supplied outside a commercial activity, and non-retroactivity for products placed on the market before the transposition date.
60. Executive Order 14365, Ensuring a National Policy Framework for Artificial Intelligence, December 11, 2025.
61. NYU Center on Technology Policy. State AI and data-center legislation counts, via Tech Policy Press, July 2026.
62. P. A. David. The dynamo and the computer: an historical perspective on the modern productivity paradox. *American Economic Review* 80(2), May 1990.
63. E. M. Goldratt and J. Cox. The Goal, 1984; E. M. Goldratt, Theory of Constraints, 1990.

64. S. Bradley, A. Hax, and T. Magnanti. Applied Mathematical Programming, ch. 4, duality and complementary slackness. Addison-Wesley, 1977.
65. W. D. Nordhaus. World dynamics: measurement without data. *Economic Journal* 83(332), December 1973, 1156–1183. doi:10.2307/2230846, JSTOR 2230846. A previous version of this bibliography gave stable ID 2230285, which is wrong.
66. D. Meadows et al. The Limits to Growth, 1972.
67. T. R. Malthus. An Essay on the Principle of Population, 1798.
68. P. Sabin. The Bet: Paul Ehrlich, Julian Simon, and Our Gamble over Earth's Future. Yale University Press, 2013. The 1980 wager on five metals over ten years, which Simon won on all five.
69. W. S. Jevons. The Coal Question, 1865.
70. P. Tetlock and D. Gardner. Superforecasting: The Art and Science of Prediction, 2015.
71. A. C. Doyle. The Sign of the Four, 1890.
72. M. Meeker, J. Simons, D. Chae, and A. Krey. Trends: Artificial Intelligence. BOND Capital, May 30, 2025.
73. L. Aschenbrenner. Situational Awareness: The Decade Ahead, June 2024.
74. D. Kokotajlo et al., AI Futures Project. AI 2027, 2025; timeline revised December 2025.
75. S. Nadella, interview on the Bg2 Pod, late 2025, on power rather than compute as the binding constraint.
76. Anthropic. Economic Index, 2025–2026.
77. W. D. Nordhaus. Are We Approaching an Economic Singularity? Information Technology and the Future of Economic Growth. *American Economic Journal: Macroeconomics* 13(1), 2021; and *Two Centuries of Productivity Growth in Computing*, Journal of Economic History 67(1), 2007. The closest existing estimate of the output elasticity of information capital, with standard errors.
78. P. Aghion, B. F. Jones, and C. I. Jones. Artificial Intelligence and Economic Growth. In *The Economics of Artificial Intelligence*, NBER, 2019. The Baumol result: growth is governed by the task that is hardest to improve, not the easiest.
79. P. M. Romer. Endogenous Technological Change. *Journal of Political Economy* 98(5), 1990; and C. I. Jones, R&D-Based Models of Economic Growth, *JPE* 103(4), 1995. The nonrivalry machinery that any elasticity above one requires.
80. E. Brynjolfsson, D. Rock, and C. Syverson. The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. *American Economic Journal: Macroeconomics* 13(1), 2021.
81. D. W. Jorgenson, M. S. Ho, and K. J. Stiroh. *Information Technology and the American Growth Resurgence*. MIT Press, 2005. Output elasticity of IT capital estimated as its factor income share.
82. G. S. Olley and A. Pakes, *Econometrica* 64(6), 1996; J. Levinsohn and A. Petrin, *Review of Economic Studies* 70(2), 2003; D. A. Ackerberg, K. Caves, and G. Frazer, *Econometrica* 83(6), 2015; A. Gandhi, S. Navarro, and D. Rivers, *JPE* 128(8), 2020. The production-function identification literature that the panel of Section 2.6 must satisfy.
83. P. Goldsmith-Pinkham, I. Sorkin, and H. Swift. Bartik instruments: what, when, why and how. *American Economic Review* 110(8), 2020. The exogenous-shares branch of the shift-share literature, and the branch Section 2.6's design falls back on once time effects are included.
84. K. Borusyak, P. Hull, and X. Jaravel. Quasi-experimental shift-share research designs. *Review of Economic Studies* 89(1), 2022. The quasi-random-shocks branch, which requires many shocks; Section 2.6 has four categories.
85. R. Adão, M. Kolesár, and E. Morales. Shift-share designs: theory and inference. *Quarterly Journal of Economics* 134(4), 2019. Inference under few effective shocks, and the reason Section 2.6 states what its design cannot detect.

A Reproducibility

The model is published as a repository, not only as scripts. The constraint parameters live in `params/constraints.json` and each verdict in `params/verdicts/`, so a reader who disputes an expansion rate changes one number and sees which verdicts flip. `python -m vn.cli` runs every verdict; `tests/test_backtest.py` asserts the backtest of Section 15 against outcome and will fail if a parameter change silently turns the Tier 2 miss into a hit.

Running the engine against the paper reveals a divergence of more than 0.10 on four verdicts, and the divergence is published rather than tuned away: the stated probabilities carry judgment the engine does not model, chiefly the elasticity dependence of Section 2 and the base-rate bounding of uncomputable cells. Reconciling the two so that every probability is either an engine output or a labeled adjustment to one is open work.

The simulation suite supporting Sections 14, 15 and the parameter bands throughout is published as `sim.py` with a fixed seed. It comprises a Monte Carlo over the parameter ranges in Table 4; a bounding experiment on `nc` treatment; an enumeration-robustness sweep; the backtest of Section 15; and the withdrawn Residual Ratio weight envelope, retained so the withdrawal is auditable. `sim.py` also emits the per-verdict posterior shifts that assign each of the thirty

verdicts to a category in Table 5; twelve of those assignments, seven Computed and five Decorative, are recorded in the repository rather than named in this document.

Table 9 of Section 2 needs no code to reproduce. Section 2.2 publishes the engine in closed form, $\sigma_i(\alpha) = g/(R_i/O_0)^{1/\alpha}$ with $g = 25.96$, together with every input and the implied required output for all six forecasts, so any cell in the table can be checked with a calculator and the thresholds follow by setting $\sigma = 1$. `elasticity.py` runs the same arithmetic across the full range and emits Table 9, Table 10 and the sensitivity table of Section 2.5. **A reader who disputes the conclusion must state their own α in order to do so**, which is the point of publishing the engine rather than the result. The intelligence price series, the withdrawn Residual Ratio inputs and the Capture Ratio inputs are all published as machine-readable files with a provenance flag on every point, so that every withdrawal in this document can be checked against the inputs that forced it. Readers are invited to contest any of it, and errors reported will be logged publicly whether or not they change a verdict.

B Amendment Log

Version 4.1 of October 2, 2026, against Version 4.0 of the same date, and the earlier rows against Versions 3.1 and 3.0. One row per correction, in the format of *Placement Before Profiling*, Appendix F.

Table 45: Amendment log. Rows dated 2026-10-01 are the Version 3.0 to 3.1 corrections; rows dated 2026-10-02 carry the 3.1 to 3.2, 3.2 to 3.3, 3.3 to 3.4, 3.4 to 4.0 and 4.0 to 4.1 corrections.

Date	Section	Change	Reason
2026-10-01	2.2; Table 9; Table 10; Headline 1; S19	Attempted to rebuild attributable output O from AI services revenue alone, excluding hardware vendors. Not constructible from disclosed data. Old series retained, labeled EST, with a stated caveat that the numerator includes hardware revenue and that α above 1 may be a price artifact. S19 now specifies a hardware-excluded numerator. Q1, Q2, Q4 and Q10 carry the caveat	O was built largely from NVIDIA data-center segment revenue, which is spending on capacity rather than output from it. No hyperscaler files an AI revenue line, the two largest frontier labs had only confidential draft registration statements as of this date, no annual 2023 figure exists for either, and calendar 2026 is incomplete
2026-10-01	End of 1.8; pointer in Section 13	Replaced the single falsification condition with two: a closing of the fixed-capability against frontier price divergence across two consecutive annual readings, and new entry above 3 GW into a named Tier 3 complement below 90 percent of its cap, tracked as S20	The published condition compared the complement basket against the intelligence index, and Section 1.5 establishes that the basket cannot be constructed. The law was unfalsifiable as stated
2026-10-01	1.8 opening; 1.8 closing; Q15; Q25	Removed all four remaining uses of the withdrawn ratio R and substituted the fixed-capability against frontier price divergence of 5.2x to 56x	R was withdrawn in Section 1.4. Continuing to reason from it was inconsistent with its own withdrawal
2026-10-01	Abstract; Contents; Section 15; Figure 21; Table 42; Appendix A	Replaced "pre-registered" with "backtest, parameters fixed at historical as-of dates, constructed in 2026" wherever it describes B1 to B5. Retained "pre-registered" only for the Call, the signature bet and the signposts	B1 to B5 were constructed in 2026, after their outcomes existed. Describing them as pre-registered overstated the instrument
2026-10-01	Abstract; Contents; Headline 8; Section 15; Table 42	Brier now reported by tier everywhere, with the aggregate only beside it: Tier 2 = 1.000 ($n = 1$), Tiers 3 and 5 = 0.035 ($n = 4$), aggregate 0.228	Section 15 required reporting by tier and never in aggregate, while the abstract, contents and headline results carried the aggregate alone
2026-10-01	Abstract; Contents; Headline 8b; 1.3; Table 5; Figure 5; Figure 19	Defined four categories once (computed, decorative, judged, empirical), made them a single partition of all thirty verdicts, and reconciled every count, caption and headline to it: 14 computed, 8 decorative, 5 judged, 3 empirical. The falsifier that computed verdicts should resolve with lower variance is retained	The counts disagreed. The text said the engine decides eight of thirty; Table 5 implied 22 engine-run; Figure 5 showed 8 / 14 / 8. Computed and decorative had been reported as overlapping labels, which let one verdict appear in two totals

Date	Section	Change	Reason
2026-10-01	Headline 7; 4.1; Figure 19 caption	Q1 survivor probability set at 0.53 throughout. 0.62 withdrawn as arithmetically impossible under the rubric: with $c = 0.45$ and $n = 5$ it requires $P_{\text{engine}} = 1.13$. 0.55 withdrawn as unsupported	Three values were in circulation: 0.62, 0.53 and 0.55. Only 0.53 reproduces under the document's own rubric: it implies $P_{\text{engine}} = 0.93$, where 0.62 implies 1.13
2026-10-01	Objection 6	Q14 survivor history stated as 0.60 to 0.52 to 0.41, matching Section 6.1	Objection 6 described a cut from 0.75 to 0.60, which the history in Section 6.1 does not support
2026-10-01	1.8; Table 7; 2.2; R1	Distinguished the two quantities explicitly: approximately $110B is attributable output, the numerator of the elasticity in Section 2.2; approximately $77B is captured revenue, the numerator of K in Table 7	The two figures were both presented as 2026 attributable revenue in different sections without distinction
2026-10-01	Headline 2; 1.8	Uncaptured surplus corrected to the table: it grew from approximately $112B to approximately $172B, an increase of about $60B	Headline 2 said about $25B and Section 1.8 said $70B to $95B. Neither matched the table
2026-10-01	Q15	Frontier price decline corrected to 1.34 to 2.25 times per year	Q15 stated approximately 2.9 times per year against the measured band of Section 1.4
2026-10-01	Table 4, Data row	Deleted "80% interval 2026 to 2032"	Reference [5] states that the interval is not in the source and has been withdrawn. The table contradicted its own reference
2026-10-01	1.2, Energy paragraph	Changed "two to six orders of magnitude of headroom" to "about two orders of magnitude"	Reference [2] supports about two orders and records that the six-orders claim was withdrawn
2026-10-01	Tables 1 to 9	Re-matched every caption to its content. Captions had been rotated by one position from Table 1 through Table 10; new captions written for the confidence rubric (Table 1) and the two price series (Table 6), and two orphaned captions retired. Tables 10 to 41 checked and found correctly matched	Each caption sat on the content of the table before it: the "three inferential layers" caption sat above the confidence rubric, the "constraint set" caption above the inferential layers, the "Capture Ratio" caption above the price series, and so on through Table 10
2026-10-01	Throughout	Corrected every section and table cross-reference: "1.2a" to 1.2 or 1.3 according to referent; "1.3a to 1.3c" to 1.5 and 1.6; "1.3b" to 1.6; confidence coefficients and `nc` convention to Section 1; consumer surplus to 1.8; the two-prices pointer to 1.4; the law's falsifier to 1.8; "Table 1" to Table 4 where it means the parameter table and Table 2 where it means the inferential layers; "Table A" to Table 9; "revised, Sec. 14" to Section 15	Several references pointed to sections that do not exist, and others to the wrong target
2026-10-01	Contents; Figure 22; Section 3 opening; Objections 2, 7, 8; Interim Residual	Signpost count reconciled to twenty, objection count to eight, and the question count to thirty, with the hundred-question battery identified as the earlier version of the program where the number is historical	"Nineteen tracked signposts" against a register of S1 to S20; "six objections" against eight; "the hundred questions" and "a hundred applications" against thirty carried here
2026-10-01	Figure 14 caption	Caption rewritten with dollar amounts escaped: $77T of new output against a $16T ceiling. All other figure captions checked for the same failure; none found	Dollar signs had been parsed as math delimiters, collapsing the caption to an unreadable run of characters

Date	Section	Change	Reason
2026-10-01	Section 14; Headline 3; Q19; R5; Figures 13 and 20	Re-checked the METR 80 percent horizon basis and held it at 27 to 32 minutes, labeled SPEC. Threshold, probability and adjudication rules unchanged. Reference [9] records the re-check date	METR has published no revision since May 8, 2026, the state of the source Version 3.0 relied on, and the newest model named there is not generally available under adjudication rule 3. One gap stated: the per-model values live in machine-readable files that were not reachable in this pass
2026-10-01	After the abstract	Added a conflict disclosure naming Galadora Technologies, General Diffusion, Reckon Research and Lodebearing, and stating that the complement-migration law and the signature bet point in the direction of those interests	The memorandum argued a thesis in which the author holds direct financial interests without saying so
2026-10-01	Throughout, including captions and references	Converted British spellings to American: data-centre, behaviour, rigour, labelled, relabelled, favour, programme, defence, grey. Running-text dates written as "July 23, 2026"	House standard, applied inconsistently in Version 3.0
2026-10-01	Table 4; R4; Q8; Q17; R5	Collapsed duplicated citation markers ("[58] [58]" twice, "[53] [53] [53]" once, "[53] [53]" once) and corrected "horizon [9]s" to "horizons [9]"	Typesetting artifacts
2026-10-01	Q14; Section 18; Table 4; reference [55]	Labeled the ">74% of new web pages AI-generated" figure EST and named its nearest traceable source and its limits. Labeled the MacroPolo elite-researcher figures EST and recorded the tracker's discontinuation in reference [55]. Labeled the 29.6 GW installed-capacity figure EST and recorded the 20 to 31 GW spread across published sources	Neither figure traces to a stable primary. Published values for installed AI data-center capacity in Q4 2025 span roughly 1.6x and none is disclosed
2026-10-01	Section 18	Forecast register resolved count corrected from six to seven	Five backtest results plus two entries resolved by revision is seven, which the same paragraph states
2026-10-01	Headline Results heading	"Eleven claims" corrected to "Twelve claims"	The list carries twelve items, numbered 1 to 11 with an 8b
2026-10-01	1.3	Empirical redefined: the category now means the engine's output does not settle the verdict, not that no engine ran. Q13's 96.5 percent engine output and its reason for being set aside are stated in the definition	The definition read "no engine," which contradicted Section 5.3, where the engine runs on Q13 and returns 96.5 percent survival
2026-10-01	Table 5 note	The division of the remaining eight verdicts into five judged and three empirical is now labeled a reconstruction rather than a model output	Three empirical verdicts are named in the text and the remaining five judged were obtained by subtraction. Only the 14 / 8 split is the model's own
2026-10-01	4.1	Q1 candidate (ii) marked as not stated rather than carried at 0.09	Version 3.0 stated no probability for candidate (ii). Version 3.1 supplied 0.09, which is forced by subtraction from the other four and was not written by the author. A figure that exists only to close a distribution is not a forecast, and the distribution is not summed in the first place
2026-10-01	1.3; Table 5 note	Engine reach corrected from twenty-two to twenty-three questions and no-engine from eight to seven; unnamed category memberships corrected from ten to twelve	Both counts were broken by this version's own redefinition of Empirical, which records that the engine ran on Q13. Seven Computed plus five Decorative is twelve, not ten

Date	Section	Change	Reason
2026-10-01	1.8	Disclosed that captured revenue is attributable output times exactly 0.70 in both years, that the factor's basis is unstated, and that K would read 0.39 against the other series	The two quantities an earlier amendment described as distinct are one series and not two
2026-10-01	Q14	Live tail for candidate (ii) set at 0.28 throughout	Version 3.0 gave 0.20 in the prose and 0.28 in the distribution one page later
2026-10-01	Section 17	"Five of the original thirteen" corrected to five of the twenty, with the register's composition by kind stated	The signpost count survived uncorrected in the body text
2026-10-01	Table 9 caption	Leftover "Table A." label removed from the caption	Residue of the Table A to Table 9 renumbering
2026-10-01	4.1; Headline 7; Contents; Q1 bottom line; Figure 14	Majority-of-GDP gap corrected to 4.8x and 0.68 orders of magnitude	\$77T against the stated \$16T ceiling is 4.81. The withdrawn ninefold is the ratio to this memorandum's own band midpoint of \$8.75T, not to the ceiling the sentence names
2026-10-01	2.3; Abstract; Headline 1; Figure 10 caption	Goldman recorded as failing marginally at $\alpha = 1.4$ and clearing at about 1.41; "reverses every forecast" corrected to four of the six; Acemoglu restated as a bet on a floor of about 0.85	Table 9 prints 0.98 for Goldman against a survival rule of $\sigma \geq 1$. McKinsey and Acemoglu clear throughout the observed range. Every cell rises monotonically in α
2026-10-01	5.2; Headline 5; Figure 15	Chokepoint time requirement corrected to 7.4 years at the midpoint, 5.6 to 10.8 across the band	The published 5.2 lies below the fastest rate in its own stated range of $r_{3c} = 0.10$ to 0.20
2026-10-01	Headline 10; 5.2; Figure 16 and its caption	Migration ratio corrected from an order of magnitude to five to six fold, in the text, the caption and the regenerated artwork	Packaging at 0.60–1.10 against grid equipment at 0.10–0.20 gives 6.0 and 5.5 at the two ends
2026-10-01	Section 1	Rubric scope stated: it governs the surviving candidate only, tails are engine posteriors, the twenty-five verdicts with no coefficient are identified, and the two scales are labeled wherever both appear	The mixing imposes a floor of $(1 - c)/n$, and several published tails sit below it. Version 3.0 presented the rubric as universal and applied it to five verdicts
2026-10-01	Section 1; Q2; Table 42	Non-Tier-0 zeros rewritten as below 0.05 percent; backtest confidences identified as stated forecasts rather than survival likelihoods	Only Tier 0 bars, which is the framework's central structural claim, and Version 3.0 published 0.0 percent at Tier 1 and Tier 5
2026-10-01	Table 6	Annualized column identified as the repository's fitted rate rather than an endpoint ratio; divergence lower bound corrected from 4 to 5.2	Compounding Total over Window gives 2.59x, 1.23x and 10.13x against the published rates, and 11.8 over 2.25 is 5.24
2026-10-01	1.5; 1.6; Objection 4; Section 15	Four argument-level concessions added: the three-attempts inference reconciled to one root; the supply-response test restated as a check rather than a validation; the no-recovery assumption behind the compounding bound stated; two of five backtests identified as near-definitional	Each was a load-bearing step asserted rather than argued, and the first was stated in two mutually exclusive forms five paragraphs apart
2026-10-01	Table 43, S8, S10, S14, S16, S19, S20	S16 recorded as FIRED on this document's own r_{data} revision; S19 marked dormant; S20 quantified at 3 GW and 90 percent of cap; S8 and S10 annotated with their current state	S16's trigger is any r_c moving by a factor above three, and Section 15 moves one by twelve to fifteen. S19 requires a series Section 2.2 shows cannot be built. S20's condition excluded nothing, since clearing prices cannot exceed the cap

Date	Section	Change	Reason
2026-10-01	Figures 1, 2, 5–22	All twenty regenerated from corrected values, with their captions rewritten to match. Figures 1, 11 and 12 additionally had quantities removed that appeared nowhere in the body: a constraint rate of 0.22 for a "delivered capacity" that is not a row of the constraint set, three capital-gap ratios, and a 99 percent packaging concentration. Figures 6 and 17, and 13 and 20, were byte-identical pairs and are now distinct charts	Version 3.0's artwork carried superseded probabilities, contradicted its own captions, labeled an arrow 9x between bars in a ratio of 4.8, retained the word pre-registered and the aggregate-only Brier, printed a withdrawn 5.2-year requirement and an order-of-magnitude migration ratio, and illustrated four claims with two images. A first correction pass reached seven figures and left six others still printing withdrawn values; all twenty are now current
2026-10-01	1.2; 1.8; Q17; Q2; Section 14; Section 18; Objection 2; Table 42; Contents	Smaller corrections: the ten parameterized rates against six constraints acknowledged; "captured by nobody" corrected to "by no producer"; the 23 percent rescoped to vision tasks in Q17; Q2's scenario reconciled with its own elimination; adjudication rule 4's direction of bias corrected and disclosed; the capital-revenue gap's revenue side identified as estimated rather than filed; the enumeration example flagged as not satisfying its own rule; B5's Brier corrected to 0.000; the contents' section list and objection-concession count corrected	Each was found by the post-correction audit
2026-10-01	New Appendix C	Post-correction adversarial review recorded, with the disposition of every finding	A reader is entitled to know what an adversarial pass found, not only that one was run
2026-10-01	Figure 1; Figure 7; Figure 12; Appendix B	Figure 1's horizon corrected from 4.4 to 1.4 years; Figure 7's per-year caps drawn at each year's own clearing price instead of one flat line; Figure 12's in-image note corrected to four rows of which three carry a share; the open-items register corrected from two items to three	Figure 1 paired the 2026-delivery shortfall with the 2030 horizon, which made Matter 3a appear to clear a deadline Section 5.2 says it does not clear. Figure 7's single flat cap invented a cap level that appears nowhere in the body and drew two of three capped bars as clearing below it. The register's claim that neither open item was a defect in what is published contradicted Section 1.8, which records the Capture Ratio's undisclosed 0.70 factor as exactly that
2026-10-01	Figure 1; Figure 20	Figure 20's note corrected to under two years; Figure 1's constraint row renamed to match Table 4	Figure 20 asserted two different durations for the same interval, one inside the plot and one in the note beneath it. Figure 1 named Tier 3d differently from Table 4
2026-10-01	Figure 19 caption; Appendix B; Appendix C	Figure 19's caption scoped to the Computed verdicts, which is the scope Table 5 gives the phrase; the thirteenth pass's account of what the twelfth had missed corrected in both appendices; the register closed at seventeen passes	The caption said the seven marked verdicts were those whose arithmetic is published, where other verdicts publish arithmetic too; Table 5 scopes the phrase to Computed and the caption had dropped the scope. The account of the twelfth pass's miss was false against this appendix's own list three lines above it: one of the two figures was named there. A correction to a claim in this appendix creates a further claim, so the sequence is closed here by decision

Date	Section	Change	Reason
2026-10-01	Headline 9; 1.5; Figure 19 legend	The phrase *three unrelated reasons* corrected to three different proximate reasons sharing one root, in Headline 9 and at the opening of Section 1.5; Figure 19's legend scoped to Computed, matching its caption	Section 1.5 asserted that the three failures were unrelated and then, in the next sentence, gave the rule that applies when they share one root, and named that root twice more in the same section. Headline 9 carried the same contradiction inside a single sentence
2026-10-01	Figures 1, 2, 5 to 22	Regenerated from corrected values; rendering history in the repository changelog	The rendering sequence, which concerned halos, panels, label positions, contrast and pass counts rather than any published quantity, is recorded verbatim in `CHANGELOG-figures.md` and is not reproduced here
2026-10-01	Title block; Appendix B	Version set to 3.1 with today's date; this amendment log added	Release of the corrected version
2026-10-02	Headline 2; 1.8; Table 7; Figure 8; Appendix B	**The 0.70 factor is derived and the Capture Ratio is restored.** The scaling is a population correction: the consumer-surplus denominator measures US consumers, while attributable AI revenue is reported globally, so the numerator is scaled to the US share of approximately 70 percent to place both sides on one population. Section 1.8 states the derivation, Table 7 publishes K at 0.27 and 0.31 on that basis with a robustness band, and Headline 2 carries the seventy-percent finding with the floor argument. The unscaled alternative is dropped	A global numerator over a US denominator is not a ratio of anything, so the correction is required rather than optional, and an exact constant in both years is what a population adjustment looks like. The unscaled computation published in the first draft of this version reintroduced the mismatch it was meant to expose. The US share is the one free input and is labeled EST; at any value between 60 and 80 percent K stays inside 0.24 to 0.34
2026-10-02	4.1; Q2; Q12; Q14; Q23; Headline 7	The circular defense of Q1's survivor probability removed. 0.53 is now defended on the rubric alone: $c = 0.45$, $n = 5$, floor $(1 - c)/n = 0.11$, so 0.53 implies $P_{\text{engine}} = 0.93$ where 0.62 implies 1.13. Candidate (ii) is marked not stated rather than carried at 0.09. The survivor and the tails are labeled **Probability (rubric)** and **Computed (engine)** at Q1, Q2, Q12, Q14 and Q23. Headline 7 no longer derives the rubric survivor from an engine tail	0.53 was justified because the candidate distribution sums to 1.00, but candidate (ii) was obtained by subtraction, so the distribution sums to 1.00 by construction and carries no information. Section 1 also states that rubric-scale and engine-scale probabilities are separate scales that are not summed, which the argument violated. The same argument survived in one amendment-log row and is removed there too
2026-10-02	Disclosure box	Every sentence after the statement of interests deleted, including "Read that as a disclosure and as a credential" and "The judgment stands on the author's hands being on the machine"	A disclosure states an interest so a reader can discount for it. Arguing that the interest is a qualification asks the reader to do the opposite, inside the paragraph meant to protect them
2026-10-02	Appendix C; Appendix B; repository	Appendix C cut to one page, keeping the seven substantive classes and their dispositions and one sentence on the figures. The account of artwork passes, halos, glyph boundaries, lozenges, legend frames, contrast ratios and pass counts deleted, with the figure-rendering rows of Appendix B collapsed into one. The removed material is reproduced verbatim in `CHANGELOG-figures.md`. Appendix C's first sentence now states that the review was run **by an AI model**, and that this does not satisfy Section 18's requirement for documented adversarial review by people with domain expertise	The production history of the artwork is not a result and was occupying more of the document than any verdict. Naming the reviewer matters more: a machine check on arithmetic and rendering is not peer review, and an appendix that did not say so invited the reader to count it as one

Date	Section	Change	Reason
2026-10-02	Abstract; Contents item 2; Headline 1; 2.5; 18	The elasticity result reframed around the measurement that does not exist. Headline 1 now leads with the missing dataset and carries 1.70, 1.17 and 1.36 as supporting evidence on a series that counts hardware revenue as output. Section 2.6 gains a specification of the dataset that would settle it: AI services output against deployed compute, by firm or sector, with hardware vendors excluded. Table 9, Table 10 and every value are unchanged	The paper's most consequential result is that the number cannot currently be measured. Leading with three readings of a series the same section calls unfit presents the symptom as the finding
2026-10-02	Title block; Appendix B; Appendix C	Version set to 3.2; this block of rows added; "Carried forward unresolved" rewritten to list only what remains open after this release	Release of the corrected version
2026-10-02	Contents item 2; Headline 1; Table 9; Table 11; Figures 9 and 10	Contents item 2 rewritten to the adopted band, where it had carried the superseded 1.17 to 1.70 reading as the operative result. Two cells of Table 9 set to the values the published engine returns: 0.41 for PwC at $\alpha = 1.2$ and 3.69 for Acemoglu at 1.4. The sensitivity caption and footnote scoped to the efficiency-consistent rows, where they had claimed every row against their own final row. The halving claim corrected: at 100 GW five of six fail and Acemoglu survives. Figures 9 and 10 reordered so each file renders at its own number, and the threshold plot given its own caption	Publishing the engine in closed form makes every cell checkable, which means every cell must match it; the two that did not were carried from a prior rounding. A claim scoped to every row that its own table contradicts is exactly the defect this document exists to catch
2026-10-02	Title block; Appendix B	Version set to 3.3 with today's date	Release of the corrected version
2026-10-02	2.1–2.6; Table 9; Table 10; new sensitivity table; Figures 9 and 10; Headline 1; Abstract; 4.1	**The compute convention is made consistent across the engine and the measurement.** The forward engine credits permitted capacity with four years of efficiency gains; the historical elasticity was measured against raw physical capacity. Both now use efficiency-adjusted compute, and the elasticity reads **0.77 to 1.26** where Version 3.1 published 1.17 to 1.70. Table 10 publishes both columns. The engine is published in closed form, $\sigma = g/(R/O_0)^{1/\alpha}$ with $g = 25.96$, with every implied R_i and a threshold column added to Table 9. A sensitivity table covers the capacity anchor, the installed-base treatment and the convention	The same symbol carried different units on the two sides of the section: α was estimated against gigawatts and applied to effective compute. The correction also removes a theoretical impossibility, since an output elasticity of 1.4 on a single input implies that input is paid more than the output it produces, while 0.77 to 1.26 straddles constant returns. **Four of six published forecasts now fail at every point in the band rather than reversing inside it, and three of the four fail under every efficiency-consistent modeling choice tested**
2026-10-02	4.1; 2.5; S19	This memorandum's own \$7 to \$10.5T band is among the four that fail, and Section 4.1 states the exposure rather than carrying the verdict quietly. Section 2.6 publishes a falsification: an estimated α below 1.27 on the specified panel eliminates the band	The band requires 1.27 at its lower bound against a measured range topping out at 1.26. It is carried rather than withdrawn because the measurement rests on three arc observations on a numerator this document calls an accounting identity, and the confidence coefficient already prices the exposure at the largest single deduction in the rubric

Date	Section	Change	Reason
2026-10-05	Title block; citation line; extract title block; extract provenance line	**The byline, publisher line and contact address are changed to Mind Meets Matter** and `sidney@mindmeetsmatter.com` in both documents. The Ashby Institute no longer appears in either	The institute was the issuing body for Versions 3.0 to 4.0. Those versions stand as issued; this one and everything after it are issued under Mind Meets Matter, and a publisher line that disagrees with the byline is the kind of inconsistency this register exists to catch
2026-10-03	2.6; 3.4; 4.1; Figure 11	**Six values that the Version 4.1 anchor change did not reach.** The power calculation in Section 2.6 still printed $\mathrm{SE} \approx 0.20$ and a minimum detectable difference of 0.55, which came from multiplying by the design effect where the stated formula takes its square root; the corrected values are 0.14 and 0.39, and the extract already carried them. **Both falsification statements still named 1.27**, which was this memorandum's lower-bound threshold on the superseded anchor; on the services anchor it is 1.20. The capital gap still read ∼18x against Bain's $2T requirement, which was the ratio against $110B; against $140.4B it is 14.2x, and Figure 11 was still rendering 6.6x, 18x and a $110B bar	**An anchor change propagates to every ratio taken against it**, and these were the ones that were not reached because they sit outside Section 2: a power calculation, two falsification thresholds in different sections, and a figure. The separation claims survive the correction, since 0.8 against 1.4 is still detectable at 0.39 and 1.00 against 1.20 is still not
2026-10-02	Abstract; Contents 2; Headline 1; 2.4; 2.5; Figures 9 and 10; extract abstract and Section 1	**"Eliminates nothing" is withdrawn as a finding and replaced with the disagreement between the two series.** The hardware-inclusive series fails four of six and is an accounting identity; the services series fails none and divides lab revenue by all AI data-center capacity, so a rising lab share raises the measured elasticity with no change in production. **Neither adjudicates, and the thresholds, 0.76 to 1.45, are the durable result.** The rising readings are attributed to the reduced form of this memorandum's own engine rather than to the six forecasts, which state no α	Version 4.0 replaced one unusable verdict with another and reported the second as a result. **A denominator covering capacity whose output is not in the numerator is a mix measure**, and a numerator that is hardware revenue is a price index; the thresholds are the only object in the section that neither defect touches
2026-10-02	2.4	**A reductio is added.** At $\alpha = 1.99$ the engine permits about $92T of 2030 AI-attributable output on $O_0 = \$140.4$B and $g = 25.96$, and the majority-of-GDP claim requires 1.94, inside the observed range. **A series that cannot eliminate a claim Section 4.1 rejects by 4.8 times is measuring diffusion and mix, not a production elasticity**, and the rise from a $1.7B base is the signature of Tier 1 absorption	The services reading is not merely uncertain, it is of the wrong kind, and the cleanest demonstration is that it fails to exclude an outcome the memorandum excludes elsewhere on independent grounds
2026-10-02	2.5; Table 12; extract Table 3	**A bias ledger is published.** Four uncorrected biases on the services series with their direction for the forecasts: output deflation favors, the 2026 partial-year run rate favors, a GDP contribution exceeding producer revenue favors and makes every threshold an upper bound, and the rising lab share of capacity runs against and is unmeasured	Three of four run one way and the fourth is the one the dataset of Section 2.6 exists to measure. Printing the directions is what licenses the refusal to report either verdict

Date	Section	Change	Reason
2026-10-02	Headline 2; 1.8; Figure 8; Tables 9 and 10; thresholds; Figures 9 and 10	**Stale Version 3.4 values recomputed on** $O_0 = \$140.4$**B** and every hardware reference removed from the Capture Ratio. Table 9's cells and all six thresholds are recomputed; the "six floors span" and "adopted band spans" sentences, the Figure 9 and Figure 10 captions and the incremental-reading sentence are rewritten; Figures 9 and 10 regenerated. $K = 0.36$, so the uncaptured share is about 64 percent, **stated as a point estimate and not a bound because the two biases now run in opposite directions**: the services series excludes hyperscaler AI revenue and understates capture, the surplus excludes enterprise surplus and overstates it. "At most thirty cents" becomes "about thirty-six cents"	A table recomputed on a new anchor and a headline left on the old one is the defect this register exists to catch. The floor language survived from a construction where both biases pushed the same way, and the services series removed that property
2026-10-02	Abstract; Headline 1; 2.4	**Attribution corrected.** "The method is the thing that found it" and "the method eliminated its own headline" are replaced with "a check of the anchor against NVIDIA's fiscal 2025 10-K found it"	The method did not find the anchor error. An external filing did, and crediting the framework for a catch it did not make is the kind of claim this memorandum is written to refuse
2026-10-02	1.8; 2.1–2.6; Tables 8, 10 and 11; Figures 9, 10 and 11; Headline 1; Headline 2; Contents item 2; Abstract; 4.1	**The output anchor is rebuilt from reported figures and the central result of Section 2 reverses.** The \$110B anchor of Versions 3.1 to 3.4 was carried forward with a citation to a debt-issuance source that could not support it, and it was below the reported annualized run rate of either OpenAI or Anthropic on its own. **Table 8 now publishes an AI services revenue series provider by provider**, 1.7, 7.0, 29.7 and 140.4 \$B for 2023 to 2026, compiled by Epoch AI [3, 5] from company disclosure and reporting, excluding hardware entirely and excluding hyperscaler AI revenue because an unknown share of it is lab revenue already counted. **On that series the measured elasticity reads 1.11, 1.33 and 1.99, rising monotonically**, and **no forecast is eliminated, on any of the ten conventions Table 11 tests**. Headline 1, Contents item 2, the abstract, Figures 9, 10 and 11 and the Q1 exposure paragraph of Section 4.1 are rewritten to the reversed result. Headline 2 moves from a seventy percent floor to a two thirds floor, and the hardware double count in the Capture Ratio is removed, since the series no longer contains hardware	**An anchor that enters every threshold cannot rest on a citation that reports a different quantity**, and this one was checkable from outside: two providers' public run rates exceeded it. The identity argument of Version 3.4, which held the series was hardware revenue and therefore an accounting identity, does not apply to a services numerator and is withdrawn; the price channel it pointed at survives in the deflated columns of Table 10, where it is measurable rather than fatal. **The method eliminated its own headline**, which is the strongest test it has been put to, and the four versions that published the elimination are left in the record rather than rewritten

Date	Section	Change	Reason
2026-10-02	2.4; 2.5; Table 11; Headline 1	**The invariance is published as the result in place of the eliminations.** Table 11 now varies the efficiency factor from 1.3 to 1.5, permitted capacity from 150 to 250 GW, the base from 25 to 35 GW, both readings of the Bain figure and both output deflators, each applied to both sides, and records that no forecast fails in any of them. The reason is stated: every threshold is $\ln(R_i/O_0)/\ln g$ with a common g and O_0, so no choice of either can reorder them, and a convention applied to both sides moves the permitted multiple and the measured band through the same channel	A sensitivity table that varies inputs one at a time and reports levels invites the objection that banding every unsourced input would dissolve the result. It would dissolve the levels. It does not touch the comparison, and the comparison is what the section claims. **Reporting the invariance is both stronger and more honest than reporting a point verdict**
2026-10-02	Title block; Appendix B; Figures 8, 9 and 10; table cross-references	**Version set to 3.4**, this block of rows added, and the corrections propagated out of the prose. **Three figures were regenerated**: Figure 8 dropped the "capture rising" annotation and its panel title, Figure 9 dropped "Acemoglu clears throughout" and the 1.70 band label, and Figure 10 dropped the "narrower than the uncertainty" claim and the constant-returns argument. **Fifty-seven table cross-references were shifted by one**, since inserting the revenue-construction table as Table 8 renumbered every table after it	A memorandum whose discipline is logging what it changed cannot leave the withdrawn claims printed inside its own artwork, where a reader who looks at the pictures reads the superseded version. The cross-reference drift was introduced by this release and caught on audit
2026-10-02	2.3–2.6; Table 10; Table 11; Figures 9 and 10; Headline 1; Abstract; Contents item 2	**Symmetric deflation.** The adopted convention deflates compute by 40 percent a year and leaves output nominal, which Section 2.6 itself says biases the estimate down. Table 10 gains two columns deflating output by the frontier price series of Section 1.4, at 1.34x and 2.25x per year, giving 1.49, 1.08, 1.15 (band 1.08 to 1.49) and 1.89, 1.56, 1.81 (band 1.56 to 1.89). Table 11 gains the matching rows: **PwC alone fails throughout at 1.34x and none fails at 2.25x**	The asymmetry was an unstated assumption that AI output prices were flat across 2023 to 2026, which Section 1.4 of this same memorandum shows they were not. **The elimination result depends on the output deflator and no deflator choice is defended by the series**, which is now stated wherever the band is used
2026-10-02	2.4; 2.5; Headline 1; Abstract; Contents item 2; extract Sections 1 to 3	**Identity versus elimination.** Every statement that the observed band eliminates a forecast unconditionally is withdrawn. The primary result is now the threshold column α^*, 0.84 to 1.52, which is arithmetic on published claims and requires no observed elasticity. Band comparisons are stated conditionally: if the identity series were informative about production, four of six would fail on the adopted convention, and under a symmetric deflator one fails or none does. "An elimination, not an expression of uncertainty" and "eliminated now, not conditionally" are removed	Section 2.6 establishes that the available series is an accounting identity, $O \equiv p_C \cdot C$, so every deviation of the measured elasticity from one is a movement in the price of compute. **A quantity that is identically a price movement cannot eliminate a forecast.** Version 3.3 made the identity argument and then drew the stronger conclusion anyway
2026-10-02	2.1; 2.3; extract Section 3	**Theory bound.** The claims that the corrected band moves the result toward theory, and that Section 2.1 says a single-input output elasticity has to sit near constant returns, are deleted. Section 2.1's factor-share bound of 0.4 to 0.8 is kept and restated as applying to a multi-input output elasticity	Section 2.6 states that the single-input specification recovers scale rather than an output elasticity, so the reduced form of Table 10 and the factor-share bound are not the same object and cannot be checked against each other

Date	Section	Change	Reason
2026-10-02	1.8; 2.2; note [28]; extract citation note	**The output anchor.** The 2023 to 2026 attributable-revenue series is published as a table, component by component with sources and labels, and all three descriptions of it are made identical. **The series includes accelerator revenue, which is now stated in both documents**, so Section 2.6's identity premise stands. Two defects are recorded rather than patched: NVIDIA fiscal 2025 data-center revenue alone was $115.2B per its 10-K, above the $110B 2026 series total, and citation [28] is a debt-issuance source that does not support a revenue series	Version 3.3 described the same series as NVIDIA data-center revenue in Section 1.8, as issuance data plus lab run-rates in the extract, and cited it to [28] in Section 2.2. An anchor that enters every figure in Tables 8, 9 and 10 cannot have three descriptions, and the $115.2B filing is a direct contradiction of the level
2026-10-02	2.4; 2.5; Table 9 note; extract Table 1 and Section 3	**The lower bound margin.** "Eliminated now" is replaced throughout with: at the edge of the band, 1.27 against 1.26, inside the stated precision, and surviving on the incremental reading of the 200 GW figure, where the threshold falls to 1.22	A gap of 0.01 against a bound with no computable standard error is not an elimination, and reporting it as one claimed a resolution a three-observation series does not have
2026-10-02	Headline 2; 1.8; Figure 8 caption	**Capture Ratio restated as a floor only.** Headline 2 now reads that at least roughly seventy percent of measured generative AI value is captured by no producer, and states why the bound holds in one direction: the revenue side includes hardware sales and the surplus side excludes enterprise surplus, both of which overstate capture. **"The direction is now known" is deleted**: K's rise reflects revenue, largely hardware, growing faster than surplus, and is not evidence about capture. "The single most important number for anyone underwriting AI returns" and "every informal version of this comparison in circulation makes exactly that error" are deleted. The 70 percent US share is kept EST with its 60 to 80 percent band and listed under needs author input, and the table note now states that hardware revenue in Captured double counts inside Created	A ratio whose numerator and denominator are both biased upward supports a floor and not a level, and a rise driven by a hardware-led numerator says nothing about the share of AI value a services producer can price against. The deleted claims asserted more than the construction carries

Date	Section	Change	Reason
2026-10-02	2.4; 2.5; Table 11; Figure 10; Headline 1; Abstract	**Second referee pass, four corrections, and the result comes out stronger.** (1) "Acemoglu clears throughout" was false: its threshold is 0.84 against a band that starts at 0.77, so neither survivor clears the whole band, and the sharper true statement is now published, that **at two of the three elasticities the series has actually produced all six forecasts fail**. (2) **The algorithmic-progress and efficiency-factor rows were judged against the adopted band while changing the forward multiple**, which is the mixed-convention error this section was written to catch, committed inside the table that catches it. Every row now carries its own band, computed by applying that row's convention to the historical series as well. **Corrected, crediting algorithmic progress rescues nothing**, because it raises g and lowers the measured elasticity through the same channel, and PwC, Goldman and this memorandum's upper bound fail on all nine comparable rows rather than on six. (3) The surplus-inclusive anchor is reclassified from a row that overturns the result to a channel that cannot be scored against a revenue-estimated band, and is ruled off the table with the reason stated. (4) "Every forecast's threshold travels further than the band is wide" was false; it is three of six	The first pass added a sensitivity table and the second found the table making the error it was added to detect. Applying a convention to the forward multiple but not to the historical series is the same unit inconsistency corrected at the section level in the previous entry, reappearing one level down. **The honest computation is the stronger one**: the three-forecast core now survives every convention tested, where the previous draft conceded two channels that overturned it. **Superseded by the symmetric-deflation row of Version 3.4**, which found that claim scoped to conventions that leave output nominal

Date	Section	Change	Reason
2026-10-02	2.2; 2.4; 2.5; 2.6; Table 10; Table 11; Figure 10; Headline 1; Abstract; References	**Nine corrections from an adversarial referee pass, four of them material.** (1) The 2023 to 2024 raw-capacity elasticity is 1.7051 and was printed 1.70; the raw band is 1.17 to 1.71 throughout. (2) The sensitivity table's Version 3.1 row contradicted itself, naming no forecast as failing throughout while naming only two survivors; at $g = 25.96$ on 1.17 to 1.71 all six survive, and the row is replaced by the only fully raw case, $g = 6.76$, on which five fail and Acemoglu survives. (3) The claim that withdrawing the efficiency credit eliminates all six held only when $g = 6.76$ was evaluated against the efficiency-consistent band, which is the mixed convention this section condemns. (4) **The output anchor was never varied although it enters every threshold**, and Section 2.6 asks the measurement to use a surplus-inclusive numerator the anchor excludes; a surplus-inclusive row is added and clears all six. (5) **Algorithmic progress was absent**; a row is added, and at 1.15x per year only PwC is eliminated. (6) Efficiency-factor rows at 1.3 and 1.5 added. (7) The 200 GW is Bain's increment and the engine reads it as a total; both readings are now stated and the incremental one restores this memorandum's lower bound. (8) "Narrower than the uncertainty in the parameter" was false, since the floors span 0.68 and the band spans 0.49. (9) PwC's figure is a 2030 annual contribution, not a cumulative total. Section 2.6 rebuilt: time effects and non-compute inputs added to the estimating equation, the two shift-share branches named and chosen between, what the design cannot detect stated, and references 83 to 85 added	A document that publishes its engine in closed form invites exactly the arithmetic check that found items 1 to 3, and a sensitivity table that omits the input every threshold depends on is not a sensitivity table. Items 4 and 5 overturn the result rather than perturbing it, which is why both are published as rows rather than argued away; the elimination is real on the stated conventions and conditional on them, and the distance between those two statements is now the finding
2026-10-02	Disclosure; 2.5	**The result that would damage the author's disclosed positions is now named**: an estimated elasticity at or above 1.4, which clears most of Table 9 and retires the constraint thesis. The falsification on this memorandum's own band is restated as already satisfied on present evidence rather than pending, and its threshold is shown moving from 1.27 to 1.22 across the capacity readings	A disclosed conflict is neutralized by a stated losing condition, not by disclosure alone, and a pre-registered trigger that the document's own published band already satisfies is a result, not a test

Date	Section	Change	Reason
2026-10-02	2.1; 2.5; References	The central claim is narrowed to what it can carry and the production-economics literature is cited: the elasticity of output with respect to *deployed physical compute capacity* is what no published AI forecast reports, where elasticities of output with respect to information capital have been estimated for decades. Six references added, 77 to 82. Section 2.1 adds the factor-share bound; Section 2.6 adds the estimating equation, the simultaneity problem and a shift-share instrument built from this memorandum's own constraint series	Version 3.2 asserted a universal negative over a literature it did not cite, and Epoch's GATE model, which states compute-output parameters explicitly, was already in this memorandum's own bibliography at [46]. The narrowed claim is defensible and more useful

Carried forward unresolved. Four items remain open after this release, and the output anchor is no longer among them. First, the 40 percent annual efficiency correction is cited to a price-performance series in FLOP per dollar while the engine requires energy efficiency in FLOP per watt; Table 11 varies it from 1.3 to 1.5 as its disclosure, and no verdict changes across that range. Second, the 200 GW is a demand projection used as a supply ceiling, and both readings of it are now rows of Table 11. Third, the 70 percent US-share coefficient behind the Capture Ratio is EST and unsourced. Fourth, the per-verdict posterior shifts that would name all thirty members of the Table 5 categories live in `sim.py`; twelve assignments are recorded in the repository rather than here, and the counts are unaffected. **The hyperscaler share of AI services revenue is a new open item**: Microsoft alone disclosed a $37B AI run rate, and separating it from lab revenue already counted is the attribution problem Section 2.6 specifies, which makes every level in Table 8 a lower bound.

The Capture Ratio, carried in Version 3.1 as a defect in what is published, is resolved. Its 0.70 scaling factor is a population correction placing a globally reported numerator on the US population its denominator measures, which Section 1.8 now states and derives. The ratio is published at 0.27 and 0.31 on that basis. **The US-share coefficient is the one input a reader should test**, it is labeled EST, and the finding is robust across every plausible value of it.

Two further items are open work rather than defects and are named where they arise: a clean demonstration of the enumeration bound (Section 1.1), and the reconciliation of engine outputs with rubric-adjusted probabilities (Appendix A). The thirty-way computation of the Table 5 categories is the first item above and is not counted twice. **Documented adversarial review by people with domain expertise remains outstanding**, as Section 18 records and as Appendix C states of itself.

C Post-Correction Adversarial Review

After the corrections of Appendix B were applied, this version was audited **by an AI model with no part in producing the corrections**, against the rendered document rather than its source. **This does not satisfy the requirement in Section 18 for documented adversarial review by people with domain expertise and no stake in the conclusion, which remains outstanding.** It is a machine check on arithmetic, internal consistency and rendered output, and should be read as that. It found defects in seven classes the original correction specification did not cover. Each class is recorded with its disposition, because a reader is entitled to know what the check found and not only that one was run.

The arithmetic, four corrections. The majority-of-GDP gap is **4.8x and 0.68 orders of magnitude**, not ninefold and 0.94 orders; the withdrawn figure was the ratio to this memorandum's own band midpoint rather than to the $16T ceiling the sentence names, and it appeared in four places. The chokepoint requirement is **7.4 years at the midpoint rate**, not 5.2, which lay below the fastest rate in its own band. The migration ratio is **five to six fold**, not an order of magnitude, which matters because this document calls that ratio the entire content of the migration argument. HBM content per accelerator rose **3.6-fold**, not 2.4.

The claims that overstated their own tables, three corrections. Goldman **fails marginally** at $\alpha = 1.4$, where Version 3.0 said it survived. The observed elasticity range reverses **four of the six** published forecasts, not every one. Acemoglu's estimate is a bet that α clears about 0.85, the lowest floor any of the six requires, not that α sits below 1.0.

The rubric, three corrections. Its scope is stated: it governs **the surviving candidate's probability and nothing else**, because the mixing imposes a floor of $(1 - c)/n$ and several published tails sit below that floor as engine outputs. The two scales are labeled at the five verdicts that state both. The twenty-five verdicts publishing no coefficient are identified as such.

The zeros, two corrections. A candidate at Tier 1 or Tier 5 is written as below 0.05 percent rather than as 0.0, because only Tier 0 bars and that distinction is the framework's central structural claim. The backtest's 0 and 100 percent are stated forecasts at historical as-of dates, not survival likelihoods.

The register, five corrections. **S16 is recorded as fired**, on this document's own revision of r_{data} by a factor of twelve to fifteen, the first and largest trigger on record and one Version 3.0 did not claim. S19 is marked dormant, because the numerator it requires does not exist. S20's condition is quantified at 3 GW and 90 percent of cap, where the prior wording excluded nothing. S8's first half is recorded as already met, and S14's constraint column is explained.

The arguments, four concessions added. Section 1.5's inference rule and its diagnosis are reconciled: three different *proximate* failures with one root. Section 1.6's supply-response test is restated as an empirical check on the tier ordering rather than a validation of it. Objection 4 states the sharper form of itself that Version 3.0 omitted. Section 15 records that two of its five backtests are close to definitional, leaving three informative questions and one miss.

The figures. All twenty regenerated figures render the body's values. The rendering history is recorded verbatim in `CHANGELOG-figures.md` in the repository, because it concerns the production of the artwork rather than any claim made here.