GLOBAL MACRO RESEARCH

The Toll Road Moves

How cheapening intelligence collides with national bottlenecks

Forecast Opinions

ID	Forecast	P(YES)	Resolve by	Resolution
FC-01	In 2030, BIS authorization will still be required for an ordinary US export to mainland China of an Nvidia H200 data center accelerator.	90%	2030	Test the rule for an H200 shipped by an unlisted US exporter to an unlisted civilian Chinese end user.
FC-02	The Stanford AI Index 2030 will show the US lead over China's highest rated public model at no more than 5.0%; a tie or Chinese lead counts YES.	80%	2030	Use Stanford's reported human preference comparison or its named successor.
FC-03	By 2031, the US, EU, UK, or China will have kept a quantitative cross-border model-weight control in force for at least 365 consecutive days.	42%	2031	The binding rule must use a numeric training compute or capability trigger.
FC-04	By 2029, the US will sign another conditional government-to-government advanced-AI access agreement beyond the Saudi and UAE arrangements.	77%	2029	The new agreement must condition advanced accelerators, or accelerators bundled with deployment infrastructure and services.
FC-08	The first qualifying IEA estimate published by 2031 will put global 2030 data-center electricity consumption at least 1,000 TWh.	68%	2031	Use an IEA estimate or explicit nowcast.
FC-11	By 2028, a US state utility commission will approve a qualifying large load cost allocation or take-or-pay tariff.	88%	2028	A general tariff must cover one customer or project of at least 100 MW and allocate at least 80% of dedicated cost or impose ten years of take-or-pay.
FC-16	By 2028, national Census BTOS AI use will reach 30.0% for three consecutive releases.	70%	2028	The estimate must remain at or above 30.0% for six weeks.
FC-18	The 2030 OEWS estimate of US software developer employment will exceed the 2026 estimate.	55%	2031	Compare the May published estimates for occupation.
FC-19	Revised US nonfarm business output per hour will grow at least 2.0% annually from 2026 Q2 through 2030 Q2.	65%	2031	Compute the geometric annual rate from the latest revised BLS index levels.
FC-24	TSMC's second Arizona fab will enter volume production by 2030.	92%	2031	TSMC must officially state that Fab 2 began volume production by year-end 2030.
FC-31	By 2029, two independent general purpose model developers will report qualifying model assisted production system improvements of at least 10%.	69%	2029	Two developers under different ultimate parents must document new post opening production cases on the same workload end-to-end cost, throughput, or training time denominator.
FC-32	By 2029, METR will publish a qualifying public model software-task horizon of at least 72 hours.	69%	2029	The central p50 estimate must appear in METR's public tracker or data after review. Private previews do not count.
FC-33	By 2029, an official OSWorld 2.0 evaluation will clear both the 50% full-set and 10% long duration binary completion thresholds.	57%	2029	A publicly available agent must meet both binary tests without human intervention. Partial reward does not count.
FC-34	By 2030, the US, EU, UK, or China will have kept a quantitative foreign remote compute access rule in force for at least 365 consecutive days.	63%	2030	The rule must require denial, a government license, or quantitative reporting above a numeric compute, accelerator, or capability trigger.

1 | Overview of the probabilities

The probabilities are my estimate that one exact proposition resolves to yes under a precommitted rule. It may sound obvious, but it changes the answer. A dramatic vendor result may be excellent evidence that a capability is moving and still be irrelevant to what metric actually resolves the forecast. A broad policy announcement may look important and still fail the threshold I set for resolution.

I'll touch on each forecast in more detail later in the paper.

The shape of the event can be important. Some events are one of many outcomes. FC-11 needs one qualifying state tariff from a growing set of proceedings, which is why it can be 88% even though no single commission is certain. Other events are conjunctions. FC-33 requires both 50% binary completion across OSWorld 2.0 and 10% in the long-duration group. Progress can be fast while the conjunction remains difficult, which is why the probability is only 57%.

I also consider the distance from the goal and the time available to reach it. FC-24 is 92% because TSMC says construction of Arizona Fab 2 is complete and targets volume production in the second half of 2027, while the forecast allows through 2030. The remaining 8% is qualification, yield, customer, and schedule risk. FC-32 is 69% for the opposite reason. The eligible public METR anchor is around twelve hours, so the forecast needs an accepted sixfold move to reach 72 hours.

Next comes the causal clock. Models and software can move in months. Data centers, fabs, and power systems move in years. Policy can change quickly on paper and slowly in enforcement. Organizations often move after the technology but before national productivity. The macro clock is not uniformly slow… capex, imports, relative prices, and hiring flows can move almost immediately, while broad productivity, occupational employment stocks, and the fiscal return arrive much later. I separate those responses because putting them all on one clock hides what I am trying to forecast. A benchmark slope cannot be converted into GDP, labor, or electricity as if every intermediate market clears instantly.

The models in this report do something more useful. They expose the parameters that change a sign, and reject impossible combinations. They tell me that employment cannot be inferred from productivity without an output response; that an efficiency gain can increase energy pressure when demand rebounds; and that imports can arrive before productivity. They do not generate the probabilities.

This is also where the vibe of the economy can play a part. Labs are using models to cut production costs and firms are committing enormous capital; utilities and regulators are suddenly fighting over who pays for power infrastructure; yet verified long duration computer use, national adoption depth, and realized productivity still lag. If the capability story were only marketing, the policy response should not be broadening this quickly. If frictionless takeoff were already here, field reliability and national productivity should not be this far behind. Seeing both sides at once is why I favor controlled diffusion.

Put together, my base case is scarcity first, reallocation second, broad productivity later. The shock is a falling cost of quality adjusted cognitive work. Lower cost expands the desired quantity of intelligence. The extra demand hits equipment, imports, sites, and power before firms reorganize enough to produce measured output. Governments then govern the scarce and observable complements. Firms and countries absorb at different speeds, so market share and rents move before the aggregate numbers look clean. The forecasts are intended to be observable cuts through this sequence.

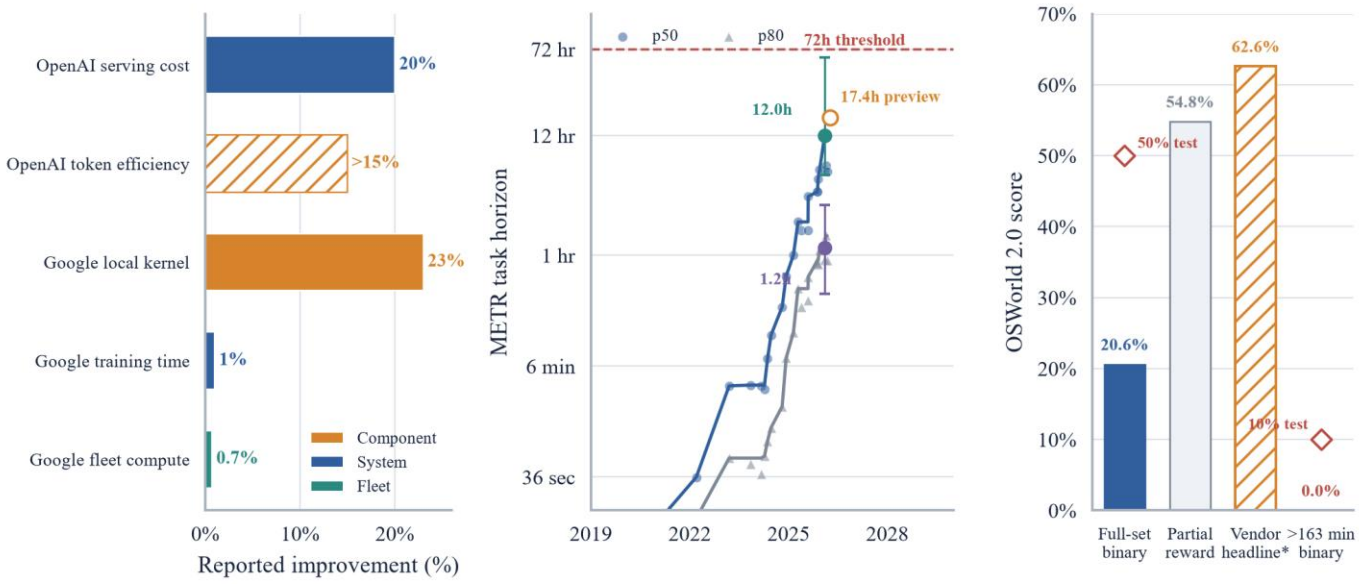
2 | Start with the supply shock

The economic primitive is a lower cost for a verified useful outcome. Token cost is only one input. An attempted workflow also consumes tools, latency, retries, and human review; the accepted result still has a probability of being wrong. A cheap attempt is not actually cheap labor if the firm pays someone to discover silent failure afterward.

This distinction makes recursive improvement easier to discuss. A model proposes a kernel, scheduling rule, evaluator, or decoding method. Humans choose the objective, test the output, reject bad proposals, and integrate the survivors. The improved system is then cheaper or faster, which funds more experiments. That is a positive economic loop even when it is bounded by hardware, training runs, verification, integration, and calendar time.

Exhibit 1 | Model-assisted engineering is beginning to lower the cost of producing AI

Model-assisted engineering is real, but local gains attenuate at the end-to-end boundary and long workflows remain a separate hurdle.



The exhibit should be read as a supply curve shift. Production cases establish that the loop exists; the gap between component and system outcomes, and between software task horizon and computer use completion, determines how quickly it reaches the wider economy.

Sources: OpenAI and Google production disclosures; METR Time Horizon 1.1; OSWorld 2.0.

OpenAI reported a 20% lower end-to-end serving cost and more than 15% better token generation efficiency from separate GPT-5.6 Sol assisted efforts. Google reported a 23% local kernel improvement that became a 1% reduction in total training time, plus 0.7% of fleet compute recovered from scheduling. These are selected proofs, and not estimates of average engineering productivity. Still, Amdahl's law does not care how exciting the local result is; a component that occupies a small share of the system cannot move the whole system by a large amount.

Now follow the logical consequence. Lower verified task cost makes more research, coding, simulation, and product experiments economically viable. It can erode the force of a static hardware threshold because accessible chips produce more useful output. It can also increase accelerator and power demand because firms spend the savings on more attempts, longer contexts, stronger models, and previously uneconomic work. Efficiency is therefore both a substitute for hardware at the task level and a complement to hardware at the market level.

This is why FC-31, FC-32, and FC-33 are not the same forecast. FC-31 is 69% because OpenAI and Google establish that the mechanism exists, but neither case can resolve the event. The yes requires two new developers under different parents to publish post-opening production gains of at least 10% on a same workload end-to-end denominator. Disclosure selection and component to system attenuation create most of the 31% no case. I also give 69% to a 72 hour public software horizon, but only 57% to the computer use conjunction. The base case is that coding, research, and systems work accelerate first. Broad office autonomy waits for state management, error recovery, and verification to catch up.

3 | The spending arrives before productivity

Once the desired quantity of intelligence rises, the shock leaves the frontier labs. It becomes demand for accelerators, memory, networking, datacenters, generators, transformers, construction, and finance. Firms do not need proof that productivity has risen; they only need the expected future margin on AI services to exceed capital, power, financing, and policy costs. Capacity must be ordered years before realized demand is visible. The cost of an idle data hall is obvious, but the cost of missing a capability discontinuity may be existential. That makes defensive over ordering rational. The near term macro analogy is an investment specific technology shock, not a neutral productivity shock: the economy has to build and import the capital before it can reorganize around it.

Exhibit 2 | The buildout enters the national accounts before the productivity dividend

Current investment and import leakage are observed; domestic-capacity and TFP paths are transparent scenarios used to discipline the timing.

Equipment and construction demand can lift domestic activity while net exports weaken. Localization reduces import dependence only as projects commission; broad productivity follows a separate organizational and diffusion lag.

Sources: Federal Reserve AI-buildout data and 1993-2019

The selected companies in the Federal Reserve panel reported roughly $156 billion of total capex in 2026 Q1. That is not just an AI flow and it does not equal domestic fixed investment or operating capacity. Separately, the Federal Reserve estimates that about 90% of the relevant high technology equipment category is imported. In its accounting proxy, related net exports subtracted 0.45 percentage point from annualized 2026 Q1 real GDP growth, offsetting contributions from software, computers, data centers, and power. AI can support domestic demand and worsen the current account at the same time. The accounting is simple: the current account equals national saving minus investment. If the investment boom is not matched by higher saving then foreign capital finances the gap. Strategic power and an import heavy capital boom can coexist.

A one sigma investment specific shock raises investment roughly 2% in the first year, output more gradually by about 0.5%, the current account deficit by about 0.2 percentage point of GDP, and import price growth by about 1 percentage point over the first year; TFP rises later.

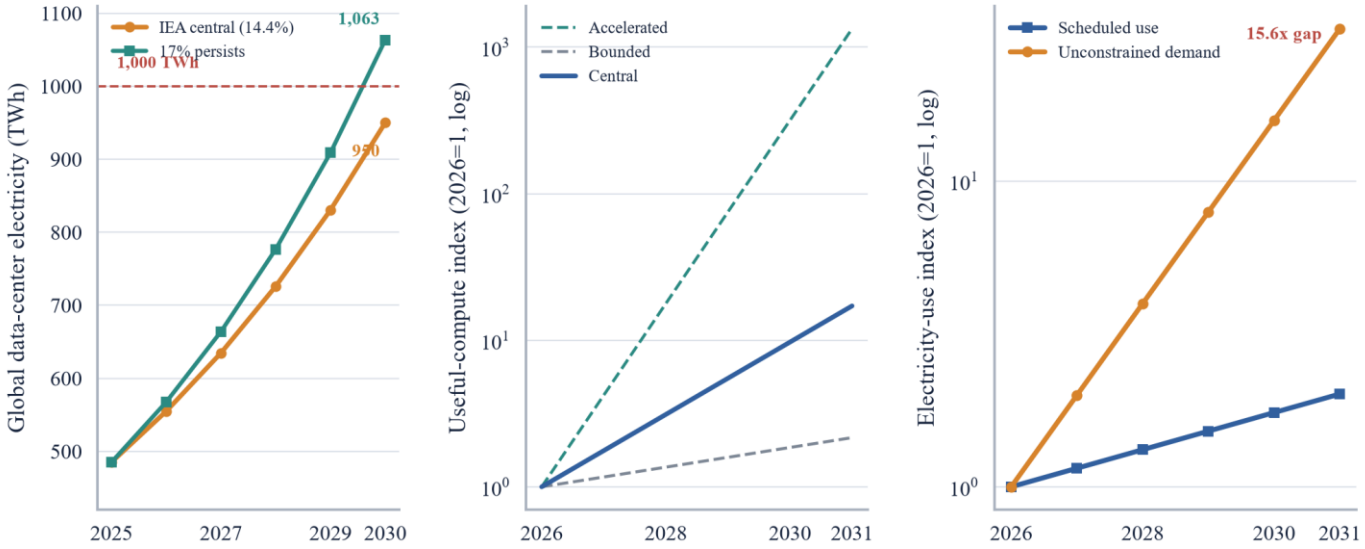
Modern mercantilism changes the geography and incidence of this sequence. Subsidies can pull a fab or data center onshore, but the project still imports tools, components, intellectual property, and sometimes power equipment before it becomes operating capacity. Redundancy improves resilience while raising capital intensity and duplicating fixed costs. The near term beneficiaries are not only model firms. They include the owners of fabrication, memory, electrical equipment, sites, engineering capacity, and finance.

FC-24 is the cleanest test. Announced is not ordered; ordered is not constructed; constructed is not connected and qualified; qualified is not utilized. Each transition deserves a realization haircut, and the haircut depends on where the forecast starts. I put TSMC Arizona Fab 2 at 92% because construction is complete, management targets 2027 volume production, and the event allows a slight buffer. What the probability does not say is that the United States becomes semiconductor autonomous. One commissioned fab changes resilience and bargaining power; it does not reproduce the entire ecosystem.

4 | Power is where the rebound has to clear

People often assume that making inference more efficient will then electricity demand. That is true only if the desired quantity of AI work rises by less than efficiency. If lower cost creates more than a proportional increase in useful tasks, or if each task becomes more compute intensive, total physical demand still rises. The engineering gain survives; it is spent on quantity and quality rather than conservation.

Exhibit 3 | If workload growth outruns efficiency, desired compute must clear against deliverable power



The question is how the market clears when desired compute grows faster than sites, interconnection, equipment, and generation.

Sources: IEA Energy and AI; FERC large-load proceedings.

The IEA estimates about 485 TWh of data center electricity consumption in 2025 and a central 950 TWh in 2030. FC-08 asks if the first qualifying estimate reaches 1,000 TWh. That requires about 15.6% annual growth from the 2025 number, compared with roughly 14.4% on the central path and around 17% in 2025. I leave the event at 68% because workload growth is strong, but efficiency, duplicate queue requests, cancellations, and delayed projects can keep realized use below the announced demand.

The constraint is a local and contractual shortage of reliable power at the right site, voltage, date, and risk allocation. A project can have generation in the region and still wait for interconnection, transformers, transmission, permits, or an acceptable tariff. The market clears through higher prices, long term contracts, take or pay commitments, queue attrition, geographic migration, model compression, and priority access.

This is why FC-11 is 88%. FERC has required the regional operators to justify or reform large load rules, and multiple states have proceedings. One post opening tariff is a one of many event. The residual 12% comes from the resolution I set, at least 100 MW and either 80% dedicated cost allocation or ten years of take or pay or equivalent exposure. A generic data center tariff is not enough.

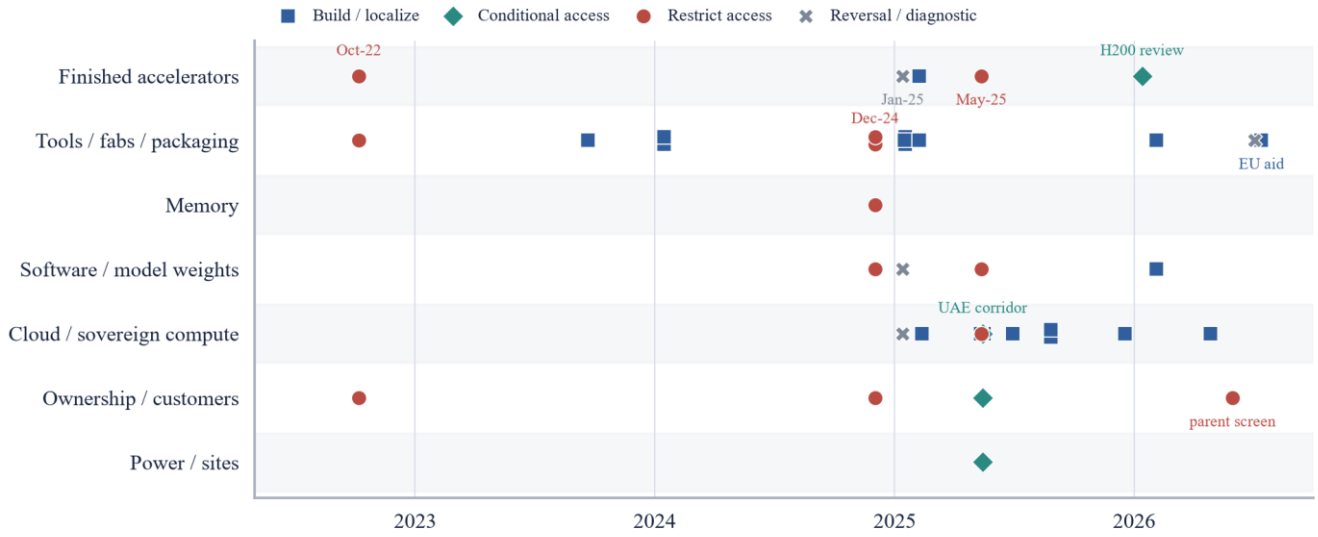
Power and construction are among the first places the shock can raise relative prices. That is not yet the same as persistent aggregate inflation. A lasting inflation process requires broader wage, expectation, or second round pass through, which the current evidence does not identify. Simultaneous private and public buildout raises the demand for saving and can put upward pressure on the n eutral real rate, but only if expected returns and investment demand outrun the saving response. The medium term disinflation story is possible.

5 | Mercantilism governs the bottleneck

What does an export control have to accomplish to work? The answer changes the forecast. A rule can fail to stop absolute Chinese capability and still delay a frontier system, widen a relative lead, screen a customer, move a factory, or give the United States bargaining power over an allied compute corridor. Treating all of those objectives as containment is when the policy debate gets confusing.

Exhibit 4 | The control perimeter is following the scarce and observable parts of the stack

Dated policy actions have moved beyond finished accelerators toward equipment, memory, software, ownership, customers, cloud access, sites, power, and subsidized capacity.



Policy migrates toward layers that are slower, attributable to a customer, or easier to finance and condition.

Sources: Official policy releases.

This makes FC-01 at 90% and FC-02 at 80% entirely realistic. I expect an H200 export license to survive in 2030, and I also expect the defined nonnegative US lead on the public-model metric to be no more than 5%. Hardware controls can tax the accessible compute stock and preserve a relative frontier advantage without creating a lasting public model capability gap. Common software gains alone do not erase a relative compute advantage; they just make it harder for that advantage to produce a large, persistent gap on this particular denominator.

What can the state actually observe and enforce? Model weights are portable, thresholds age quickly, and the base rate for durable cross border software controls is low, so FC-03 remains 42%. Remote compute is more observable: a cloud provider can know the account, customer, accelerator class, capacity, and location. That makes a quantitative access or reporting rule more plausible, which is why FC-34 is 63%. Bilateral corridors are easier still because chips, financing, ownership, security, and end use terms can be bundled in one agreement; FC-04 is 77%.

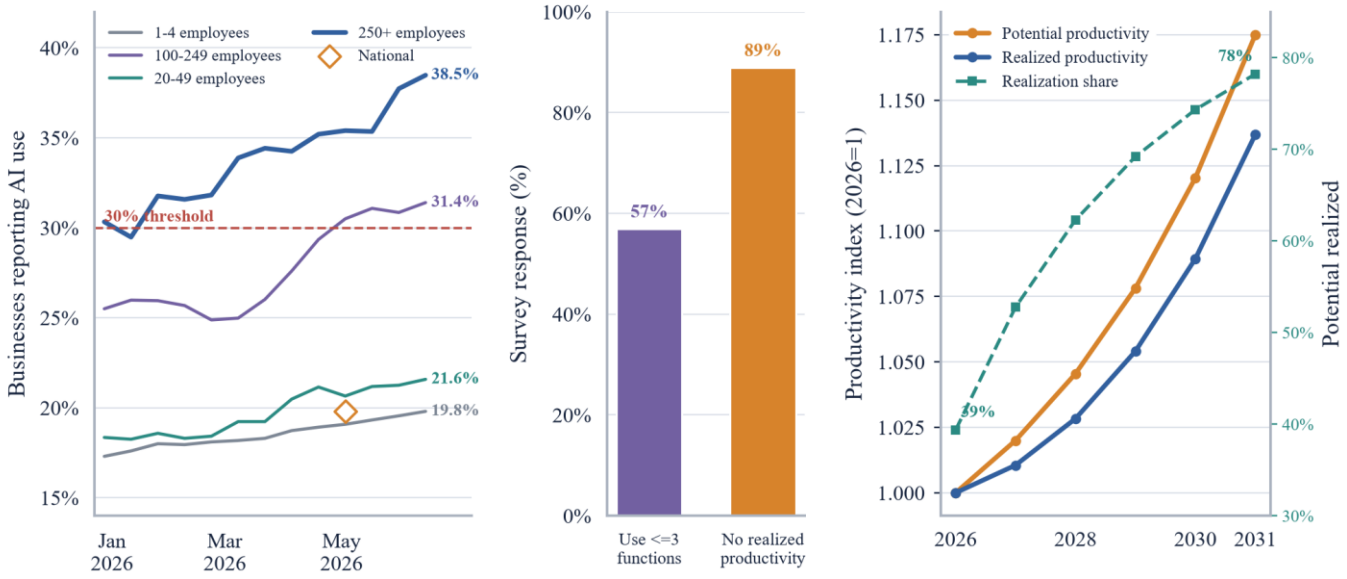
The macro result is then controlled diffusion. Trusted jurisdictions can receive financed access and accept monitoring or security conditions. Chokepoint owners can capture fabrication, memory, equipment, and intellectual property rents. Energy rich hosts can exchange power and capital for a place inside the stack. Countries that possess only one complement can still lose the income if foreign capital, models, cloud platforms, or equipment suppliers own the rest.

So where does modern mercantilism actually have leverage? It is strongest where the controlled layer is binding, observable, enforceable, and difficult to substitute. It will redistribute rents, locations, and strategic options more reliably than it contains knowledge. The cost is duplication, allied subsidy competition, smaller production runs, higher capital intensity, and a larger public balance sheet role. The state can tax intelligence at the toll road. It cannot make software obey a border in the same way that a physical wafer does.

6 | Absorption is the missing national account

Buying access is fairly easy. Changing production is the difficult part. A model can perform a task and still create no measured output if the worker does not trust it, the data is inaccessible, the workflow cannot change, the evaluator is weak, or management will not delegate authority.

Exhibit 5 | Access is broadening while organizational absorption remains shallow



Productivity requires breadth across functions and depth in data, permissions, process redesign, training, and verification. The cost of building those complements is paid before much of the benefit appears.

Sources: Census BTOS and Microstructure of AI Diffusion; multicountry firm survey; phase-two organizational-capital model. Survey denominators differ; modeled paths are scenarios.

This is where the denominator gets us in some trouble. National BTOS use is 19.8%, not 38.5%. The higher figure is the large firm series. Large firms have data, software, integration teams, legal capacity, and enough repeated workflow volume to justify fixed adoption costs. Small firms often buy the same frontier capability but cannot build the stack as quickly. FC-16 is therefore 70%, not a near certainty: the national series must rise another ten points.

Even adoption is only the extensive margin. Fifty seven percent of adopters use AI in three or fewer functions, and 89% of surveyed firms in a separate study report no realized productivity impact. Each additional AI using function is associated with a 1.4 point higher probability that a firm reports increased sales and a 1.2 point higher probability that it reports an employment decrease. The estimates are descriptive, but they show why shallow use and deep redesign should not be averaged into variable.

A national adoption checkbox reaches output only through several gates. Access must spread, use must deepen across functions, management must delegate authority, reliability must make review economical, and the resulting firm gains must aggregate rather than merely transfer share between competitors. If any gate remains close to zero, broad adoption can coexist with weak national productivity.

The organizational J-curve connects this to FC-19. Firms expense process mapping, data work, security, evaluation, and training before the new production system is fully used. The central research scenario raises the realized share of potential productivity from roughly 39% to 78% over five years. Those values show how much organizational timing is required before a frontier task gain can plausibly appear in output.

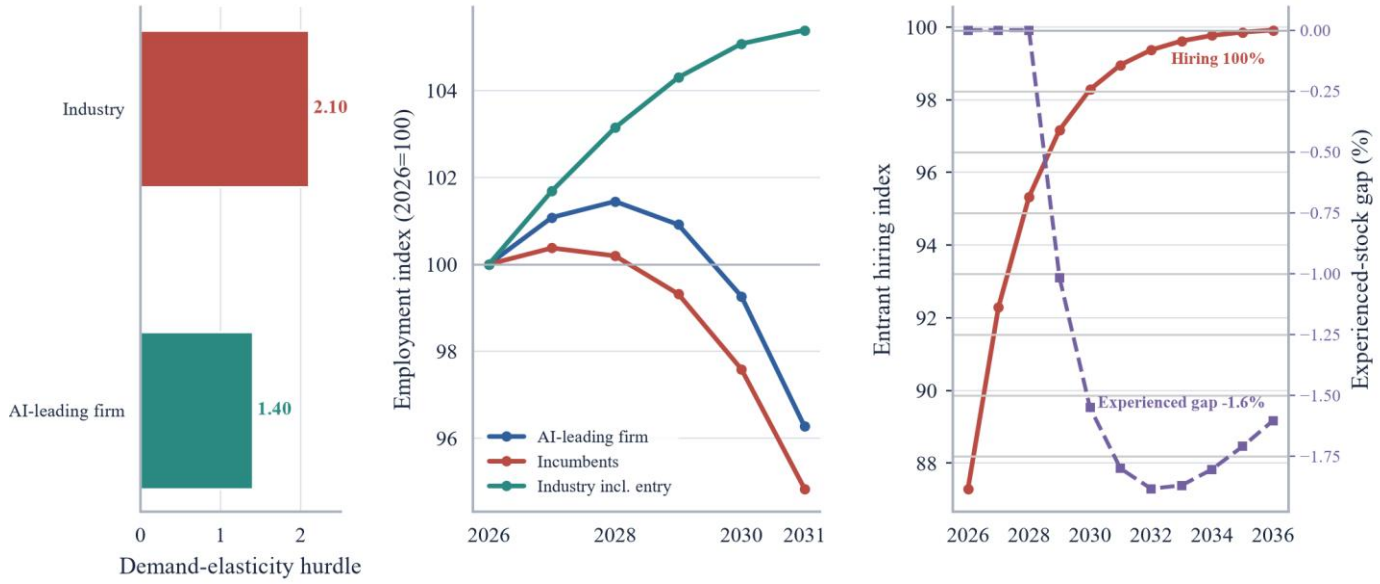
This is also a national power problem. A country can buy accelerators and still import the model, cloud service, engineering, and profits. Durable gain requires physical access, organizational absorption, and a way to retain income. FC-19 remains 65%, but it is one of the portfolio's most subjective probabilities. In 2026 Q1, high exposure sectors showed 1.49% quarter on quarter log productivity growth on a four quarter moving-average measure, against 0.17% for low exposure sectors; that frequency differs from FC-19's four-year geometric rate, and the comparison does not identify AI causally. Against that, 89% of surveyed firms still report no realized effect, functional use is shallow, and the event requires four years of at least 2% aggregate output per hour growth. A yes would not prove AI caused it: capital deepening, reallocation, cyclical hours, and benchmark revisions can move the same index. I could not get a clean historical base rate for four year 2% windows, so the scenario model disciplines timing rather than generating 65%.

7 | Maximum output is not maximum employment

My original instinct was that a much more productive firm would not settle for producing the same amount with fewer people. It would lower prices, enter adjacent functions, launch products that were previously uneconomic, and take share. I still think that is the right way to analyze a winning firm. It does not settle the industry or national labor result.

Exhibit 6 | With average hours fixed, employment rises only when output outruns labor productivity

The leader, the incumbent industry, new entrants, and the junior experience pipeline can move in different directions at the same time.



Share capture makes the employment hurdle easier for an AI-leading firm, but one firm's gain is another firm's loss at the industry level. Aggregate labor depends on demand elasticity, price pass-through, new scope, entry, and the conversion of junior cohorts into future expertise.

Sources: phase-two labor-demand, firm-reallocation, and junior-pipeline models; Census early-career study; Federal Reserve job-posting panel; Babina et al. Models are sign and timing diagnostics, not employment forecasts.

The accounting starts with hours: labor hours growth equals real output growth minus output per hour growth. Headcount adds average hours per worker. A 20% task gain says nothing about employment until we know what happens to price, demand, product scope, and market share. Under the central model assumptions, industry employment requires demand elasticity above roughly 2.10, while a leader taking share needs roughly 1.40. Those thresholds hold only under the model's assumed pass-through, margin capture, productivity realization, and scope effects.

The evidence supports both sides of the mechanism. In a 2010-2018 sample of US public nontechnology firms, a one sigma increase in AI skilled worker share was associated in the fully controlled specification with 31.7% higher sales, 33% higher employment, and 36% higher market value. A related industry specification found roughly a 1.4 point increase in top firm sales share. This predates generative AI and is a reference class for skill led expansion and reallocation. Task experiments, meanwhile, range from a 19% slowdown for experienced developers in a difficult familiar repository setting to gains above 30% for novices in support work. There is no universal worker multiplier.

The most troubling labor signal is a flow, not a layoff wave. Exposed early career hires are 12.7% below the reference path, while lower separations offset part of the stock decline. A separate Federal Reserve firm panel finds only a 0.04%-0.13% scaled job posting effect and describes it as a precise null. There is an option value mechanism here: when future task content is uncertain, a vacancy is easier to defer than an incumbent with firm specific knowledge is to dismiss. Firms may also be responding to economic and sector shocks, so the current evidence cannot assign the whole gap to AI.

Junior labor is also investment. Today's entrants become tomorrow's verifiers, supervisors, originators, and domain experts. Cutting the cohort can improve near term margins and create a delayed experienced worker shortage. AI tutoring could offset the smaller cohort, or delegation could remove the repetitions that build judgment. The evidence supports AI workflow skill, not youth itself: younger workers may be more fluent with the tools, but domain judgment, evaluation ability, and authority to redesign a proce ss are separate complements. FC-17 at 67% tests whether early career compression persists. FC-18 at 55% shows that software demand expansion and agent substitution are almost balanced; yes can still be below the no AI counterfactual because the event compares published employment levels, not causal paths. My base case is maximum output with selective headcount compression and substantial firm reallocation,

8 | What the collision produces

So what does this actually look like in the macro data? AI is a positive supply shock to cognitive work, but the response over the next few years is not a frictionless productivity boom. Lower cost expands the desired workload. Desired workload pulls investment, imports, power, and scarce engineering forward. Modern mercantilism redirects access and rents across countries. Organizational capital determines which installed systems become output. Firm winners take share before the aggregate economy catches up.

In the **first stage**, the visible macro variables are capex, equipment imports, construction, electricity contracts, grid proceedings, and policy conditions. This stage can widen the current account deficit, support selected industrial prices, and keep capital demand strong. The important country exposures are ownership of the frontier stack, fabrication and memory chokepoints, reliable power, project finance, and access to trusted corridors.

In the **second stage**, firms separate. Large and technically capable organizations absorb first. Some use AI to reduce cost; others use it to expand products and take share. Concentration can rise even if entry rises, and a leading firm can hire while incumbent industry labor falls. Early career hiring can weaken before measured productivity because waiting has option value when future task content is uncertain.

In the **third stage**, commissioned capacity first raises capital services and can lift labor productivity through capital deepening. Within firm redesign and reallocation are what can lift measured TFP later. Exposed service prices can fall and real income can rise, while labor share and income distribution remain contested. Scarce complements can still generate relative price pressure. Persistent inflation requires broader wage, expectation, or second round pass through, so "AI is inflationary" and "AI is disinflationary" are incomplete statements unless the layer and date are specified.

Now ask who pays before the productivity arrives. Subsidies, guarantees, tax credits, defense procurement, grids, and public finance arrive before taxable operating income. The fiscal return depends on domestic wages, taxable profit, ownership, financing cost, and stranded asset risk. Monetary policy has to distinguish today's investment demand from tomorrow's supply; it cannot book future productivity against fixed capacity today. High rates can slow marginal projects while increasing concentration because cash rich frontier firms can self finance and smaller firms or utilities cannot.

The exchange rate is ambiguous for the same reason. Foreign capital financing profitable US capacity can support the dollar even as equipment imports widen the current account deficit. A host can record construction and GDP while remitting the eventual operating income abroad, so GDP and national income can diverge. A country that imports the hardware, model service, and engineering but fails to retain wages or profits can experience the opposite. Financing conditions therefore decide which announced projects survive and who owns the income stream.

Modern mercantilism does not stop that sequence; it changes its incidence. Frontier stack owners retain model, cloud, capital, and intellectual property rents. Manufacturing chokepoint owners monetize fabs, memory, equipment, and packaging. Energy rich trusted hosts can sell location and power. Rule-making economies can shape access without capturing production. Countries missing several complements at once become price takers, even when they can access capable models over the internet.

The view would break in recognizable ways. Repeated model assisted production gains could stop at isolated components. Refreshed agent suites could hold public horizons below twelve hours and long computer use completion near zero. National adoption could stall below 22% and remain shallow. Desired compute could prove inelastic enough that efficiency actually lowers total power demand. Export restrictions could retreat rather than migrate. Productivity could fail to rise after the investment and absorption window. A capability plateau would turn defensive over ordering into cancellations and an overbuild bust; faster long duration reliability would pull productivity and power scarcity forward together; hard fragmentation would lower global efficiency while increasing fiscal duplication.

The opposite evidence would strengthen the chain: replicated end to end production gains, reliable long workflows, national use above 24% with broader functional depth, delivered electricity above the IEA central path, operating domestic capacity, and quantitative cloud access rules. I would update each forecast through its resolution metric before allowing one data point to move the entire cluster. If several independent links fail together, the narrative itself will have to change. A vendor score does not resolve a maintainer metric; a large-firm adoption rate does not resolve the national series; an announced project is not volume production.

The scarce asset is not intelligence by itself. It is the ability to attach cheap intelligence to power, capital, trusted access, domain knowledge, and an organization capable of changing… and then retain part of the income. That is where AI and modern mercantilism collide. The result is controlled diffusion: scarcity first, reallocation second, and broad productivity later.

Selected sources

Technology, agents, and firms	Macro, energy, and policy
OpenAI - GPT-5.6	Federal Reserve - Technology shocks and the US current account
OpenAI - Building abundant intelligence	Federal Reserve - The AI buildout and the economy
METR - Measuring AI ability to complete long tasks	Federal Reserve - AI adoption and firms' job postings
OSWorld 2.0 paper	Census - AI use by business size
Stanford - 2026 AI Index technical performance	BIS - China semiconductor license-review policy
TSMC - Q1 2026 transcript	IEA - Key questions on energy and AI
BLS - Major-sector productivity	FERC - Large-load integration action
BLS - Occupational Employment and Wage Statistics	Census - Early-career hiring in QWI