# The Economic Leverage of Global AI Hegemony

Gonzalo Huertas *
IMF

April 2026

**Abstract**

This paper estimates the macroeconomic impact of removing access to frontier AI services from a country that has grown to rely on them as an input to its economy. Applied through the production network framework of Baqaee and Farhi (2019) and calibrated to a representative small open economy with levels of AI embeddedness consistent with current trends, the headline GDP loss of "AI sanctions" is approximately 0.8 percent. The range of possible outcomes can vary significantly, however, depending on the future convergence path of competitor models and on the substitutability of AI across sectors. In the most severe scenario –with very high sectoral complementarity and a widening AI frontier gap– the economic impact of AI sanctions can be as high as 2.2 percent of GDP, similar in magnitude to the 1973 OPEC oil shock.

*Early draft, do not circulate. The views expressed herein are those of the author and do not represent the views of the IMF, its Executive Board, or IMF management.

# 1 Introduction

When the world's state of the art AI services are provided by close-weight firms concentrated in a single country, the global supplier can deny other countries access to those models. This paper estimates the macroeconomic impact of a sudden cutoff from frontier AI services in a sample country that has grown to rely on them as an intermediate input in domestic production.

I model a scenario where a small open economy imports AI services from the global frontier supplier, such that any interruption in these services would force domestic firms to switch to alternative, less productive models. I ground the exercise on the production network framework in Baqaee and Farhi (2019), using revenue forecasts from the current top global AI service providers (OpenAI, Anthropic, and Google) and productivity measures from AI capability evaluators (METR, SWE-Bench, GPQA) to calibrate the first and second-order effects of an AI access shock. This scenario is particularly suited to represent unilateral sanctions or export controls by a global hegemon in AI services, but could alternatively be interpreted as representing cybersecurity measures.

The headline estimate is that this shock would reduce the targeted economy's GDP by 0.77 percent. However, the full set of results comprise a wide envelope, from as little as 0.12 percent at the most benevolent end of the parameter space, to as high as 2.27 percent at the most pessimistic corner where complementarity between AI and its other sectors is very high and the capability gap between frontier and non-frontier models is especially wide.

The exercise requires constructing two empirical primitives that the existing literature has not jointly produced: the pre-shock Domar weight of global frontier AI services in the representative country's GDP, and the productivity loss from forced substitution to the next best alternative model. together with the macro elasticity of substitution between AI and the rest of the economy. I build the Domar weight bottom-up because AI services are not yet a separate line in standard input-output accounts. I project the revenues of the three leading global providers through 2028 from disclosed run rate trajectories and assign a country shares consistent with observed regional adoption patterns.

The productivity loss is calibrated through two complementary estimators that lean on different empirical approaches: one combines METR's task-horizon reliability curves with the Anthropic Economic Index's task-length distribution of enterprise API usage, the other provides a more conservative lower bound by aggregating a basket of standardized benchmarks (SWE-Bench Verified, GPQA Diamond, OSWorld, Chatbot Arena) with weights grounded in the US Bureau of Labor Statistics' Standard Occupational Classification (Handa et al., 2025). The macro elasticity between AI and other sectors in the economy is calibrated as a range given by available estimates from the literature, spanning short-run displacements in production networks to medium-run commodity shocks (Atalay, 2017; Boehm et al., 2019; Allcott et al., 2016).

This paper contributes to three adjacent literatures. It is, to the knowledge, the first empirical application of the Baqaee and Farhi (2019) framework to a forward-looking policy shock of a candidate general-purpose input. It bridges the task-level AI productivity literature (Noy and Zhang, 2023; Brynjolfsson et al., 2023; Eloundou et al., 2024) to a country-level Hicks-neutral shock representation, providing a transparent operational link between the two. And it offers a quantitative envelope for the economic costs of sudden discontinuation of global frontier AI access once the technology has become well embedded into an economy – a process that is gaining increasing speed both across advanced and developing markets.

The remainder of the paper is organized in four parts. Section 2 provides background information on the extent to which countries have adopted AI into their economic processes, and identifies the leading AI global suppliers and their next-best counterparts. Section 3 develops the Baqaee and Farhi (2019) framework in the form needed for a sanction-as-TFP-shock representation, and derives the two sector formula that pins down necessary parameters. Section 4 calibrates the exercise. Section 5 maps them through the framework into the final GDP impact. Section 6 concludes.

# 2 Background: AI's role in the global economy

## 2.1 The growing embeddedness of AI into production

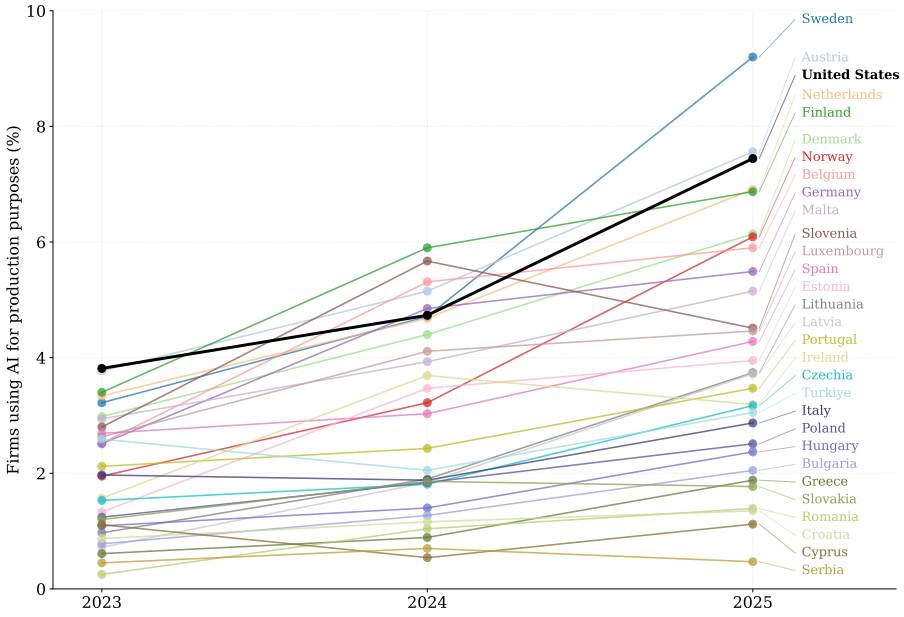
The integration of artificial intelligence into production has become, by 2026, a measurable global trend. Worker-level generative-AI use and firm-level adoption rates have shifted by an order of magnitude 2022. The pace and depth of that integration vary substantially across countries, however, with the United States sitting at the leading edge on most measures, a top tier of advanced economies catching up rapidly, and a long tail still in the early stages of the transition (Bick et al., 2026; Sajadieh et al., 2026).

The US–EU gap on enterprise AI adoption is roughly two to three years, on both the extensive margin and the intensive margins (Sajadieh et al., 2026). Data on AI use in the production of goods and services at the firm level in 2025 shows a disparity between the US at 7.4 and the EU27 at 4.1 percent, with only Sweden and Austria scoring slightly higher. Adoption is concentrated in the information-economy sectors most exposed to language models and generative tools: Information at 37 percent of firms, Professional, Scientific and Technical Services at 33 percent, and Financial Activities at 30 percent, against single-digit shares in Accommodation and Food Services and low-teens in Wholesale Trade (Allen, 2026). At the worker level, 43 percent of US employees reported using generative AI as of 2026, against a European range of 26 percent in Italy to 36 percent in the United Kingdom (Bick et al., 2026).

The embededdness of AI into the economy is rising fast. The share of firms using AI for production purposes in the US rose by 2.7 percentage points in between

2024 and 2025, with EU countries growing at an average rate of 1 percentage point over the year, and top performers showing an increase close to that of the US (Figure 1). The pace of this diffusion is consistent with previously-reported intent in corporate surveys: at the start of 2025, around 9 percent of US firms reported planning to adopt AI that year, remarkably close to the final observed adoption rate. Given that over 20 percent of US firms now expect to use AI in 2026 (Allen, 2026), the embeddedness of generative technology into the economy is poised for a sharp acceleration this year – and takeup rates are even higher when surveys include incidental use of any consumer-grade chatbot (Board of Governors of the Federal Reserve System, 2026; PricewaterhouseCoopers, 2026).

**Figure 1:** Share of firms using AI for production purposes




Source: Own calculations based on data from Eurostat and US Census Business Trends and Outlook Survey, harmonized following Bick et al. (2026).

In tandem with a rapid increase in sectoral diffusion, AI use is also growing at the intensive margin. AI-using hours accounted for 5.2% of total work hours in the US in 2026, AI, roughly double the share in the United Kingdom, Sweden and the Netherlands, and more than triple the share in Germany, France and Italy (Bick et al., 2026). The corresponding estimate of AI-attributed time savings, averaged across users and non-users, runs from 1.0% of work hours in France through a six-country European mean of 1.4% to 2.3% in the United States. In the specific case of Anthropic's Application Programming Interface (API) access, the average conversation now handles a task that would require

1.75 hours of skilled human work (Anthropic, 2026a).

The catch-up trajectory is steep. At the observed adoption rate paths, by 2028 most advanced economies are set to reach levels of AI integration comparable to the current embeddedness in the United States. On the worker side measure –the share of working-age adults reporting generative AI use– the European frontrunner gap to the current US level closes within one to two years, and the slowest-laggard gap within three to four. On the firm side production-purpose measure, the EU27 average arrives at the current US production rate by late 2027–2028, with the Nordic and small-Western-European countries approaching earlier. On the consumer and per-capita margin, several non-US economies had already moved past the United States as of February 2026.[1]

Outside Europe, comparable cross-country data is scarcer, but uptake is accelerating. Whereas early research suggested that advanced economies in Asia adopted AI in their own production substantially less than continental Europe (Calvino and Fontanelli, 2023), more recent consumer-side data shows substantial catching up. Singapore and Israel are the world's most intensive per-capita users of Anthropic's Claude family of models, with Australia, South Korea and New Zealand sitting close the top ten (Appel et al., 2025). In Türkiye, 7.5% of enterprises with ten or more employees used at least one AI technology in 2025 –a level that is roughly comparable to Poland (Türkiye İstatistik Kurumu (TÜIK), 2025).

China's adoption of AI into the economy has been especially rapid since 2025, bolstered by the local development of alternative, low-cost models. Deepseek, in particular, has been integrated into major sectors including automakers, state-owned telecommunications operators, commercial banks, and provincial and city governments (Rest of World, 2025; Sixth Tone, 2025; South China Morning Post, 2025). China stands as an exceptional case where AI adoption has relied almost entirely on local labs rather than US models – with some estimates placing DeepSeek's share of AI use in the country at 89% (Microsoft AI Economy Institute, 2026). This diffusion may be partly due to active policy, with a recent State Council initiative explicitly targeting a penetration of 70% for AI agents

[1]The worker-side closure window combines the documented US worker use growth rate with assumed European frontrunner growth. US worker use grew from approximately 28% in August 2024 (Bick et al., 2024) to 43% in early 2026 (Bick et al., 2026), equivalent to $\sim 10.6$ percentage points per year. The European frontrunner (the United Kingdom at 36% in 2026) closes the 7-percentage-point gap to the current US level in roughly 0.9 years at 8 pp/yr or 1.4 years at 5 pp/yr; the slowest laggard (Italy at 26%) closes its 17-percentage-point gap in 2.1 years at 8 pp/yr or 3.4 years at 5 pp/yr. The firm-production-purpose closure window assumes the E_AI_PPP subset tracks the broader Eurostat "any business purpose" measure at a constant subset-to-broad ratio of 0.20 (the 2025 reading of 4% over 20%). Under that ratio, broad-measure growth at the 2024–2025 pace of 6.5 pp/yr puts the EU27 production-purpose rate above the current US BTOS production rate of 7% in late 2027, while a moderation to 5 pp/yr puts the crossing in late 2028. The 35% broad-measure projection applies the same 5-pp/yr moderate-continuation rate to the 2025 EU27 base. All three projections are to the *current* (early-2026 to mid-2025) US reading on each measure, not to a contemporaneous 2028 US level: if US adoption continues to grow on each margin through the projection horizon, the contemporaneous gap closes later than the dates above.

by 2027 (State Council of the People's Republic of China, 2025).

## 2.2 The US as the global leader in frontier AI services

AI covers a wide range of products with very different features. For commercial large language models, a distinction can be drawn between those at the frontier –with the highest capabilities on length of task and industry benchmarks– and the next available substitutes, which trail by varying margins on the same dimensions. At the time of writing, the global supply of frontier AI services appears entirely concentrated among three US providers: OpenAI, Anthropic, and Google's Gemini and Vertex platforms. Their models are of the closed-weight type, where the underlying architecture and trained parameters are not publicly available, and access is provided only through the companies' APIs and managed interfaces.

Alternatives to frontier models exist and have improved rapidly: the Chinese cluster (including Deepseek, GLM, Kimi, and Qwen), Mistral, and Meta's Llama family are well established in the general public. Unlike frontier US providers, these are open-weight models: publicly available to download, run, and modify on any hardware or private server, without relying on the developer's environment.

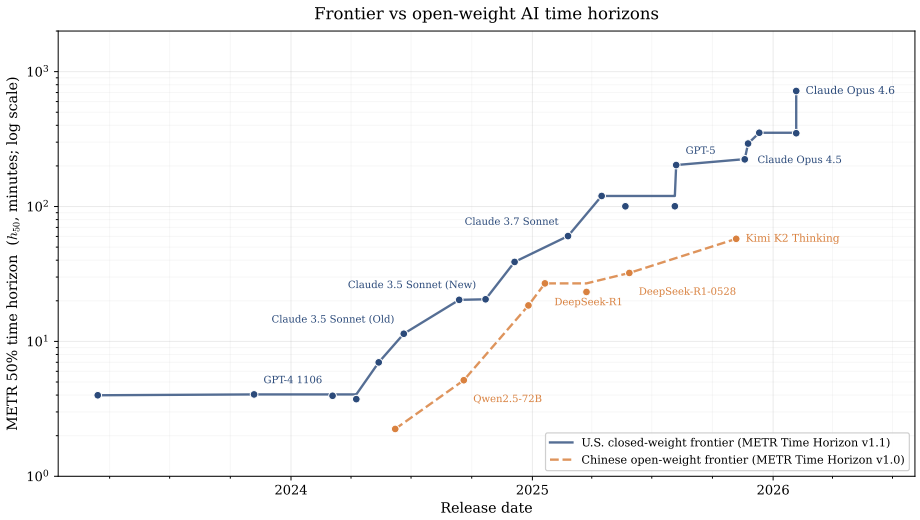
Whether closed-weight US models are technically superior to their open-weight alternatives depends on the exact type of capabilities being measured. Open-weight models, especially those of Chinese origin, are at or very near parity when it comes to single self-contained skills like factual recall, graduate level science, contest mathematics and isolated coding problems.[2] Kimi and DeepSeek sit within roughly two points of the closed-weight frontier on the American Invitational Mathematics Examination 2026, and three open-weight Chinese models occupy the top three slots in the LMArena ranking (Mathematical Association of America, 2026; LMSYS, 2026; CodeSOTA, 2026). However, once focus shifts to whether AI can sustain an extended, multi-step task in a realistic environment –executing long agentic chains, shipping a working patch to a live codebase, or maintaining coherence across extensive sequential calls– the US frontier leads by a significant margin (CodeSOTA, 2026; METR, 2025b).

It is this latter class of capabilities that is likely to determine the speed and extent with which AI becomes embedded in the wider economy. Here the frontier gap is clearest along the industry standard *METR task completion task horizon*

[2] In early 2026, the top ranks in the SWE-Bench Verified, Chatbot Arena and GPQA Diamond benchmarks were held by some combination of Anthropic's Claude Opus 4.6, OpenAI's GPT-5.4 and Google's Gemini 3.1 Pro (LMSYS, 2026; Sajadieh et al., 2026). The leading open-weight alternatives–DeepSeek V3.2, Qwen 3.5, GLM-5/5.1, Kimi K2 Thinking, Mistral Large 3 and the Llama 4 family– sat at a modest distance behind (BenchLM, 2026). On Chatbot Arena specifically, the gap between the top US and top Chinese model narrowed to roughly 2.7% in Elo terms by March 2026, down from a 17.5-31.6 percentage-point spread on the MMLU+MATH+HumanEval composite at the end of 2023 (Sajadieh et al., 2026). By this measure, the gap on less complex tasks has compressed sharply over the past two years.

metric: the length of software engineering tasks a model can reliably complete end-to-end. Figure 2 plots the three-year trajectory of this measure, comparing frontier US models against their alternatives. The divergence widened over the course of 2025 and early 2026, with the most advanced US model evaluated by METR (Claude Opus 4.6) about nine months ahead of its closest open-weight competitor.[3]

**Figure 2:** Frontier vs. open-weight AI time horizons, 2023-2026




Trajectory of the METR 50-percent time horizon ($h_{50}$, minutes, log scale) from GPT-4 (March 2023) through the April 2026 frontier, for two model families: the US closed-weight frontier (blue, GPT-4 through Claude Opus 4.6) and the Chinese open-weight frontier (orange, Qwen2-72B through Kimi K2 Thinking). Each marker is an individual release; lines trace the running maximum of $h_{50}$ within each family. All numeric values are taken directly from METR's public time-horizons dashboard (METR, 2026), which scores models against the Time-Horizon 1.1 reference suite of METR (2025b); all values reflect METR's March 2026 regularization correction (the Claude Opus 4.6 v1.1 reading of $\sim$720 minutes supersedes the preliminary 14.5-hour figure circulated before the correction). The latest Chinese open-weight model with a published METR score is Kimi K2 Thinking (November 2025); the April 2026 release wave (Kimi K2.5, Kimi K2.6, GLM-5.1, DeepSeek V4-Pro) has not yet been METR-scored, and I plot no values for these models.

The gap between the US frontier and open-weight alternatives has remained significant even after the latest wave of releases in April 2026. OpenAI's GPT-5.5 and Anthropic's Claude Mythos Preview show double-digit improvements over the previously leading Opus model on industry benchmarks like SWE-Bench

[3]The METR numbers have not yet incorporated the latest models from 2026, including Claude Mythos, GPT 5.5, DeepSeek V4, GLM-5.1 and Kimi K2.6. Other estimates of the frontier gap in 2026 suggested a range closer to six months (Epoch AI, 2026b; Bloomberg, 2026d; PolitiFact, 2026). Since METR benchmarks are an industry standard, however, I anchor the baseline discussion around the available data.

and Terminal-Bench 2.0 (Jimenez et al., 2024; OpenAI, 2026b). Evaluated on those same benchmarks, the newest Chinese open-weight alternatives, DeepSeek V4 Pro and GLM-5.1, perform fifteen to nineteen points below GPT-5.5, with DeepSeek's own technical report characterizing V4 as on a "trajectory that trails state-of-the-art frontier models by approximately 3 to 6 months" (DeepSeek-AI, 2026; Z.ai (Zhipu AI), 2026).

These three facts together –a target-relevant scale in reach by 2028, a US AI footprint at or above that scale already, and a meaningful and widening frontier lead over admissible substitutes– motivate the exercise that follows. Section 3 takes up the theoretical framework that converts these inputs into a macroeconomic answer.

# 3 The Model

The empirical exercise is grounded in the framework of Baqaee and Farhi (2019), which extends Hulten (1978)'s theorem to second order and shows how sectoral productivity shocks can have outsized macroeconomic effects in efficient, multi-sector economies with a production network. The framework suits the question for three reasons. First, the sanctions I consider are large enough that first-order approximations are likely inadequate. Second, they are concentrated in a single upstream sector – frontier AI services – that most downstream sectors cannot easily substitute away from at short and medium horizons. Third, the Domar-weight machinery at the core of the framework is the natural bridge between the input-output evidence on which the estimates rest and the aggregate GDP numbers I report.

## 3.1 Environment

Consider a perfectly competitive economy with a representative household whose consumption aggregator $C(c_1, \ldots, c_N)$ is homogeneous of degree one, where $c_i$ is final consumption of good $i$. Goods $i = 1, \ldots, N$ are produced by competitive firms using Hicks-neutral production functions

$$y_i \;=\; A_i \, F_i\big(\ell_{i1}, \ldots, \ell_{iM}, \, x_{i1}, \ldots, x_{iN}\big), \tag{1}$$

where $A_i$ is sector $i$'s total factor productivity, $\ell_{ij}$ is labor of type $j$ employed by sector $i$, and $x_{ij}$ is good $j$ used as an intermediate input by sector $i$. Labor markets may be segmented across sectors, so factor reallocation is not automatic; the special case of a single, freely mobile labor market is nested. Market-clearing is standard: $y_i = c_i + \sum_k x_{ki}$ for each $i$, and labor markets clear at sector-specific wages.

Let $A \equiv (A_1, \ldots, A_N)$. The equilibrium map from technology to aggregate consumption is denoted $C(A)$, with steady state $\bar{C} \equiv C(1, \ldots, 1)$. I treat a

sanction – a forced switch from frontier to second-tier foundation models in the sanctioned country – as a negative Hicks-neutral shock $\Delta \log A_{\text{AI}} < 0$ to the AI services sector. Competitive equilibrium is defined in the usual way, and I rely throughout on the first welfare theorem, so that $C(A)$ coincides with the planner's solution.

### 3.2 Domar weights and Hulten's Theorem

Sector $i$'s *Domar weight* is its sales share of GDP,

$$\lambda_i \equiv \frac{p_i y_i}{\text{GDP}}. \tag{2}$$

Because intermediate-input sales are counted in every industry's revenue but enter GDP only once, Domar weights need not sum to one. Their sum,

$$\xi \equiv \sum_{i=1}^{N} \lambda_i, \tag{3}$$

is the *input-output multiplier*, and captures the degree of roundabout production in the economy. Empirically, $\xi$ is on the order of 1.7–2.0 in modern advanced economies.

The foundational result of Hulten (1978) is that, to first order,

$$\frac{d \log C}{d \log A_i} = \lambda_i. \tag{4}$$

The argument is an envelope theorem consequence of first-best efficiency: at the optimum, the marginal value product of every input is equalized across uses, so small reallocations of inputs across sectors don't change aggregate output value to first order, so that the only first-order effect of raising $A_i$ is to scale up sector $i$'s output, weighted by its sales share. Two features of Hulten's theorem are worth underlining. First, given $\lambda_i$, the underlying microeconomic structure –whether sector $i$ is a central supplier or sells mostly to final demand, whether its production function is Leontief or Cobb-Douglas, how deep the input chain beneath it runs– is irrelevant for the first-order macroeconomic impact of a productivity shock to $i$; I can compute a first-order approximation of the GDP impact of an AI productivity shock without specific knowledge of the network structure or substitution elasticities. Second, the result holds for any efficient allocation under minimal assumptions: homothetic preferences, competitive markets, and constant returns at the level of final goods.

A first order approximation, however, is unlikely to be sufficient to measure the broad macroeconomic impact of a disruption to frontier AI services. For one, the shock is potentially large; losing access to frontier models is not a marginal perturbation. Task-level estimates of the productivity gap between frontier and second-tier systems lie in the 20–40 percent range at current margins (Noy and Zhang, 2023; Brynjolfsson et al., 2023; Eloundou et al., 2024) and may be growing. At the same time, AI is likely a strong macro-complement to labor and other intermediates: the production-network literature (Atalay, 2017; Boehm et al., 2019; Barrot and Sauvagnat, 2016) repeatedly estimates short-run elasticities of substitution between intermediates well below one, often close to zero. Finally, AI is a general-purpose input. Like electricity, even when its share of total sales in the economy is small, it enters the production function of a very broad cross-section of sectors, and under complementarity a shock to a universal intermediate is amplified nonlinearly.

### 3.3 Second-Order Effects

Baqaee and Farhi (2019) characterize the second-order response of aggregate output to a sectoral productivity shock in terms of two reduced-form sufficient statistics. The first is the *macro elasticity of substitution* between goods $i$ and $j$,

$$\frac{1}{\rho_{ij}} \equiv -\frac{d\log(C_i/C_j)}{d\log A_i}, \tag{5}$$

where $C_i \equiv \partial C/\partial A_i$. Using Hulten's theorem, this is equivalent to

$$\frac{d\log(\lambda_i/\lambda_j)}{d\log A_i} = 1 - \frac{1}{\rho_{ij}}, \tag{6}$$

so $\rho_{ij}$ measures how industry $i$'s Domar weight shifts relative to $j$'s when $A_i$ changes. Goods $i$ and $j$ are *macro-complements* when $\rho_{ij} \in (0,1)$ and *macro-substitutes* when $\rho_{ij} > 1$. Note that $\rho_{ij}$ is a reduced-form general equilibrium object – a weighted average of micro elasticities of substitution, as I discuss below – and need not coincide with any single deep parameter. The second sufficient statistic is the *input-output multiplier elasticity* $d\log\xi/d\log A_i$, which captures how the economy's sales-to-GDP ratio responds to a shock at $i$. When technology shocks are labor-augmenting this elasticity is zero, but for Hicks-neutral shocks it is generically nonzero because $\xi$ itself is endogenous to the equilibrium substitution between reproducible intermediates and primary factors.

With these two objects, Theorem 2.2 in Baqaee and Farhi (2019) states:

$$\frac{d^2 \log C}{d(\log A_i)^2} \;=\; \frac{\lambda_i}{\xi} \sum_{j \neq i} \lambda_j \left(1 - \frac{1}{\rho_{ij}}\right) + \lambda_i \frac{d \log \xi}{d \log A_i}. \tag{7}$$

The second-order effect of a shock to sector $i$ is the elasticity of its own Domar weight: when a negative shock raises $\lambda_i$, which happens under macro-complementarity, the economy loses more output than Hulten's theorem predicts (negative shocks are amplified); when it lowers $\lambda_i$, under macro-substitutability, the loss is smaller (shocks are attenuated). Intuitively, Domar weights reveal through market prices how critical a sector has become along the equilibrium path. The Cobb-Douglas form is a special case under which Hulten's theorem is globally exact, so that no amount of network structure generates nonlinearity. Combining first- and second-order terms yields the local approximation:

$$\log C(A) \;\approx\; \log \bar{C} \;+\; \lambda_i \log A_i \;+\; \frac{1}{2}\left[\frac{\lambda_i}{\xi} \sum_{j \neq i} \lambda_j \left(1 - \frac{1}{\rho_{ij}}\right) + \lambda_i \frac{d \log \xi}{d \log A_i}\right] (\log A_i)^2, \tag{8}$$

Equation (8) is the second-order-accurate approximation of the log GDP change induced by a shock to sector $i$. Rather than computing the second-order term directly –which would require pinning down the full $N \times N$ matrix of $\rho_{ij}$'s and the input-output multiplier elasticity, neither of which can be read off micro data– I follow Baqaee and Farhi and adopt an equivalent but computationally lighter representation. Their Proposition 8.1 states that, to second order in the size of the shock,

$$\log\left(\frac{C(A+\Delta)}{C(A)}\right) \;=\; \frac{1}{2}\big[\lambda(A+\Delta) + \lambda(A)\big]'\big(\log(A+\Delta) - \log(A)\big) \;+\; O(\|\Delta\|^3). \tag{9}$$

In words, the log output change equals the size of the shock weighted by the simple mean of the pre- and post-shock Domar weights $\lambda(A)$ and $\lambda(A+\Delta)$ –a Törnqvist index.[4] The advantage of Equation (9) is that it repackages the second-order information in a single object, the post-shock Domar weight, $\lambda(A+\Delta)$, which is considerably more tractable than calibrating the second derivative of $\log C$ through the network.

Equation (9) is nevertheless empirically demanding, since the post-shock Domar weight is itself an endogenous general equilibrium object. Whereas Baqaee and Farhi (2019) evaluate it using observed pre- and post-shock Domar weights from

[4]The Törnqvist representation sidesteps the need to track the general equilibrium adjustment of $\lambda$ between the two endpoints, replacing it with the simple endpoint mean. The approximation is exact to second order in $\|\Delta\|$, with residual $O(\|\Delta\|^3)$.

historical data, that option is not available here because the shock takes place in a hypothetical future scenario. Instead, I obtain the post-shock Domar weight by collapsing the model economy to a two sector representation of AI and non-AI combined by a CES aggregator with single elasticity $\rho$. In this environment, Domar weights obey the two sector adding-up constraint

$$\lambda_{\text{AI}} + \lambda_{\text{non-AI}} \;=\; 1, \tag{10}$$

which amounts to imposing $\xi = 1$.[5] Under the two sector representation, the general definition of the macro elasticity in (6) specialises to a single-variable ordinary differential equation for $\lambda_{\text{AI}}$ as a function of $A_{\text{AI}}$. Setting $i = \text{AI}$ and $j = \text{non-AI}$ in (6) and substituting from (10),

$$\frac{d\log\big(\lambda_{\text{AI}}/\lambda_{\text{non-AI}}\big)}{d\log A_{\text{AI}}} \;=\; \frac{d\log\big(\lambda_{\text{AI}}/(1-\lambda_{\text{AI}})\big)}{d\log A_{\text{AI}}} \;=\; 1 - \frac{1}{\rho}. \tag{11}$$

With constant $\rho$, the right-hand side in (11) is also a constant. The left-hand side is then an exact differential in $\log A_{\text{AI}}$: the log-odds of $\lambda_{\text{AI}}$ are a linear function of $\log A_{\text{AI}}$ with slope $(1 - 1/\rho)$, so (11) integrates globally along any productivity path. Integrating both sides from $A$ to $A + \Delta$,

$$\log\!\left(\frac{\lambda_{\text{AI}}(A+\Delta)}{1-\lambda_{\text{AI}}(A+\Delta)}\right) \;-\; \log\!\left(\frac{\lambda_{\text{AI}}(A)}{1-\lambda_{\text{AI}}(A)}\right) \;=\; \left(1 - \frac{1}{\rho}\right)\Delta\log A_{\text{AI}},$$

which rearranges to the logit update

$$\log\!\left(\frac{\lambda_{\text{AI}}(A+\Delta)}{1-\lambda_{\text{AI}}(A+\Delta)}\right) \;=\; \log\!\left(\frac{\lambda_{\text{AI}}(A)}{1-\lambda_{\text{AI}}(A)}\right) + \left(1 - \frac{1}{\rho}\right)\Delta\log A_{\text{AI}}. \tag{12}$$

Equation (12) is the workhorse of the empirical exercise in Section 5: given a pre-shock Domar weight $\lambda_{\text{AI}}(A)$, a productivity shock $\Delta\log A_{\text{AI}}$, and a calibrated value of $\rho$, it delivers the post-shock Domar weight that closes Equation (9). The coefficient $(1 - 1/\rho)$ is negative under macro-complementarity ($\rho < 1$), and a negative shock to AI raises the log-odds of $\lambda_{\text{AI}}$.

# 4 Empirical Strategy

To estimate the GDP impact of the removal of access to frontier AI services on a representative economy, one requires a value for $\lambda_{\text{AI}}$, the Domar weight of

[5] Note that this does not imply an assumption that there are no intermediate inputs in the sample economy. The two sector CES is used to project $\lambda_{AI}$ forward, but the pre-shock Domar weight relies on actual data reflecting intermediate sales to downstream firms. In addition, the use of $\xi = 1$ instead of the typical $\xi \approx 1.7$–$2.0$ observed in advanced economies does not materially affect the results, because $\lambda_{\text{AI}}$ is small (in the order of 0.5 to 1 percent of GDP). The logit transformation $\lambda_{\text{AI}}/(1-\lambda_{\text{AI}})$ is dominated by $\lambda_{\text{AI}}$ itself: whether the denominator is $1 - \lambda_{\text{AI}} \approx 0.995$, or, under a general $\xi > 1$ two sector specification, $\xi - \lambda_{\text{AI}} \approx \xi$, the derivative with respect to $\log A_{\text{AI}}$ that governs the logit update below reduces to $d\log\lambda_{\text{AI}}/d\log A_{\text{AI}}$ to leading order in $\lambda_{\text{AI}}$. If, however, $\lambda_{\text{AI}}$ itself were to become very large (which might be conceivable in a more distant future where AI services compose a double-digit share of total aggregate sales in the economy) then the CES approach could become a more serious limitation.

US-origin frontier AI services in the targeted country's output. I calibrate the model to the nearest horizon where global enterprise adoption of AI services would make the sector's size economically relevant.[6] At current trends, these developments appear likely by 2028 at the earliest.

To estimate the GDP response from AI sanctions one also needs the productivity impact of losing access to US frontier providers, $\Delta log A_{AI}$. The relevant object is the technological gap between frontier models and the set of next-best substitutes, weighted by how the sanctioned economy would actually use these models. I construct two alternative estimators of this gap, drawing on different measurements of AI model capabilities as economically relevant inputs to production. The first, *Estimator A* (the headline reading), exploits the fact that frontier and non-frontier models differ most sharply not in average accuracy but in how long a coherent task they can reliably complete before failing; it combines METR's reliability curves (METR, 2025b,a) with the AEI distribution of task horizons in actual enterprise usage (Anthropic, 2026a) to map the frontier-to-substitute gap on long-horizon tasks into a productivity loss. The second, *Estimator B* (a conservative lower bound), aggregates a basket of standardized benchmarks (SWE-Bench Verified, GPQA Diamond, OSWorld, and the Epoch Capabilities Index) with weights grounded in the BLS SOC classification of the empirical AEI 1P API task composition (Handa et al., 2025).

Finally, I calibrate the two sector macro elasticity $\rho$, which governs how a shock to the AI sector propagates to the rest of the economy, from historical analogues –crude oil and electricity shortages (Baqaee and Farhi, 2019; Allcott et al., 2016)– and from the short-run micro-substitution literature on intermediate inputs (Atalay, 2017; Boehm et al., 2019), reporting a range of possible values. Substituting the resulting pre- and post-shock Domar weights into Equation (9) produces the second-order GDP loss.

## 4.1 Calibrating the pre-shock Domar weight

In an economy where AI services are entirely imported, the sector's Domar weight equals the ratio of the country's AI services imports to its GDP. I calibrate the exercise to a representative small open economy with sufficient AI adoption that, by the near-term horizon of 2028, it is likely to have embedded AI services into its economy to an extent comparable to current adoption levels in the US.[7]

Since the goal is to obtain the effect for a sample representative country, I cal-

[6] More specifically, the earliest point at which published company projections and independent market forecasts converge on Domar weights large enough to produce first-order output effects of economically meaningful size.

[7] The reason to anchor around current US embeddedness of AI services is that the productivity shock primitive estimated in Section 4.2 draws on task exposure and capability gap evidence whose underlying samples (the Anthropic Economic Index, the Brynjolfsson-Li-Raymond field studies of customer service productivity, and the Eloundou et al. task exposure index) are overwhelmingly composed of US enterprises and users.

culate the AI services Domar weights for the ten current leading AI adopters, ranked by the Anthropic Economic Index's measure (Anthropic, 2026a): Singapore, Israel, Switzerland, Canada, Australia, the Netherlands, Estonia, South Korea, New Zealand, and Sweden.[8] I use the mean of these countries' Domar weights as the representative economy's own. This construction makes the exercise representative of the broad class of advanced AI importers exposed to an access shock, but it is easy to extend to other advanced and emerging economies that are not as far along in AI adoption.

Computing $\lambda_{AI}$ for each basket member reduces to two steps. First, because AI services imports are not distinguished in national accounts or balance of payments data, I estimate the global revenues of the leading AI providers from their disclosed projections (Subsection 4.1.1). The second step is to determine what share of those revenues is sold to each country composing the representative economy, and normalize by GDP (Subsection 4.1.2).

### 4.1.1 Global Revenues of the Three Frontier AI Providers

As of April 2026, frontier services in generative AI are supplied by three US companies –OpenAI, Anthropic, and Google– each of which sells to enterprises and consumers worldwide. I project their revenues through 2028 by first determining which services count as "frontier" for the purposes of the sanctions exercise; then on what basis revenue should be measured (cumulative payments or annualized run rate), and, finally, what the available evidence has to say about the level and trajectory of these revenues for each of the three providers.

The first question is which services fall inside the sanctionable AI sector. Two definitions are relevant. A narrow definition encompasses only direct payments for frontier model inference – API access and subscription revenue flowing to OpenAI's ChatGPT and API, Anthropic's Claude, and Google's Gemini and Vertex. The narrow definition corresponds most closely to a classical intermediate-input import: a firm or household purchases access to a productive service that enters as an input into its own output.

In contrast, a broad definition also includes AI-augmented platform and software-as-a-service (SaaS) products in which frontier AI capabilities are embedded as a non-separable value driver. The dominant contributor in this category is Microsoft, through Azure OpenAI Service (enterprise access to OpenAI's GPT models via the Azure cloud platform), Microsoft 365 Copilot, and GitHub Copilot; because Microsoft bundles the three in its own disclosures, I estimate them jointly below. Other broad-definition products –Salesforce Einstein and Agentforce, Adobe Firefly, and analogous offerings– are smaller in aggregate and I exclude them from the empirical totals on data-availability grounds, a choice that makes the reported broad-definition numbers a conservative lower bound.

[8] The United Kingdom and France are excluded because both countries feature a substantial native AI development environment. Google DeepMind is headquartered in London, and Mistral AI is headquartered in Paris.

These products bundle frontier AI with platform or workflow software, but a sanctions event that cut off US AI would effectively destroy their value proposition, since the AI layer is non-separable from and foundational to the offering.[9]

I adopt the narrow definition for the baseline calibration, both because it is the more conservative choice (since it likely understates the import bill and therefore the Domar weight) and because it is the definition for which firm-level revenue data are most reliably available. The broad definition serves as a robustness check, although it is arguably the more economically relevant measure; Appendix B develops the broad-definition empirical counterpart –including a bottom-up reconstruction of Microsoft's AI revenues and a full recalculation of the Domar weight and the first-order GDP response– and is referenced throughout the main text at the points where the narrow-definition results would differ materially under the broad perimeter.

The second decision is whether to measure an AI provider's sales over a calendar year according to actual revenue or annualized run rate. Actual annual revenue is the accounting standard: the total revenue a company recognizes over a calendar year under accrual accounting. Annualized recurring revenue (ARR) is a run rate metric common in private software companies, wherein one multiplies the most recent month's revenue by twelve. Much public reporting on frontier AI company revenues relies on ARR precisely because these companies grow so fast that their year-end trajectory is more informative than their year-average. For a fast-growing company, year-end ARR will substantially exceed cumulative revenue earned during that calendar year, while at the same time understating the revenue poised to be collected over the following twelve months because the company enters the new year on a much higher base.

For present purposes, the two concepts bound the Domar weight from different directions and I report both throughout. Actual annual revenue is the correct input for a Domar weight measuring contemporaneous import exposure averaged over the calendar year. Year-end ARR is the better forward-looking indicator: because it annualizes the most recent monthly revenues, it captures the rate at which AI services are being consumed at the moment a sanction would be imposed, rather than the average over a calendar year that began on a much smaller base. I report two panels of $\lambda_{AI}$ in Section 4.1.2, one built on actual annual revenues and one built on year-end ARR.

The primary data source for what follows is Epoch AI's continuously updated revenue-reporting database (Epoch AI, 2026a), which collates disclosed and reported revenue milestones for AI companies with confidence ratings and primary source citations. I supplement this with direct company disclosures, leaked planning documents, and investigative reporting where relevant.

[9] The substitutability of, say, Microsoft 365 Copilot with a non-US AI-augmented alternative is substantially lower than a headline price comparison would suggest, because switching involves not just replacing the AI model but re-integrating an entirely new software stack into enterprise workflows.

**OpenAI.** OpenAI is the largest of the three providers by both ARR and recognized revenue. Table 1 documents the ARR trajectory from late 2023 through April 2026: the figure rose from $2.0 billion in December 2023 to $5.5 billion at the end of 2024, $21.4 billion at the end of 2025, and approximately $25 billion by April 2026, the latter implied by the company's confirmation of approximately $2 billion in monthly revenue (OpenAI, 2026a). For actual annual revenues, the best available figures are $3.7 billion for 2024 (Bloomberg, 2025) and $13 billion for 2025 (The New York Times, 2026); the H1 2025 figure of $4.3 billion in actual revenues (The Information, 2025f) confirms that the full-year $13 billion is the natural sum of rapid intra-year growth.[10]

**Table 1:** OpenAI ARR Milestones

| Date | ARR | Confidence | Source |
|---|---|---|---|
| Dec. 2023 | $2.0B | Likely | Financial Times (2024) |
| Dec. 2024 | $5.5B | Likely | CNBC (2025d) |
| Jun. 2025 | $10B | Confident | CNBC (2025d) |
| Jul. 2025 | $12B | Confident | The Information (2025d) |
| Dec. 2025 | $21.4B | Likely | The Information (2026c) |
| Feb. 2026 | $25B | Likely | The Information (2026c) |
| Apr. 2026 | ∼$25B | Likely | OpenAI (2026a) |

*Notes:* ARR = one month's revenue · 12. Confidence ratings follow Epoch AI (2026a). OpenAI's April 2026 figure is implied by the company's confirmation of approximately $2 billion in monthly revenue (OpenAI, 2026a).

OpenAI's revenue comes from three customer-facing channels: consumer subscriptions (ChatGPT Plus at $20/month, Pro at $200/month), team and enterprise plans, and direct API access. By mid-2025, API revenues alone were running at approximately $3 billion ARR, representing roughly a quarter of total ARR at the time (Financial Times, 2025). Enterprise revenue has been growing faster than consumer and was on track to reach parity with consumer revenue by end of 2026.[11] The figures above also exclude Microsoft licensing revenue (CNBC, 2025d) and reflect direct customer revenues only.[12]

[10] OpenAI's own public statement in January 2026 cited "$2B ARR in 2023, $6B in 2024, and $20B+ in 2025" as growth milestones (OpenAI, 2026a). These are ARR figures, not actual annual revenues, and should not be read as implying $6 billion was earned in 2024. The Bloomberg confirmation of $3.7 billion in 2024 actual revenue (Bloomberg, 2025) is consistent with an ARR that ended 2024 at $5.5B but averaged much lower during the year.

[11] OpenAI announced plans for an advertising product in April 2026, with internal projections of $2.5 billion in advertising revenue for 2026 (Axios, 2026). Advertising revenue is excluded throughout: it represents income from selling user attention to third parties, not payment by users or enterprises for AI access.

[12] Under a revenue-sharing arrangement, OpenAI pays Microsoft approximately 20 percent of its direct customer revenues. Leaked financial documents show Microsoft received $865.8 million in such payments over the first three quarters of 2025 (Zitron, 2025), consistent with $4.3 billion in H1 2025 revenues if the second half was larger. The payment flows in both directions: Microsoft also pays OpenAI 20 percent of Azure OpenAI Service revenues. None of this changes the customer revenue figures, which are gross of the Microsoft revenue share.

For 2026 through 2028, OpenAI's internal financial projections, as reported by The Information (2025e) from investor documents prepared in Q3 2025, imply annual growth multipliers of approximately 2.3 (2026), 2 (2027), and 1.6 (2028) from the 2025 base (Epoch AI, 2025). Applied to the $13 billion 2025 actual leads to the revenue forecasts:

$$2026 = 13 \cdot 2.3 = \$30\text{B}$$
$$2027 = 30 \cdot 2.0 = \$60\text{B}$$
$$2028 = 60 \cdot 1.6 = \$96\text{B}$$

The $96 billion 2028 figure corresponds to the total revenue target in OpenAI's investor documents, of which approximately $25 billion is expected to come from advertising (Axios, 2026). Excluding advertising, the customer revenue figure relevant to the exercise is $71 billion for 2028. A subsequently revised projection targeting $280 billion in total revenue by 2030 (Startupnews.fyi, 2026) implies upside to this figure. The corresponding year-end ARR will exceed these annual revenue figures, since the ARR captures end-of-year momentum rather than the average over the year: with OpenAI's current 1.5 ARR-to-actual revenues ratio, 2028 ARR would reach $107 billion.

**Anthropic.** Anthropic's growth trajectory has been steeper than OpenAI's, reflecting both a lower base and the mid-2025 emergence of Claude Code – which drove extraordinary enterprise adoption. Table 2 documents the full ARR timeline from January 2024 through April 2026, during which the ARR rose from $87 million to $30 billion. For actual annual revenues, The Information reported Anthropic's 2025 figure at $4.5 billion (The Information, 2026b)–a level well below the year-end $9 billion ARR, consistent with the rapid back-loaded growth seen in the ARR table. Actual 2024 revenues were approximately $0.4 billion, implied by the ARR trajectory rising from $87 million (January 2024) to $1 billion (December 2024).[13] Approximately 75–80 percent of Anthropic's revenue is enterprise- and API-based, with consumer subscriptions (Claude Pro, Claude Max) accounting for the remainder (SaaStr, 2026). This enterprise-heavy mix reflects Anthropic's positioning as an API and agentic-AI provider, and makes its revenue especially representative of the B2B AI import channel relevant to the sanctions scenario.

For 2026, Anthropic's actual revenues can be derived directly from the ARR milestones in Table 2. The implied monthly revenues are approximately $0.92 billion in January (interpolated between the December 2025 ARR of $9 billion and the February 2026 ARR of $14 billion), $1.17 billion in February, $1.58 billion in March, and $2.50 billion in April–a sum of approximately $6.2 billion for the first four months alone. For the remaining eight months, I assume monthly

[13] Assuming moderately back-loaded growth through 2024–consistent with the acceleration visible in the ARR table–average monthly revenues were roughly $33M, giving a full-year figure near $0.4B. Anthropic's own internal projections from early 2024 had placed this figure at $0.9 billion (The Information, 2025c), which proved to be an overestimate.

**Table 2:** Anthropic ARR Milestones

| Date | ARR | Confidence | Source |
|---|---|---|---|
| Jan. 2024 | $87M | Confident | The Information (2024) |
| Dec. 2024 | $1.0B | Likely | CNBC (2025c) |
| Mar. 2025 | $2.0B | Confident | CNBC (2025b) |
| May 2025 | $3.0B | Likely | CNBC (2025a) |
| Jul. 2025 | $5.0B | Confident | Anthropic (2025b) |
| Oct. 2025 | $7.0B | Confident | The Information (2025a) |
| Nov. 2025 | $1B (Claude Code only) | Confident | Anthropic (2025a) |
| Dec. 2025 | $9.0B | Likely | Bloomberg (2026c) |
| Feb. 2026 | $14B | Confident | Anthropic (2026b) |
| Mar. 2026 | $19B | Confident | Bloomberg (2026a) |
| Apr. 2026 | $30B | Confident | Bloomberg (2026b) |

*Notes:* ARR = one month's revenue · 12. November 2025 entry reports Claude Code's ARR separately within the total. Confidence ratings follow Epoch AI (2026a). Anthropic's February 2026 disclosure noted that "run-rate revenue is $14 billion, with this figure growing over 10x annually in each of those past three years" (Anthropic, 2026b). *Accounting note:* OpenAI's Chief Revenue Officer alleged in an April 2026 internal memo that the $30B figure reflects gross accounting of revenue-share flows from Amazon and Google, and that a comparable net figure would be approximately $22B (Heath, 2026). See Section **??** for discussion of why the gross figure is the appropriate input for this paper's exercise.

revenues grow linearly from the April level to a December level consistent with the year-end 2026 ARR of $90 billion (Panel A), implying a December monthly run rate of $7.50 billion. Under linear interpolation, May through December contribute approximately $42.5 billion, giving a full-year 2026 actual of approximately $49 billion, which I round to 50 billion.[14]

For 2027 and 2028, I project Anthropic's actual revenues by applying OpenAI's disclosed year-on-year growth multipliers to the 2026 baseline. OpenAI projects approximately 2.0× revenue growth from 2026 to 2027 and 1.6× from 2027 to 2028 (The Information, 2025e). These multipliers represent the best available market-level deceleration profile: both companies operate in the same frontier AI segment, face similar enterprise sales dynamics, and are growing through identical product categories (API, consumer subscriptions, agentic coding). In the absence of updated Anthropic-specific projections–the November 2025 company documents (The Information, 2025b; TechCrunch, 2025) were calibrated to a $4.7 billion ARR and are thoroughly obsolete at the current $30 billion level–OpenAI's growth trajectory serves as the natural industry proxy. Applying these rates to the $50 billion 2026 estimate yields actual revenues of $100 billion in 2027 and $160 billion in 2028. Year-end ARR for 2027 and 2028

[14] Anthropic's own 2026 projection of $18 billion (The Information, 2026a), calibrated when ARR was $4.7 billion, is by now obsolete. Even under the conservative assumption that monthly revenues remained flat at April 2026 levels ($2.50 billion per month) for the rest of the year, the full-year actual cannot fall below approximately $26 billion.

is then derived using the same linear-interpolation formula employed for 2026 (actual ≈ (start-ARR + end-ARR)/2). Inverting this relation gives $110 billion at year-end 2027 and $210 billion at year-end 2028. The implied ARR-to-actual ratio falls from 1.8× in 2026 to 1.1× in 2027 and 1.3× in 2028, consistent with the mathematical consequence of decelerating intra-year growth: the year-end ARR exceeds the annual average by a diminishing margin as the monthly run rate flattens.

**Google (Gemini/Vertex AI).** Google does not separately disclose revenues from Gemini API, Vertex AI generative inference, or Google AI Studio; these are bundled within Google Cloud. This is independently confirmed by the 2026 Stanford AI Index Report, which tracks annualized revenues for all major frontier providers using the same Epoch AI data relied on throughout this paper, yet reports no standalone figure for Google Gemini or Vertex AI–precisely because no such disclosure exists (Sajadieh et al., 2026). Google Cloud reported full-year 2025 revenues of approximately $43 billion and reached an annual run rate exceeding $70 billion by Q4 2025 (Alphabet, 2026).[15]

I estimate Google's external Gemini/Vertex AI inference revenues for 2025 using two complementary approaches. The first is a cross-check via the enterprise API market-share data of Menlo Ventures (2025), which place Google at 21 percent of the enterprise AI model API market. To estimate the total market size I use the fact that Anthropic discloses 75–80 percent of its revenues as coming from enterprise and API channels (SaaStr, 2026); at Anthropic's 40 percent share of that market, this implies a total enterprise API market of approximately $8.8 billion in 2025 and a Google share of $1.8 billion. This estimate is a lower bound, since the Menlo survey covers direct enterprise API usage and does not capture consumer subscriptions.[16] The second approach uses management commentary as a ceiling. On its Q4 2025 earnings call, Google reported that AI-related Cloud revenues had reached "billions of dollars per quarter" (Alphabet, 2026)–implying at minimum $8 billion annually across all AI-related Cloud workloads. External-facing generative inference for customers (as opposed to internal tooling, ML training infrastructure, and non-generative AI products) is a significant but non-exclusive fraction of that aggregate. Discounting the

[15] Q4 2025 year-on-year growth was 48 percent. Management stated on the Q4 2025 earnings call that AI-related revenues had reached a scale of "billions of dollars per quarter" and that products built on generative models grew more than 200 percent year-on-year (Alphabet, 2026). Q1 2026 results are scheduled to be reported on April 29, 2026, after this draft was prepared.

[16] Google sells Gemini Advanced access bundled into its Google One AI Premium tier ($20/month), and even modest subscriber counts–say, 5–15 million paying globally–imply $1–3 billion in annual subscription revenue, a material fraction of which is attributable to AI access. Independent industry modelling places 2025 consumer subscription revenue from this tier at approximately $1.2 billion (Business of Apps, 2026), consistent with roughly 5–6 million AI Premium subscribers and landing at the conservative lower end of this range. Google does not disclose this tier separately in its filings; the estimate is compiled from earnings commentary and subscriber growth disclosures. This figure does not alter the central Google estimate, which already embeds consumer subscription revenue alongside enterprise API revenues.

management figure conservatively–attributing one-third to one-half of Cloud AI revenues to external generative inference–yields a range of $2.7–4 billion, broadly consistent with the lower-bound estimate, especially if consumer subscriptions are added to the Menlo baseline. Combining the two approaches gives a 2025 range of approximately $1.8–4 billion, with a midpoint of $3 billion that I carry forward as the central estimate. This is the least reliable of the three providers' figures: absent Google disclosing this line separately, any value in the $1.5–5 billion range is difficult to rule out on available evidence.

For 2026 through 2028, no equivalent internal projection is available for Google, and I derive forward estimates using a constant-market-share approach. In 2025, Google's $3 billion in Gemini/Vertex AI revenues represents approximately 14.6 percent of the $20.5 billion narrow-definition total. I project this share at a stable 13 percent going forward–a slight downward adjustment from 14.6 percent reflecting faster growth in Anthropic's and OpenAI's non-API revenue channels (consumer subscriptions, agentic coding products), where Google has weaker positioning.[17] Denoting Google's estimate $G$ and the combined OpenAI–Anthropic estimate $X$, the condition $G = 0.13(G + X)$ gives $G = (0.13/0.87)\, X = 0.149\, X$. Applying this to the OpenAI–Anthropic combined projections yields Google actuals of $12 billion (2026), $24 billion (2027), and $35 billion (2028). For year-end ARR I apply a 1.5× ARR-to-actual multiplier to these projections; the 1.5× figure is lower than the 2× used for Anthropic, reflecting Google Cloud's more mature and stable revenue base, and is consistent with the 2025 observation (Google actual $3B, year-end ARR $6B = 2×) decelerating as the business stabilises. This delivers year-end ARRs of $18 billion (2026), $36 billion (2027), and $53 billion (2028). These figures are inherently uncertain given Google's non-disclosure, and the sensitivity results in Appendix B confirm that the main conclusions are robust to Google being set to zero.

Table 3 collects the full revenue trajectory across both metrics and all three providers, from observed 2024–2025 actuals through the 2028 projections. The two panels report the same underlying revenue base under the two complementary measurement conventions described above: Panel A in year-end ARR terms and Panel B in full-year actual revenues. By 2028, the narrow-definition global revenues of the three providers reach $370 billion in year-end ARR and $266 billion in cumulative customer revenues over the calendar year. These are the global totals to which the geographic decomposition of Section 4.1.2 is applied. Appendix B develops the broad-definition counterpart, which augments these totals with an independent bottom-up estimate of Microsoft's AI revenues.

### 4.1.2 AI's Domar weight across the basket of countries

The global revenues constructed in Table 3 are payments from customers in every country in which the three providers sell. To recover each basket economy's

[17] Note that the Menlo Ventures enterprise API share of 21 percent (Menlo Ventures, 2025) applies to the direct API segment only; Google's share of the broader total including consumer and agentic revenues is lower, as the 2025 observed figure of 14.6 percent confirms.

**Table 3:** Frontier AI Services Revenue: Actual and ARR, 2024–2028

| | 2024 | 2025 | 2026 (est.) | 2027 (proj.) | 2028 (proj.) |
|---|---|---|---|---|---|
| *Panel A: Year-end ARR ($B, central estimates)* | | | | | |
| OpenAI (ex-ads) | 5.5 | 21 | 42 | 80 | 107 |
| Anthropic | 1.0 | 9 | 90 | 110 | 210 |
| Google (Gemini) | n/a | 6 | 18 | 36 | 53 |
| **Narrow total** | **∼6.5** | **∼36** | **∼150** | **∼226** | **∼370** |
| *Panel B: Full-year actual customer revenues ($B, central estimates)* | | | | | |
| OpenAI (ex-ads) | 3.7 | 13 | 30 | 60 | 71 |
| Anthropic | 0.4 | 4.5 | 50 | 100 | 160 |
| Google (Gemini) | 1.5 | 3 | 12 | 24 | 35 |
| **Narrow total** | **5.6** | **20.5** | **92** | **184** | **266** |

*Notes:* Panel A reports annualized run rates at year-end (one month's revenue · 12); Panel B reports cumulative customer revenues during each calendar year. Both panels exclude advertising revenue and investor funding. 2024–2025 Panel B figures are sourced from Bloomberg (2025), The New York Times (2026), and The Information (2026b). 2026–2028 Panel B projections: OpenAI from The Information (2025e) (ex-advertising); Anthropic 2026 derived by integrating monthly revenues implied by ARR milestones through April 2026 and linearly interpolating to a December run rate consistent with the $90B year-end ARR; Anthropic 2027–2028 by applying OpenAI's disclosed growth multipliers (2.0×, 1.6×) to the $50B 2026 baseline; Google by applying a stable 13% market share to the OpenAI–Anthropic combined total. Panel A Anthropic ARR: 2026 = $90B (Gerstner, 2026); 2027 and 2028 derived by inverting the linear interpolation formula (actual ≈ (start-ARR + end-ARR) /2): 2027 = $110B; 2028 = $210B. Panel A Google ARR: 1.5× the corresponding Panel B figure. The two panels bound the Domar weight calibration: Panel B is the conservative, national-accounts-consistent estimate; Panel A is the forward-looking indicator. Appendix B develops and reports a broad-definition counterpart that augments these totals with Microsoft's AI revenues (Azure OpenAI Service, Microsoft 365 Copilot, and GitHub Copilot).

import bill I allocate those global totals in two steps. I first remove the US domestic share, since revenues recognized inside the United States are not imports from any other country's perspective. I then allocate the non-US remainder across regions, and within each region across countries, using the geographic structure of provider revenues and the cross-country distribution of services imports. The resulting share of global revenue, applied to the totals of Table 3 and divided by each economy's GDP, delivers its Domar weight.

Although none of the three frontier AI providers publishes a geographic breakdown of their revenues, three independent sources provide consistent estimates of the US share. First is Alphabet, Google's parent company: its 2025 annual filing places the United States at 48 percent of total revenues (Alphabet, 2026). Second, Forrester Research estimates that the United States accounted for 46 percent of the global enterprise AI software market in 2024 (Forrester Research, 2024). Third, OpenRouter's 2025 State of AI report places the United States at

47 percent of global LLM token volume (OpenRouter, 2025). I take the median of 47 percent as the estimate of the US domestic share of global narrow AI revenues, assigning the remaining 53 percent to the share going to international customers, from which the representative economy's import bill must be drawn.

To calibrate the Domar weight for each country in the composing basket used to construct the representative economy, I use the four geographic segments reported in Alphabet's 2025 annual filing –United States, EMEA, APAC, and Other Americas– which account for 48, 29, 17, and 6 percent of total revenues, respectively (Alphabet, 2026). Expressed as shares of the non-US total, the three relevant segments are EMEA (56 percent), APAC (33 percent), and Other Americas (12 percent). I apply Alphabet's geography uniformly across the three providers as the best available proxy for the regional distribution of frontier AI revenue.

Within each region, I allocate revenue across countries in proportion to their services imports, which proxy the cross-country distribution of AI services purchases. Data on services is sourced from the World Bank's balance of payments series (World Bank, 2024).

For each basket member $c$, the Domar weight is therefore

$$\lambda_{AI,c} \;=\; \frac{0.53 \cdot \sigma_{\text{region}(c)} \cdot s_c \cdot R^{\text{global}}_{2028}}{\text{GDP}_c},$$

where $\sigma_{\text{region}(c)}$ is its region's share of non-US revenue, $s_c$ its share of that region's services imports, and $R^{\text{global}}_{2028}$ the global narrow-definition total from Table 3. Averaging equally across the ten members yields $\lambda_{AI} = 0.57$ percent under Panel A (year-end ARR, \$370B) and 0.41 percent under Panel B (full-year actual revenues, \$266B). Table 4 summarizes the calculation; the country-by-country detail appears in Appendix A. The two panels bound the Domar weight under the narrow definition — Panel B the conservative, national-accounts-consistent measure, Panel A the forward-looking indicator — and both are carried through to the GDP-response calculations of Section 5. The broad-definition counterpart, which adds Microsoft's AI revenues to the perimeter, is developed in Appendix B.

## 4.2 Calibrating the AI Productivity Shock $\Delta \log A_{\mathbf{AI}}$

The Törnqvist–Hulten representation in equation (9) shows that the first-order GDP effect of a sanction is the product of the AI sector's Domar weight and the Hicks-neutral productivity change $\Delta \log A_{\text{AI}}$, resulting from switching away from frontier AI models and into second-best options. Having pinned down $\lambda_{AI}$ above, this section focuses on calibrating the second primitive.

The set of second-best AI models under the sanctions scenario includes non-US closed and open-weight models (DeepSeek V4, Qwen 3.5, GLM-5.1, Mistral Large 3, Kimi K2.5) plus US-origin open-weight releases (the Llama family) whose weights have been publicly distributed.

**Table 4:** Domar weight of AI Services for the Representative Basket by 2028

| | Panel A (ARR) | Panel B (actual) |
|---|---|---|
| Global narrow AI revenue ($B) | 370 | 266 |
| $\times (1 - 0.47)$: non-US ($B) | 196.1 | 141.0 |
| Allocated across the ten basket economies and averaged equally (Appendix Table 9) | | |
| **Domar weight $\lambda_{AI}$** | **0.57%** | **0.41%** |

*Notes:* $\lambda_{AI}$ is the equal-weighted mean of the per-country Domar weights of the ten highest-ranked AI adopters outside the United States (by Anthropic Economic Index per-capita usage), excluding the United Kingdom and France. Each country's weight is its frontier AI import share of GDP, built from the global revenue totals of Table 3, the non-US share (0.53), Alphabet's regional revenue shares, World Bank within-region services-import shares, and IMF 2026 GDP. Panel B is the conservative, national-accounts-consistent measure; Panel A the forward-looking indicator. Per-country detail in Appendix A.

The productivity loss to second-best alternatives can vary significantly, depending on the weight one assigns to task difficulty and the relative importance of knowledge reasoning versus long-horizon agentic capabilities. I construct two conceptually distinct estimators: Estimator A weights the AI sector productivity shock toward agentic and long-horizon task completion measures, which most directly reflect AI as an input for production; this is the baseline. Estimator B relies on the capability gap from a basket of standardized AI benchmarks for a fixed set of tasks, and acts as a conservative lower bound. Estimator A is most applicable when the long right tail of task difficulty is the binding margin in production, whereas Estimator B is more appropriate when average task capability is especially relevant. Both arrive at the productivity shock from AI sanctions by measuring a capability ratio –the fraction of frontier AI capabilities that is retained by the second-best models– denoted by $\kappa$. As will be shown below, the productivity shock is given by: $\Delta \log A_{\text{AI}} = \log \kappa$.

To calibrate the estimators' task weights I rely on data from the Anthropic Economic Index (Handa et al., 2025; Anthropic, 2026a), which classifies a representative sample of Anthropic's API queries into occupational tasks from the O*NET database, the US Department of Labor's standard taxonomy of work activities (O*NET Resource Center, 2024). For each task, the index reports how many conversations fall within its category, the time required for a skilled human to complete it, and a Standard Occupational Classification (SOC) code (U.S. Bureau of Labor Statistics, 2018). Estimator A draws on the distribution of completion times; Estimator B draws on the task composition across SOC groups.

#### 4.2.1 Estimator A: task horizon coverage.

Estimator A measures the productivity shock from data on the length of tasks that each AI model can effectively complete in production. I refer to the frac-

tion of task hours an AI model can reliably deliver as its coverage. A model's capability ratio $\kappa$ is then the ratio of the leading second-best model's coverage to the US frontier's, and $\log \kappa$ is the Hicks-neutral productivity shock to the AI sector.

The estimator has two primitives: the reliability curve (how good an AI model is at each task length) and the distribution of task lengths in the economy (the share of AI-handled work-hours at each task length). A model's coverage is therefore given by the task hour-weighted average of its reliability curve.

Denote the reliability curve as $P(\text{success} \mid t; M)$, the probability that model $M$ successfully completes a task of length $t$, where $t$ is the time needed for a skilled human to complete it. For frontier models, METR (2025b) estimates twos separate marks: the 50-percent time horizon $h_{50}$ (where the model's success probability is one-half) and the 80-percent time horizon $h_{80}$ (where it is 0.8). Following the METR reporting convention, I assume a task length function of the form:

$$P(\text{success} \mid t; M) \;=\; \frac{1}{1+\exp\big(a(\log t - b)\big)}, \tag{13}$$

where $b = \log h_{50}$ fixes the task length at which the model succeeds half the time, and $a = \log(0.25)/\log(h_{80}/h_{50})$ determines how steeply success rates decrease with task length, and is determined by requiring $P(h_{80}) = 0.80$. Two measured horizons uniquely identify the curve. For April 2026, I use the median of METR Time-Horizon 1.1 scores, setting $h_{50}^F = 6.40$ hours and $h_{80}^F = 1.16$ hours as the US frontier (see Appendix C for the full sensitivity to this choice), and placing the second-best model capabilities at 2.27 METR-doublings behind (METR, 2026), which leads to $h_{50}^S = 1.33$ and $h_{80}^S = 0.24$.[18]

Next is the economy's distribution of task lengths. For each task category $c$ in the O*NET database, the AEI reports a conversation count, $n_c$, and a mean human completion time for that task, $\bar{t}_c$. The product $n_c \cdot \bar{t}_c$ is a measure of the economic size of that task category within total AI usage. A model's success probability across all task categories, or coverage, is obtained by weighting its reliability curve using this task-length distribution:

$$\theta(M) = \frac{\sum_c n_c\, \bar{t}_c\, P(\text{success} \mid \bar{t}_c; M)}{\sum_c n_c\, \bar{t}_c}$$

Coverage in frontier and second-best models is, respectively, $\theta(F)$ and $\theta(S)$. The change in productivity in the AI sector resulting from forcibly switching to

[18]The 2.27-doubling baseline is the most recent date-matched METR comparison, taken *within* a single METR version (v1.0) to avoid mixing scales: Claude Opus 4.5 ($h_{50}$ = 277.7 minutes) against Kimi K2 Thinking ($h_{50}$ = 57.6 minutes), both November 2025, a capability ratio of 0.207 corresponding to $\log_2(4.82) = 2.27$ doublings. (An earlier reading of 2.35 inadvertently paired the v1.1 Opus figure, 293 minutes, with the v1.0 Kimi figure.) The April 2026 Chinese releases (Kimi K2.5/K2.6, DeepSeek V4-Pro, GLM-5.1) have not been METR-scored on the task-length suite, so the central uses only what METR has directly measured.

second-best models is given by:

$$\Delta \log A_{AI} = \log(A_S) - \log(A_F) = \log \theta(S) - \log \theta(F)$$
$$\Delta \log A_{\mathrm{AI}} = \log \kappa$$

Where $\kappa \equiv \theta(S)/\theta(F)$ is the ratio of model coverages. Applying the April-2026 METR-anchored reliability curves to the AEI 1P API task-hour distribution yields $\theta_F = 0.684$ and $\theta_S = 0.398$, such that $kappa = 0.582$. The productivity impact of AI sanctions according to Estimator A is therefore:

$$\Delta \log A^{A}_{\mathrm{AI}} = -0.54$$

This very meaningful decrease is a direct result of the shape of the AEI task distribution, which places a substantial share of task-hour weight in the one-to-ten-hour band, precisely the window in which the reliability of second-best models diminishes fastest compared to the frontier curve.

For additional robustness, I bracket Estimator A's result within a confidence band given by alternative possible outcomes for the capability gap between frontier and second-tier models. On the one hand, it's possible that second-tier models will exhibit more rapid convergence to the frontier – the Epoch Capabilities Index places the recent capability distance at roughly six months (Epoch AI, 2026b). With a corresponding 1.39-doubling lag, such a scenario would yield $\Delta \log A = -0.29$. On the other hand, divergence between the frontier and its competitors is also possible; recent statements by Anthropic representatives, for instance, place Chinese models up to twelve months behind the Mythos iteration. This estimate would imply a 2.56-doubling lag, leading to $\Delta \log A = -0.63$. Within this convergence-to-divergence spectrum, the range for the AI sector's productivity loss is:

$$\Delta \log A^{A}_{\mathrm{AI}} \in [-0.63, -0.29], \qquad \text{central value } -0.54.$$

#### 4.2.2 Estimator B: task mix weighted capability gap.

As an alternative measure for the productivity loss of switching to second-tier models, I construct an estimator based on unit-task benchmarks. This captures a scenario where AI as an input to production is mostly used at short horizons, rather than as a substitute for multi-hour blocks of skilled labor. In the remainder of the paper, I retain the results from Estimator B as a conservative lower bound.[19]

If the best non-US system delivers a fraction $\kappa_B \in (0, 1)$ of the frontier's task completion value per query, then the output of the AI services sector per unit

[19] There are two reasons why Estimator B is more conservative than Estimator A. First, most unit-task benchmarks are short-horizon instruments whose success ratio gap systematically understates the economic gap for an input that is consumed in larger-than-unit blocks. Second, these benchmarks are rapidly approaching saturation as the capabilities of AI models exceed beyond originally planned measures.

of input falls by $\log \kappa_B$. I rely on a basket of four benchmarks that are common in the literature: SWE-Bench Verified (Jimenez et al., 2024) as a coding proxy, GPQA Diamond (Rein et al., 2023) as a science reasoning proxy, OSWorld (Xie et al., 2024) as an agentic tool proxy, and the Epoch Capabilities Index (Epoch AI, 2026c) as a generalist composite. As in Estimator A, the goal is to construct a coverage ratio. In this case, we have a gap for each benchmark: the fraction of the frontier's capability that can be reached by the second-tier models. A ratio of 1.0 means parity. To arrive at a single coverage ratio, I take a weighted average of the four benchmark gaps.

The benchmark gap scores are obtained directly from the primary public sources. The weights are derived from the AEI task composition, by mapping each SOC task group to the benchmark whose published scope most directly covers it: SOC 15 (Computer and Mathematical) to SWE-Bench, SOC 19 (Life, Physical, and Social Science) to GPQA, SOC 43 (Office and Administrative Support) to OSWorld, and every other major group to the Epoch Capabilities Index as the generalist residual. Applying this mapping to the 2026 AEI data and weighting by conversation count, SOC 15 accounts for 62.6 percent of API traffic, SOC 19 for 4.4 percent, SOC 43 for 7.4 percent, and the residual of 25.5 percent falls into the generalist (Epoch Capabilities Index) bucket. These are the weights used for each score's contribution to an AI model's capability. Because the benchmarks are close to saturation, I measure each gap not on the raw success scores, but on the tasks that are still failed: the ratio of the frontier's error rate to the substitute's, $(1 - p_F)/(1 - p_S)$. This puts the comparison where the economic value of frontier capability concentrates, on the difficult tasks at the margin of what models can do (Eloundou et al., 2024; METR, 2025b). It equals one at parity and falls as the second-tier substitute fails relatively more. The weighted error-rate gaps give a coverage ratio of:

$$\kappa_B \;=\; 0.626 \cdot 0.708 \;+\; 0.044 \cdot 0.640 \;+\; 0.074 \cdot 0.760 \;+\; 0.255 \cdot 0.952 \;=\; 0.77$$

where the generalist composite, not being a pass/fail benchmark, enters at its index ratio. Estimator B's productivity shock is therefore:

$$\Delta \log A_{\mathrm{AI}}^{(B)} = \log(0.77) = -0.26.$$

Table 5 collects the productivity shock calibrations from Estimators A and B. Note that, because they measure different margins, they cannot be pooled: Estimator A measures the gap at the task-hour distribution of actual API usage; Estimator B measures the gap at the average task, with task-mix weights grounded in the BLS SOC classification of the AEI task composition. Estimator A is this paper's headline reading because AI as input to production is, at the current margin, consumed in multi-hour blocks of skilled-labor-equivalent work.

### 4.3 Calibrating AI's post-shock Domar weight

In order to compute the second-order effects of AI sanctions using the Törnqvist–Hulten representation in equation (9), one needs the value of the post-shock

**Table 5:** Estimates of $\Delta \log A_{\text{AI}}$ under a Frontier Model Sanction

| Estimator | Convergence | Central | Divergence |
|---|---|---|---|
| A – Task-horizon coverage | −0.29 | −0.54 | −0.63 |
| B – Task-mix-weighted capability gap | | −0.26 | |

*Note:* Estimator A combines the logistic METR reliability curve with the AEI 1P API task-hour distribution; its three values are the convergence (1.39-doubling, Epoch ECI), central (2.27-doubling, measured), and divergence (2.56-doubling, Amodei (2026)) lags. Estimator B aggregates four standardized benchmarks (SWE-Bench Verified, GPQA Diamond, OSWorld, and the Epoch Capabilities Index) with weights grounded in the BLS SOC classification of the AEI 1P API task composition, each benchmark's gap measured on its error rate.

Domar weight of the AI sector. Unlike the pre-shock weight, which can be derived from revenue data, the post-shock weight is an object that must be inferred. Recall from section 3.3 that it is possible to obtain this post-shock weight from the original Domar weight and the macro elasticity of substitution $\rho_{ij}$ (the ability of the targeted economy to switch from AI services to alternative intermediates into production). What's needed, therefore, is a calibration for the $\rho_{ij}$ parameter.

I draw on two separate strands of evidence in the literature for the plausible range of the macro elasticity. First, the structural exercise in Baqaee and Farhi (2019) which finds $\rho = 0.4$ during the 1970s oil shock. Oil is an input with few immediate substitutes whose marginal product is diffused across the economy through intermediate use, making it an adequate comparator for the specific case of AI services as an intermediate input. Second, the micro literature on short-run input substitution –Atalay (2017) on US manufacturing industry-level production functions, Boehm et al. (2019) on the Tohoku earthquake propagating through Japanese multinational supply chains, and Allcott et al. (2016) on the 2012–2016 Indian electricity crisis– consistently estimates intermediate-input elasticities clustering in the range 0.1–0.3 at horizons of one to three years.[20] I set a baseline of $\rho = 0.4$ and an envelope $\rho \in [0.2, 0.6]$ that spans short-run micro and medium-run structural benchmarks. Note that values of $\rho$ closer to unity would collapse equation (12) to $\lambda(A + \Delta) = \lambda(A)$ and return to Hulten's floor; values below 0.2 imply Leontief-like complementarity, which seems unlikely given the existence of pre-AI workflows.

Applying equation (12) to the four central cases in Section 4.2 and to the endpoint and central calibrations for $\rho$ produces a range of post-shock AI Domar weights, summarized in Table 6.

[20] The structural macro literature on non-homothetic demand and long-run reallocation (Comin et al., 2021) finds elasticities that can be higher, closer to 0.6–1.0, once firms and households have had a decade or more to adjust.

**Table 6:** Post-shock Domar weights (in percent of GDP)

| | $\rho = 0.6$ | $\rho = 0.4$ (baseline) | $\rho = 0.2$ |
|---|---|---|---|
| Panel A (ARR) × Estimator A ($\Delta = -0.54$) | 0.81 | 1.27 | 4.74 |
| Panel A (ARR) × Estimator B ($\Delta = -0.26$) | 0.68 | 0.84 | 1.60 |
| Panel B (actual) × Estimator A ($\Delta = -0.54$) | 0.59 | 0.92 | 3.45 |
| Panel B (actual) × Estimator B ($\Delta = -0.26$) | 0.49 | 0.60 | 1.15 |

*Note:* Results calibrated using central estimates for $\Delta \log A_{AI}$.

# 5 Results: The GDP Impact of AI Sanctions

Drawing on the estimated Domar weights of the AI sector and the size of the productivity shock, I now estimate the impact of AI sanctions on the sample economy's output. I proceed in three steps. First, subsection 5.1 computes the first-order Hulten effect, obtained by the pre-shock Domar weight and the productivity shock before any general equilibrium sectoral reweighting. The subsection that follows computes the Baqaee-Farhi second-order impact. The final subsection then collects the full battery of results across the two Domar weight panels and the two estimators of $\Delta \log A_{\text{AI}}$, decomposes the headline into the contributions of each primitive, and discusses robustness.

## 5.1 First-Order Effect

Applying equation (4) to the two primitives constructed in 4.1.2 and 4.2 yields the first-order results collected in Table 7, below. Note that the Domar weight enters with two values: $\lambda_{AI} = 0.57$ percent when the 2028 AI sector is measured at year-end annualized run rates (Panel A) and $\lambda_{AI} = 0.41$ percent when it is measured at full-year actual revenues (Panel B). The productivity shock enters with three possible values (convergence, central, and divergence) for Estimator A and a single value for Estimator B. The resulting eight-cell grid traces the full envelope of the first-order response of GDP of the target economy to AI sanctions.

**Table 7:** First-order GDP response, in percent of GDP

| | Estimator A | | | Estimator B |
|---|---|---|---|---|
| | Conv. | Central | Div. | |
| $\Delta \log A_{\text{AI}}$ | −0.29 | −0.54 | −0.63 | −0.26 |
| Panel A ($\lambda_{AI} = 0.57\%$) | −0.17 | −0.31 | −0.36 | −0.15 |
| Panel B ($\lambda_{AI} = 0.41\%$) | −0.12 | −0.22 | −0.26 | −0.11 |

*Note:* Cells report $\Delta \log Y^{(1)} = \lambda_{AI} \cdot \Delta \log A_{\text{AI}}$. Panel A uses year-end ARR for the Domar weight; Panel B uses full-year actual revenues (see Table 4). Estimator A is the task-horizon coverage estimator (the headline reading); Estimator B is the task-mix-weighted capability gap.

Two features of Table 7 set the stage for the second-order analysis that follows. First, the first-order magnitude is modest. The central readings span −0.11 to −0.31 percent of GDP.[21] Second, the variance across the grid is large. Most of this differences stem from the two estimators of the productivity shock, which are measuring different margins of the task-difficulty distribution. Recall, however, that the preferred estimate of the productivity shock is the one obtained in Estimator A, given the preeminence of long duration AI tasks as a defining characteristic of its utility for complex production processes.

## 5.2 Second-Order Effects

Bringing the previously calibrated values for $\lambda(A)$, $\lambda(A+\Delta)$, and $\Delta \log A$ into equation (9) leads to the the second-order results displayed in Table 8, below. Results are provided for each pairwise combination of data panels (depending on which revenue stream is assumed for frontier AI services, underlying their initial Domar weight) and each Estimator for the resulting productivity loss in the AI sector once sanctions are imposed.

**Table 8:** Total GDP loss under the Törnqvist approximation (in percent of GDP).

| | $\rho = 0.6$ | $\rho = 0.4$ | $\rho = 0.2$ |
|---|---|---|---|
| Panel A × Estimator A | −0.37 | −0.50 | −1.43 |
| Panel A × Estimator B | −0.16 | −0.18 | −0.28 |
| Panel B × Estimator A | −0.27 | −0.36 | −1.04 |
| Panel B × Estimator B | −0.12 | −0.13 | −0.20 |

Note that the amplification of downstream industry effects from the shock –the Törnqvist total divided by the Hulten first-order floor reported in Table 7– is governed almost entirely by $\rho$ and $\Delta \log A$ as long as the economy is in a small-Domar-weight regime; that is, for pre-shock levels of $\lambda$ below a threshold level of approximately 5%.[22] This amplification is particularly high (approaching a fivefold increase) for $\rho = 0.2$, when sectoral production gets closer to Leontief. This exceeds the nearly fourfold amplification that Baqaee and Farhi (2019) document for the 1973 oil shock, despite the starting Domar weight being much

[21] This is unsurprising. Baqaee and Farhi (2019) report that the first-order effect of the 1970s oil shock delivered an output loss of about 0.6 percent.

[22] The Hulten first-order floor is $\lambda\Delta$. The Törnqvist total is $\frac{1}{2}(\lambda + \lambda')\Delta$. The amplification ratio is therefore

$$\text{amp} \;=\; \frac{\frac{1}{2}(\lambda + \lambda')\,\Delta}{\lambda\,\Delta} \;=\; \tfrac{1}{2}\Big(1 + \tfrac{\lambda'}{\lambda}\Big)$$

From equation (12), $\lambda' = \sigma(\text{logit}\,\lambda + (1 - 1/\rho)\Delta)$. When $\lambda$ is small, I have:

$$\log \lambda' \;\approx\; \log \lambda + (1 - 1/\rho)\Delta \;\Longrightarrow\; \frac{\lambda'}{\lambda} \;\approx\; e^{(1-1/\rho)\Delta},$$

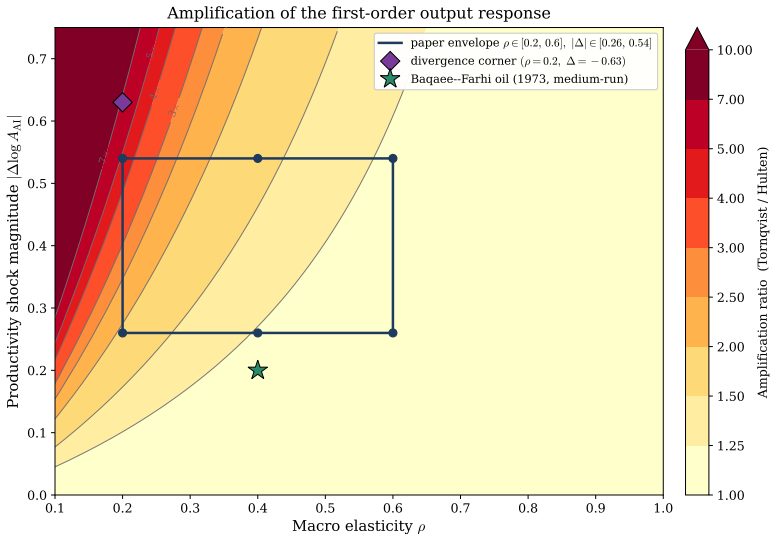
which doesn't depend on $\lambda$. Plugging back,

$$\text{amp} \;\approx\; \tfrac{1}{2}\Big(1 + e^{(1-1/\rho)\,\Delta}\Big)$$

smaller than oil's five-percent share at the time, because the capability gap ($\Delta \log A$ at Estimator A's central) is meaningfully larger than the price change that scales the calibration of oil.

Figure 3 visualises the amplification ratio over the full $(\rho, \Delta)$ plane and locates the envelope of results inside it. The contours fan out sharply when $\rho$ drops below 0.3 and $|\Delta|$ exceeds 0.4; at $\rho = 0.4$ and $|\Delta| = 0.20$, which is roughly where Baqaee and Farhi (2019)'s medium-run oil counterfactual sits.

**Figure 3:** Amplification ratio of the Törnqvist second-order response relative to the Hulten first-order response.



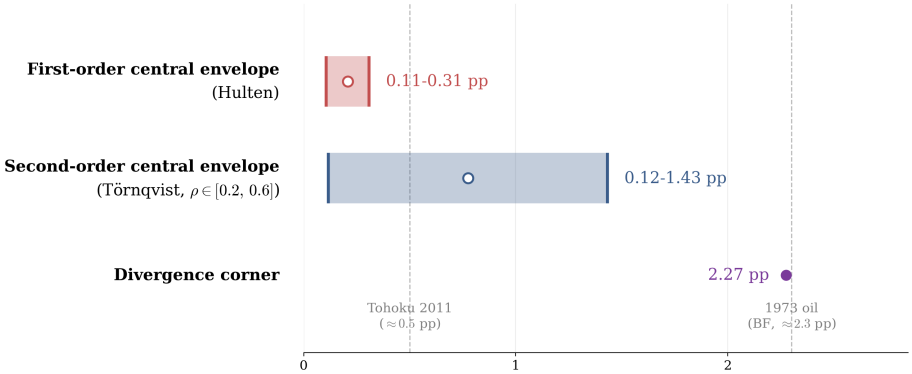

*Note:* Ratios plotted over the parameter space $(\rho, |\Delta \log A_{\text{AI}}|)$. The rectangle marks the paper's central envelope $\rho \in [0.2, 0.6] \times |\Delta| \in [0.26, 0.54]$, spanning Estimator B's reading up to Estimator A's central. The six navy dots on the edges of the box are the central cases tabulated in Table 8; the purple diamond is the divergence corner at $\rho = 0.2$, $\Delta = -0.63$; the green star marks Baqaee and Farhi (2019)'s medium-run oil calibration for reference.

Another important consideration is that the baseline scenario assumes a sanctioned economy that has reached levels of AI embeddedness into its production broadly comparable to current US levels, which I project around 2028, but $\Delta \log A$ is estimated from the technological gap between frontier US models and open-weight alternatives from current studies. If future developments lead to an even wider AI gap, then a 2028 shock could well be larger than Estimator A's central calibration. The divergence column at $\Delta = -0.63$ partly spans this possibility.

## 5.3 The Full GDP Impact of AI Sanctions

I now bring all values from the preceding sections together into a final range of results, shown on a common axis in Figure 4. The first-order effect provides a floor on the impact of AI sanctions, and the addition of second-order effects brings in the additional (and much larger) impact as downstream sectors reallocate and the relative price of AI services shifts.

**Figure 4:** GDP Impact of AI sanctions (in percent of GDP).




*Note:* The first-order Hulten row reports the effect under the central productivity calibration. The second-order Törnqvist measure uses the central productivity calibration and a range of applicable macro elasticities ($\rho \in [0.2, 0.6]$); the divergence-corner is the Törnqvist value at $\rho = 0.2$ and $\Delta = -0.63$. Dashed vertical lines mark two historical benchmarks: the 2011 Tohoku earthquake and the 1973 oil shock.

Some observations bear highlighting. The range of estimates is an interval from 0.12 to 1.43 percent of GDP. Depending on the underlying assumptions, a sudden denial of frontier AI services could have a minor effect on the targeted economy, or it could inflict the kind of damage that might push a number of countries into recession. The headline estimate is 0.77 percent of GDP, which sits somewhere between the supply-chain disruption caused by the 2011 Tohoku earthquake and the 1973 oil shock.

Under certain conditions, though, it is possible to obtain stronger effects. If the production network behaves closer to Leontief (a small macro elasticity), and if the drop in the productivity of the AI sector were sufficiently large (as in the divergence calibration for Estimator A), the economic impact estimate rises to 2.27 percent of GDP. This last number, however, relies on strong assumptions about the underlying nature of the complementarity between AI and the rest of the economy, and should not be interpreted as the main takeaway result. That said, the headline estimate range for the second-order central envelope also assumes away certain frictions –in particular, the cost, in resources and time, for firms to switch from frontier AI services to feasible alternatives– which

could introduce additional economic losses.

# 6 Concluding Remarks

The premise of this paper is that artificial intelligence services have, in the span of three years, crossed the threshold from consumer good to general-purpose production input, and that the concentration of frontier AI services in a single global supplier makes them, in principle, a sanctionable channel of strategic leverage. Whether that leverage is large enough to matter is an empirical question, to which this paper provides an answer in the form of a quantitative envelope, calibrated to a representative advanced small open economy.

At the baseline calibration, the impact is close to 0.77 percent of GDP, but the range of possible outcomes is much larger: under convergence assumptions, the impact would be small enough to be lost in normal quarterly noise. At the divergence corner, where complementarity is severe and the AI frontier capability gap is wide, the impact could come close to 2.27 percent of GDP – more than enough to induce a recession in most affected economies.

The results rely on a number of simplifying decisions. First, to estimate the post-shock weight of the AI sector I rely on a two sector CES environment, where the input-output network features a single macro elasticity $\rho$, so the empirical heterogeneity in how different using industries substitute away from AI is folded into one scalar. Second, the framing is static: the pre-shock Domar weight and the productivity gap are both calibrated to the 2026 frontier, so any further widening of the frontier AI capability gap would move the baseline results toward the divergence end of the interval. Third, the Hicks-neutral TFP-shock representation smooths over the regime-shift character of an actual sanction — the possibility that, at sufficiently large $|\Delta \log A|$, firms abandon AI-dependent workflows discontinuously rather than continuing along a degraded production function.

Several extensions follow naturally. The most direct is to replace the two sector CES aggregator with a multiple sector representation; this would condition the macro elasticity on the empirical distribution of AI usage across industries. A second is to embed the exercise in a forward diffusion path with endogenous capability growth between the calibration date and the shock date, converting the paper's level envelope into a trajectory. A third, more ambitious extension would be to relax the Hicks-neutral framing in favour of a model that allows discrete workflow abandonment at sufficiently large $|\Delta \log A|$, which would generate the regime-shift behaviour the smooth elasticity cannot.

The methodology developed here is not specific to any one country. Any economy undergoing widespread adoption of AI into its production processes, and relyingon global AI services, can be considered. As enterprise AI adoption deepens and the capability gap between frontier and non-frontier systems continues to shift, the same exercise can be re-run on updated data, with the envelope

tightening as the underlying primitives become more cleanly identified.

# References


Allcott, H., Collard-Wexler, A., and O'Connell, S. D. (2016). How do electricity shortages affect industry? Evidence from India. *American Economic Review*, 106(3):587–624.

Allen, J. S. (2026). Monitoring AI adoption in the U.S. economy. *FEDS Notes*. Published April 3, 2026. Reports BTOS firm-weighted AI adoption of 18 percent at end-2025 (broad measure), with sector adoption ranging from 8 percent (Accommodation and Food Services) to 37 percent (Information). Forward-looking BTOS adoption: over 20 percent of firms expect to use AI in the first half of 2026. RPS work-related GenAI use 41 percent (Nov 2025).

Alphabet (2026). Alphabet Q4 2025 earnings call and press release. https://www.cnbc.com/2026/02/04/alphabet-googl-q4-2025-earnings.html. Q4 2025 earnings. Google Cloud on annual run rate exceeding $70B in Q4 2025 (48% YoY growth). Management: AI-related revenues "billions of dollars per quarter"; products built on generative models grew more than 200% year-on-year.

Amodei, D. (2026). Remarks on the US–China model capability gap. Reported in PYMNTS, 5 May 2026. "the Chinese models are six to 12 months behind Anthropic's models".

Anthropic (2025a). Anthropic acquires Bun as Claude Code reaches $1B milestone. https://www.anthropic.com/news/anthropic-acquires-bun-as-claude-code-reaches-usd1b-milestone. December 2025. Confirms Claude Code reached $1B ARR within six months of general availability (May 2025). Announces acquisition of Bun JavaScript runtime.

Anthropic (2025b). Anthropic raises Series F at $183B post-money valuation. https://www.anthropic.com/news/anthropic-raises-series-f-at-usd183b-post-money-valuation. September 2025. Series F closing announcement ($13B raised, led by ICONIQ). Concurrent reporting placed Anthropic's ARR at $5B+.

Anthropic (2026a). The anthropic economic index: Economic primitives. Technical report, Anthropic. Available at https://www.anthropic.com/research/anthropic-economic-index-january-2026-report.

Anthropic (2026b). Anthropic raises $30 billion in Series G funding at $380 billion post-money valuation. https://www.anthropic.com/news/anthropic-raises-30-billion-series-g-funding-380-billion-post-money-valuation. February 12, 2026. Series G announcement ($30B raised, led by GIC and Coatue). Discloses run-rate revenue of $14B, "growing over 10× annually in each of the past three years". Eight of the Fortune 10 are Claude customers.

Appel, R., McCrory, P., Tamkin, A., McCain, M., Neylon, T., and Stern, M. (2025). Anthropic economic index report: Uneven geographic and enterprise

AI adoption. arXiv 2511.15080, Anthropic. Privacy-preserving analysis of 1 million Claude.ai conversations and 1 million API transcripts across 150+ countries. Documents geographic concentration of AI usage in high-income countries; the U.S. accounts for 25.4 percent of global Claude.ai conversation traffic in February 2026.

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics*, 9(4):254–280.

Axios (2026). Scoop: OpenAI projects $100 billion in ad revenue by 2030. https://www.axios.com/2026/04/09/openai-100-billion-in-ad-revenue. April 9, 2026. Reports $2.5B in 2026 advertising revenue projection, growing to $25B by 2028.

Baqaee, D. R. and Farhi, E. (2019). The macroeconomic impact of microeconomic shocks: Beyond Hulten's theorem. *Econometrica*, 87(4):1155–1203.

Barrot, J.-N. and Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *Quarterly Journal of Economics*, 131(3):1543–1592.

BenchLM (2026). Open-source LLM benchmark comparison. Technical report, BenchLM.

Bick, A., Blandin, A., and Deming, D. J. (2024). The rapid adoption of generative AI. NBER Working Paper 32966, National Bureau of Economic Research, Cambridge, MA. Methodology source for the worker-side daily-bin imputation of AI hours used in Bick et al. (2026). Reports US adoption of generative AI at home and at work, hours of use per week, and time-savings imputations from a Real-Time Population Survey module fielded in June, August and November 2024.

Bick, A., Blandin, A., Deming, D. J., Fuchs-Schündeln, N., and Jessen, J. (2026). Mind the gap: AI adoption in Europe and the U.S. NBER Working Paper 34995, National Bureau of Economic Research, Cambridge, MA. Forthcoming, *Brookings Papers on Economic Activity*, Spring 2026. Reconciles BTOS and Eurostat ICT Enterprise Survey methodologies; finds that U.S. workers and U.S. firms (on the directly comparable production-rate question) lead Europe in AI adoption, with the U.S. in the top tier of European countries on broader Eurostat measures.

Bloomberg (2025). OpenAI expects revenue will triple to $12.7 billion this year. https://www.bloomberg.com/news/articles/2025-03-26/openai-expects-revenue-will-triple-to-12-7-billion-this-year. March 26, 2025. OpenAI's own 2025 full-year revenue projection.

Bloomberg (2026a). Anthropic nears $20 billion revenue run rate amid pentagon feud. https://www.bloomberg.com/news/articles/2026-03-03/anthropic-nears-20-billion-revenue-run-rate-amid-pentagon-feud. March 3, 2026. Reports Anthropic ARR at $19B, citing CEO Dario Amodei

at Morgan Stanley TMT conference. Notes $6B added in February 2026 alone.

Bloomberg (2026b). Anthropic tops $30 billion run rate, seals Broadcom deal. https://www.bloomberg.com/news/articles/2026-04-06/broadcom-confirms-deal-to-ship-google-tpu-chips-to-anthropic. April 6, 2026. Reports Anthropic ARR at $30B alongside announcement of Broadcom deal to supply Google TPU chips.

Bloomberg (2026c). Anthropic's revenue run rate tops $9 billion as VCs pile in. https://www.bloomberg.com/news/articles/2026-01-21/anthropic-s-revenue-run-rate-tops-9-billion-as-vcs-pile-in. January 21, 2026. Reports $9B ARR at end of 2025.

Bloomberg (2026d). DeepMind CEO says Chinese AI firms are 6 months behind the west. https://www.bloomberg.com/news/articles/2026-01-20/deepmind-ceo-says-chinese-ai-firms-are-6-months-behind-the-west. January 20, 2026. Reports Demis Hassabis's interview with Emily Chang at the World Economic Forum in Davos. Hassabis: Chinese AI firms "haven't been able to innovate beyond the cutting edge" and "remain about six months behind the frontier AI of the leading western labs." Characterises the response to DeepSeek-R1 as a "massive overreaction.".

Board of Governors of the Federal Reserve System (2026). Monitoring AI adoption in the U.S. economy. https://www.federalreserve.gov/econres/notes/feds-notes/monitoring-ai-adoption-in-the-u-s-economy-20260403.html. FEDS Notes, April 3, 2026.

Boehm, C. E., Flaaen, A., and Pandalai-Nayar, N. (2019). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake. *Review of Economics and Statistics*, 101(1):60–75.

Brynjolfsson, E., Li, D., and Raymond, L. R. (2023). Generative AI at work. Working Paper 31161, National Bureau of Economic Research. Revised 2025.

Business of Apps (2026). Google gemini revenue and usage statistics (2026). https://www.businessofapps.com/data/google-gemini-statistics/. Accessed 2026. Estimates 2025 consumer subscription revenue from the Gemini Advanced (Google One AI Premium) tier at approximately $1.2 billion, compiled from Google earnings commentary, subscriber growth disclosures, and industry modelling. Google does not disclose this tier separately in its financial filings.

Calvino, F. and Fontanelli, L. (2023). A portrait of AI adopters across countries: Firm characteristics, assets' complementarities and productivity. OECD Science, Technology and Industry Working Paper 2023/02, Organisation for Economic Co-operation and Development, Paris. Distributed-microdata harmonisation of national firm AI surveys across Belgium, Denmark, France, Germany, Ireland, Israel, Italy, Japan, Korea, Portugal and Switzerland for the 2017–2020 reference period. Provides the first directly comparable cross-

country firm-AI adoption rates including Japan, Korea and Israel.

CNBC (2025a). Anthropic hits $3 billion in annualized revenue on business demand for AI. https://www.cnbc.com/2025/05/30/anthropic-hits-3-billion-in-annualized-revenue-on-business-demand-for-ai.html. May 30, 2025. Reports Anthropic ARR at $3B driven by enterprise and API demand.

CNBC (2025b). Anthropic reaches $2 billion in annualized revenue. https://www.cnbc.com/2025/03/12/anthropic-reaches-2-billion-annualized-revenue.html. March 2025. Reports Anthropic ARR at $2B.

CNBC (2025c). Google agrees to new $1 billion investment in Anthropic. https://www.cnbc.com/2025/01/22/google-agrees-to-new-1-billion-investment-in-anthropic.html. January 22, 2025. Reports Anthropic's December 2024 ARR reached $1 billion, approximately 10x year-over-year growth.

CNBC (2025d). OpenAI hits $10 billion in annual recurring revenue fueled by ChatGPT growth. https://www.cnbc.com/2025/06/09/openai-hits-10-billion-in-annualized-revenue-fueled-by-chatgpt-growth.html. June 9, 2025. Confirms June 2025 ARR at $10B and December 2024 ARR at $5.5B. Notes figures exclude Microsoft licensing revenue. Revenue mix: consumer subscriptions, team/enterprise plans, and direct API access.

CodeSOTA (2026). LLM leaderboard 2026: GPT-5 vs claude Opus 4.6 vs gemini 3 (MMLU, GPQA, HLE, LiveCodeBench). https://www.codesota.com/llm. Cross-benchmark snapshot of frontier closed-weight and Chinese open-weight model scores on MMLU, GPQA Diamond, AIME 2026, HumanEval, HLE and SWE-Bench Verified/Pro, accessed April 2026. Source for the empirical ranking claims that Qwen leads on GPQA Diamond and that the open-weight gap on AIME 2026 is approximately two points.

Comin, D., Lashkari, D., and Mestieri, M. (2021). Structural change with long-run income and price effects. *Econometrica*, 89(1):311–374.

DeepSeek-AI (2026). DeepSeek-V4: A million-token context that agents can actually use. Hugging Face technical-report blog, https://huggingface.co/blog/deepseekv4. DeepSeek's own technical-report companion for the V4 release of April 24, 2026. Reports SWE-Bench Verified 80.6% for V4-Pro (within 0.2 percentage points of Claude Opus 4.6 at 80.8% and matching Gemini 3.1 Pro at 80.6%) and Terminal-Bench 2.0 67.9% for V4-Pro (compared with Opus 4.6 at 65.4% and Gemini 3.1 Pro at 68.5%). Contains the verbatim self-assessment that V4 is on a "developmental trajectory that trails state-of-the-art frontier models by approximately 3 to 6 months," subsequently reported by TechCrunch (2026).

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of large language models. *Science*, 384(6702):1306–1308.

Epoch AI (2025). OpenAI is projecting unprecedented revenue growth. https://epoch.ai/gradient-updates/openai-is-projecting-unprecedented-revenue-growth/. Documents OpenAI's projected 2.3×, 2×, 1.6× annual growth multipliers through 2028. Accessed 2025.

Epoch AI (2026a). AI companies revenue reports (structured database). https://epoch.ai/data/ai_companies_revenue_reports.csv. Continuously updated structured database of disclosed and reported revenue milestones for AI companies, with confidence ratings and primary source citations. Accompanying analysis at https://epoch.ai/data-insights/openai-revenue/. Accessed April 2026.

Epoch AI (2026b). Chinese AI models have lagged the US frontier by 7 months on average since 2023. https://epoch.ai/data-insights/us-vs-china-eci. Published January 2, 2026. The Epoch Capabilities Index (ECI) puts the 2023–2025 gap between leading Chinese models and the U.S. frontier at a 7-month average, with a 4-month minimum and a 14-month maximum. As of the publication date no Chinese model had matched the ECI of OpenAI's o3 (released April 2025). Notes that nearly all leading Chinese models in the comparison are open-weight while frontier U.S. models remain proprietary.

Epoch AI (2026c). The Epoch capabilities index (ECI). https://epoch.ai/eci. Composite capability index across benchmarks; open-vs-closed frontier gap of about four months. Accessed May 2026.

Financial Times (2024). OpenAI's annual revenues surpass $2 billion. https://www.ft.com/content/81ac0e78-5b9b-4f33-8601-d715e576cfea. February 2024. Reports OpenAI reached an ARR of $2B in December 2023, equivalent to $166.6M per month.

Financial Times (2025). OpenAI API revenue nears $3 billion annualised rate. https://www.ft.com/content/openai-api-revenue-mid-2025. Mid-2025. Reports OpenAI's direct API (developer) revenue running at approximately $3B ARR by mid-2025, representing roughly one-quarter of total ARR at that point.

Forrester Research (2024). Forrester: US set to dominate AI enterprise software market. https://www.computerweekly.com/news/252525729/Forrester-US-set-to-dominate-AI-enterprise-software-market. Forrester estimates the United States accounts for 46 percent of the global enterprise AI software market in 2024, and 41 percent of global technology spending overall. Reported by *Computer Weekly*. Accessed April 2026.

Gerstner, B. (2026). Altimeter capital: Anthropic's revenue run rate could triple in 2026. https://www.msn.com/en-us/money/news/altimeter-s-brad-gerstner-says-anthropic-s-revenue-run-rate-could-triple-in-2026-as-ai-worries-hit-tech-stocks/ar-AA20KOrC. March–April 2026. Altimeter Capital CEO Brad Gerstner speaking on the All-In podcast. States he "would not be shocked" if Anthropic exits 2026 at $80–100B ARR.

Notes Anthropic generated $6B in revenue in February 2026 alone and that the company has over 1,000 customers spending more than $1M per year.

Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., Mueller, J., Hong, J., Ritchie, S., Belonax, T., Troç, K. N., Olsson, C., Amodei, D., and Ganguli, D. (2025). Which economic tasks are performed with AI? evidence from millions of Claude conversations. Technical report, Anthropic. Available at https://www.anthropic.com/research/economic-index.

Heath, A. (2026). Read OpenAI's latest internal memo about beating the competition—including Anthropic. https://www.theverge.com/ai-artificial-intelligence/911118/openai-memo-cro-ai-competition-anthropic. April 12, 2026. Reports on a four-page internal memo by OpenAI Chief Revenue Officer Denise Dresser. Key claim on Anthropic: "Their stated run rate is inflated. They use accounting treatment that makes revenue look bigger than it is, including grossing up rev share with Amazon and Google. Our analysis shows that this overstates their run rate by roughly $8 billion (at the current $30 stated). We report Microsoft revshare net, which is more inline with standards we would be held to as a public company.".

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, 45(3):511–518.

Jimenez, C. E., Yang, J., Wettig, A., Yao, S., Pei, K., Press, O., and Narasimhan, K. (2024). SWE-bench: Can language models resolve real-world GitHub issues? In *International Conference on Learning Representations (ICLR)*.

LMSYS (2026). Chatbot arena leaderboard. Technical report, LMSYS. Accessed April 2026.

Mathematical Association of America (2026). American Invitational Mathematics Examination (AIME). https://maa.org/maa-invitational-competitions/. Annual high-school invitational mathematics competition administered by the MAA. Repurposed in 2024–2026 as an out-of-distribution reasoning benchmark for frontier language models. Performance figures cited in this paper draw on the AIME 2026 problem set.

Menlo Ventures (2025). The state of generative AI in the enterprise 2025. https://menlovc.com/2025-the-state-of-generative-ai-in-the-enterprise/. Enterprise model API market shares: Anthropic 40%, OpenAI 27%, Google 21%. Accessed 2025.

METR (2025a). Preliminary evaluation of DeepSeek and Qwen models. Technical report, METR. Available at https://evaluations.metr.org/deepseek-qwen-report/.

METR (2025b). Task-completion time horizons of frontier AI models. Technical report, METR. Available at https://metr.org/time-horizons/.

METR (2026). AI time horizons dashboard. https://ai-time-horizons.

vercel.app/. Interactive companion to METR Time Horizon 1.1, plotting METR 50- and 80-percent time horizons for U.S. closed-weight and Chinese open-weight models. Accessed April 2026; reports Claude Opus 4.6 at 718.81 minutes, Gemini 3.1 Pro at 384.15 minutes, GPT-5.4 at 341.74 minutes on the U.S. frontier; Kimi K2 Thinking at 57.62 minutes on the best Chinese open-weight model with a public METR score.

Microsoft AI Economy Institute (2026). Global AI adoption in 2025: A widening digital divide. Technical report, Microsoft Research. Microsoft AI Diffusion Report, January 8, 2026. Reports country-level market shares for DeepSeek by chatbot-traffic share: 89% in China (the highest national share recorded), 56% in Belarus, 49% in Cuba, 43% in Russia, 25% in Iran, 23% in Syria, and 11–14% across several Sub-Saharan African economies; DeepSeek uptake in North America and Europe is described as "low." Companion blog post at https://blogs.microsoft.com/on-the-issues/2026/01/08/global-ai-adoption-in-2025/.

Microsoft Corporation (2025). Microsoft cloud and AI strength drives second quarter results. https://www.microsoft.com/en-us/investor/earnings/fy-2025-q2/press-release-webcast. FY25 Q2 earnings call, January 29, 2025. Amy Hood: "Our AI business has surpassed an annual revenue run rate of $13 billion, up 175 percent year-over-year." Fiscal Q2 ends December 31; the $13B reading corresponds to year-end 2024 ARR. Microsoft discontinued the AI-business ARR disclosure after this release.

Microsoft Corporation (2026). Microsoft cloud and AI strength drives second quarter results. https://news.microsoft.com/source/2026/01/28/microsoft-cloud-and-ai-strength-drives-second-quarter-results-3/. FY26 Q2 earnings call, January 28, 2026. Reports 15 million paid Microsoft 365 Copilot seats and 4.7 million paid GitHub Copilot subscribers. Management did not restate the aggregate AI-business ARR metric disclosed a year earlier.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

O*NET Resource Center (2024). O*NET 29.0 Database: Task Statements. Technical report, National Center for O*NET Development for the U.S. Department of Labor, Employment and Training Administration. Available at https://www.onetcenter.org/database.html.

OpenAI (2026a). A business that scales with the value of intelligence. https://openai.com/index/a-business-that-scales-with-the-value-of-intelligence/. January 2026. OpenAI company blog post. Discloses year-end ARR milestones: "$2B in 2023, $6B in 2024, $20B+ in 2025". These are ARR figures, not full-year cumulative revenues. Confirms approximately $2B in monthly revenue.

OpenAI (2026b). Introducing GPT-5.5. https://openai.com/index/

introducing-gpt-5-5/. Released April 23, 2026. First fully retrained base model since GPT-4.5. Terminal-Bench 2.0 82.7%, GDPval 84.9% across 44 knowledge-work occupations, FrontierMath Tier 4 39.6% (Pro). See https://techcrunch.com/2026/04/23/openai-chatgpt-gpt-5-5-ai-model-superapp/.

OpenRouter (2025). State of AI 2025: 100T token LLM usage study. https://openrouter.ai/state-of-ai. Empirical study of over 100 trillion tokens processed across more than 300 models and 60 providers. Geography derived from billing-region data. United States: 47.17 percent of global LLM token volume; North America: 47.22 percent; Europe: 21.32 percent. Accessed April 2026.

Perspectives Plus (2026). Microsoft's AI numbers: The good, the bad and the ugly. https://www.perspectives.plus/p/microsoft-ai-numbers-good-bad-ugly. Analyst commentary on Microsoft's decision to stop disclosing AI-business ARR after FY25 Q2 and on the pace of Microsoft 365 Copilot enterprise uptake. Argues the discontinuation is consistent with deceleration below the 175 percent year-on-year path implied by the January 2025 disclosure.

PolitiFact (2026). Trump says the U.S. leads China "by a lot" in AI. experts say that lead is measured by a few months. https://www.politifact.com/factchecks/2026/jan/29/donald-trump/trump-says-the-us-leads-china-by-a-lot-in-ai-exper/. January 29, 2026. Compiles industry, government and academic estimates of the U.S.–China frontier-AI gap as of early 2026: David Sacks (White House AI czar) "three to six months"; Jensen Huang (Nvidia) "nanoseconds"; Matt Sheehan (Carnegie) "six to 18 months"; Demis Hassabis (Google DeepMind) "six months.".

PricewaterhouseCoopers (2026). PwC 2026 AI performance study: Three-quarters of AI's economic gains are being captured by just 20 percent of companies. https://www.pwc.com/gx/en/news-room/press-releases/2026/pwc-2026-ai-performance-study.html.

Rein, D., Hou, B. L., Stickland, A. C., Petty, J., Pang, R. Y., Dirani, J., Michael, J., and Bowman, S. R. (2023). GPQA: A graduate-level google-proof Q&A benchmark. *arXiv preprint arXiv:2311.12022*.

Rest of World (2025). China's AI frenzy: DeepSeek is already everywhere—cars, phones, even hospitals. https://restofworld.org/2025/china-embeds-deepseek-ai-in-everything/. Documents the rapid integration of DeepSeek's open-weight models across Chinese sectors in early 2025: more than twenty automakers (BYD, Geely, Zeekr among them), at least nine securities firms, three state-owned telecommunications operators, and consumer-electronics, hospital and local-government deployments. Identifies the rate, scale and breadth of integration as "unprecedented" relative to prior open-weight adoption episodes.

SaaStr (2026). Anthropic just passed OpenAI in revenue. while spending 4x less to train their models. https://www.saastr.com/anthropic-just-passed-openai-in-revenue-while-spending-4x-less-to-train-their-models/. Accessed April 2026. Reports 75–80% enterprise/API revenue mix.

Sajadieh, S., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Santarlasci, L., Pava, J., Maslej, N., et al. (2026). The AI index 2026 annual report. https://aiindex.stanford.edu/report/. Stanford, CA. April 2026. Section 4.2 tracks annualized revenues for leading frontier AI companies—OpenAI, Anthropic, xAI, and Mistral AI—using Epoch AI data, but reports no standalone revenue figure for Google Gemini or Vertex AI, confirming the absence of a separate public disclosure for those products.

Sixth Tone (2025). Seek and deploy: China's tech giants race to adopt DeepSeek AI. https://www.sixthtone.com/news/1016609. Reports that within weeks of the DeepSeek-R1 release in January 2025, the model had been deployed in over 70 local governments and across the central state-owned-enterprise system supervised by SASAC, including Sinochem and China Southern Power Grid. Beijing municipal government began implementing DeepSeek on March 4, 2025 as a 24-hour business-consultation service.

South China Morning Post (2025). China's M&A deals to rise from decade-low as DeepSeek, AI spur interest, experts say. https://www.scmp.com/business/banking-finance/article/3300608/chinas-ma-deals-rise-decade-low-deepseek-ai-spur-interest-experts-say. Discusses the cross-sector impact of DeepSeek adoption on Chinese commercial banks, brokerages, and state-owned enterprises in the early months of 2025, including the role of SASAC and the broader implications for M&A activity. Confirms the integration of DeepSeek into financial-sector workflows alongside the automaker and telecommunications adoption documented in Rest of World (2025) and Sixth Tone (2025).

Startupnews.fyi (2026). OpenAI projects $2.5 billion in ad revenue for 2026, $100 billion by 2030: Report. https://startupnews.fyi/2026/04/12/openai-projects-2-5-billion-in-ad-revenue-this-year-100-billion-by-2030-report/. April 2026. Reports revised $280B total revenue target by 2030.

State Council of the People's Republic of China (2025). Opinions of the State Council on deeply implementing the "Artificial Intelligence+" initiative. https://english.www.gov.cn/policies/latestreleases/202508/27/content_WS68ae7976c6d0868f4e8f51a0.html. State Council guideline issued August 2025. Sets explicit penetration targets for "new-generation intelligent terminals and AI agents" across six priority areas (science and technology, industry, consumption, public welfare, governance, global cooperation): above 70% by 2027, above 90% by 2030, with a fully "intelligent economy and intelligent social development" phase targeted for 2035.

TechCrunch (2025). Anthropic expects B2B demand to boost revenue to $70b in 2028. https://techcrunch.com/2025/11/04/anthropic-expects-b2b-demand-to-boost-revenue-to-70b-in-2028-report/. November 4, 2025.

TechCrunch (2026). DeepSeek previews new AI model that "closes the gap" with frontier models. https://techcrunch.com/2026/04/24/deepseek-previews-new-ai-model-that-closes-the-gap-with-frontier-models/. Reports a "developmental trajectory that trails state-of-the-art frontier models by approximately 3 to 6 months" for DeepSeek V4 in reasoning benchmarks.

The Information (2024). Anthropic's revenue run rate hit $87 million in January. https://www.theinformation.com/briefings/anthropic-revenue-run-rate-87-million-january. 2024. Reports Anthropic's January 2024 ARR at $87M, a baseline from which the company grew to $30B ARR by April 2026. Rated "Confident" in Epoch AI revenue database.

The Information (2025a). Anthropic approaches $7 billion revenue run rate, surpassing OpenAI in enterprise API. https://www.theinformation.com/briefings/anthropic-approaches-7-billion-revenue-run-rate. October 2025. Reports Anthropic ARR at approximately $7B, surpassing OpenAI's enterprise API run rate. Consistent with Sacra and SaaStr ARR trajectory data.

The Information (2025b). Anthropic projects $70 billion in revenue, $17 billion in cash flow in 2028. https://www.theinformation.com/articles/anthropic-projects-70-billion-revenue-17-billion-cash-flow-2028. November 2025. Base case $70B; bull case $148B by 2028.

The Information (2025c). Anthropic projects soaring growth to $34.5 billion in 2027 revenue. https://www.theinformation.com/articles/anthropic-projects-soaring-growth-to-34-5-billion-in-2027-revenue. Reports Anthropic's projected 2024 revenue of $918M, and 2025 projections of $2B base case and $4B optimistic case.

The Information (2025d). OpenAI's annualized revenue hits $12 billion. https://www.theinformation.com/briefings/openai-annualized-revenue-hits-12-billion. July 2025. Reports OpenAI ARR at $12B.

The Information (2025e). OpenAI's own forecast predicts $14 billion loss in 2026 but Nvidia-style $100 billion revenues by 2029. https://finance.yahoo.com/news/openais-own-forecast-predicts-14-150445813.html. Investor documents prepared Q3 2025, reported by The Information. Summarised via Yahoo Finance.

The Information (2025f). OpenAI's revenue in first half of 2025 reached $4.3 billion. https://www.theinformation.com/briefings/openai-revenue-first-half-2025. 2025. Reports OpenAI's H1 2025 actual cumulative revenues at $4.3 billion, consistent with rapid intra-year growth toward the $13 billion full-year 2025 figure.

The Information (2026a). Anthropic hikes 2026 revenue forecast 20%, delays when it will go cash flow positive. https://www.theinformation.com/articles/anthropic-hikes-2026-revenue-forecast-20-delays-will-go-cash-flow-positive. 2026. Reports Anthropic's revised 2026 internal revenue forecast of $18 billion (up 20% from prior forecast of $15B). Also projects $55 billion in 2027 revenues.

The Information (2026b). The math behind Anthropic's mad revenue growth. https://www.theinformation.com/newsletters/dealmaker/math-behind-anthropics-mad-revenue-growth. 2026. Reports Anthropic's 2025 full-year actual revenue at $4.5 billion—nearly 12 times 2024 actual revenue. Notes the vast majority of this is enterprise and API revenue.

The Information (2026c). OpenAI tops $25 billion in annualized revenue as Anthropic narrows gap. https://www.theinformation.com/articles/openai-tops-25-billion-annualized-revenue-anthropic-narrows-gap. February 2026. Reports OpenAI ARR crossed $25B at end of February 2026, up from $21.4B at year-end 2025. Weekly active users at 910M; paying business users surpassed 9M.

The New York Times (2026). OpenAI's soaring revenue is shadowed by billion-dollar losses. https://www.nytimes.com/2026/02/10/technology/openai-revenue-losses.html. February 2026. Reports OpenAI's full-year 2025 actual revenue at approximately $13 billion.

Türkiye İstatistik Kurumu (TÜIK) (2025). Survey on information and communication technology (ICT) usage in enterprises, 2025. https://data.tuik.gov.tr/. 2025 wave of TÜIK's annual ICT-in-enterprises survey, harmonised with Eurostat's enterprise-AI methodology to permit cross-country comparison with EU-27 member states. Reports 7.5% of Turkish enterprises with ten or more employees used at least one AI technology in 2025, up from 4.4% in 2024 and 2.7% in 2021. Adoption concentrated in the information and communication sector (47.1%) and in firms with 250 or more employees (24.1%).

U.S. Bureau of Labor Statistics (2018). 2018 Standard Occupational Classification system. Technical report, U.S. Department of Labor. Available at https://www.bls.gov/soc/2018/home.htm.

World Bank (2024). Imports of services (BoP, current US$). World Development Indicators, indicator BM.GSR.NFSV.CD. https://data.worldbank.org/indicator/BM.GSR.NFSV.CD. Balance-of-payments services imports by economy; 2023 values. Accessed via the World Bank API, 2026.

Xie, T., Zhang, D., Chen, J., Li, X., Zhao, S., Cao, R., Hua, T. J., Cheng, Z., Shin, D., Lei, F., et al. (2024). OSWorld: Benchmarking multimodal agents for open-ended tasks in real computer environments. *arXiv preprint arXiv:2404.07972*.

Z.ai (Zhipu AI) (2026). GLM-5.1: Open-weight frontier model. https://huggingface.co/zai-org/GLM-5.1. Released April 7, 2026. SWE-Bench

Pro 58.4% (briefly state-of-the-art), Terminal-Bench 2.0 63.5%, AIME 2026 95.3%. 744B-parameter MoE with 40B active per token. Trained on Huawei Ascend 910B without Nvidia GPUs; Zhipu AI has been on the US Entity List since January 2025.

Zitron, E. (2025). Exclusive: Here's how much OpenAI spends on inference and its revenue share with Microsoft. https://www.wheresyoured.at/oai_docs/. November 2025. Based on leaked financial documents. Reports Microsoft received $493.8M in revenue-share payments from OpenAI in 2024 and $865.8M in the first three quarters of 2025, implying a 20% revenue-share rate.

# A The Basket of Leading Adopters

This appendix reports the country-by-country calibration underlying the representative Domar weight of Section 4.1.2. The basket comprises the ten economies with the highest per-capita AI usage outside the United States, as measured by the Anthropic Economic Index (Anthropic, 2026a) (February 2026 snapshot), excluding the United Kingdom and France. For each member, Table 9 reports its usage index, its Alphabet revenue region, its 2026 GDP, and the resulting Domar weight under both revenue panels. Each weight is computed as $\lambda_{AI,c} = 0.53 \cdot \sigma_{\text{region}(c)} \cdot s_c \cdot R^{\text{global}}_{2028}/\text{GDP}_c$, with $\sigma_{\text{region}}$ taken from Alphabet's reported segments and $s_c$ the country's share of its region's total services imports (World Bank BoP, 2023). The equal-weighted mean across the ten members is the headline calibration, $\lambda_{AI} = 0.57$ percent (Panel A) and 0.41 percent (Panel B). The spread across members — in particular between the small, highly AI-intensive open economies and the larger members — substantiates the claim of Section 4.1 that the exercise is not specific to any one country.

**Table 9:** Per-Country Domar weights for the Basket of Leading Adopters

| Country | AUI | Region | GDP 2026 ($B) | $\lambda_{AI}$ (A) | $\lambda_{AI}$ (B) |
|---|---|---|---|---|---|
| Singapore | 4.68 | APAC | 660 | 2.33% | 1.67% |
| Israel | 4.40 | EMEA | 720 | 0.18% | 0.13% |
| Switzerland | 4.11 | EMEA | 1,147 | 0.46% | 0.33% |
| Canada | 3.70 | Other Americas | 2,507 | 0.34% | 0.24% |
| Australia | 3.48 | APAC | 2,124 | 0.23% | 0.16% |
| Netherlands | 3.41 | EMEA | 1,450 | 0.48% | 0.34% |
| Estonia | 3.29 | EMEA | 52 | 0.54% | 0.39% |
| South Korea | 3.07 | APAC | 1,931 | 0.39% | 0.28% |
| New Zealand | 3.04 | APAC | 279 | 0.32% | 0.23% |
| Sweden | 2.85 | EMEA | 760 | 0.42% | 0.30% |
| **Equal-weighted mean** | | | | **0.57%** | **0.41%** |

*Notes:* AUI = Anthropic AI Usage Index, per-capita usage relative to the global mean (February 2026; the United States reads 3.87 for reference). Region = Alphabet geographic segment, with region share of non-US revenue $\sigma_{\text{region}}$ of 0.56 (EMEA), 0.33 (APAC), and 0.12 (Other Americas). GDP from the IMF World Economic Outlook (2026). Each Domar weight is $\lambda_{AI,c} = 0.53 \cdot \sigma_{\text{region}(c)} \cdot s_c \cdot R^{\text{global}}_{2028}/\text{GDP}_c$, where $s_c$ is the country's share of its region's services imports (World Bank, 2024) and $R^{\text{global}}_{2028}$ is the narrow-definition global total ($378B Panel A, $270B Panel B) from Table 3. Panel A uses year-end ARR; Panel B full-year actual revenues.

# B Sensitivity: The Broad Definition of the Frontier AI Sector

This appendix develops the empirical counterpart of the broad definition of the frontier AI sector introduced in Section 4.1.1. The broad definition augments

the narrow perimeter of the three frontier providers (OpenAI, Anthropic, and Google) with Microsoft's AI-augmented platform and productivity products–specifically, Azure OpenAI Service (AOAI), Microsoft 365 Copilot, and GitHub Copilot. The three are bundled in Microsoft's own financial reporting as its "AI business." I build the Microsoft AI revenue trajectory bottom-up from the three components, anchor the 2024 level on disclosed seat counts and analyst decompositions, and project forward by inheriting the ARR growth trajectory of OpenAI from Section 4.1.1 in the absence of Microsoft-specific forward guidance. The resulting broad-definition Domar weight is then carried through the full battery of calculations executed in Section 5: the Hulten first-order response of Section 5.1, and the post-shock Domar weight and Törnqvist second-order total of Section 5.2.

## B.1 Why the Disclosed $13 Billion AI Business Figure may not be the Appropriate Anchor

On its fiscal-Q2 2025 earnings call (calendar Q4 2024), Microsoft disclosed that its "AI business" had reached a $13 billion annualized run rate, growing 175 percent year-on-year (Microsoft Corporation, 2025). At first glance this would seem the natural input for the broad-definition total. It is not, for two reasons.

First, the $13 billion figure aggregates substantially more than the three broad-definition products. Microsoft's "AI business" umbrella bundles AOAI, Microsoft 365 Copilot, and GitHub Copilot together with the remainder of Azure's AI services–including pre-LLM Cognitive Services (computer vision, speech transcription, document intelligence, form recognition, anomaly detection) and AI-related Azure infrastructure consumption. Only the first three are "frontier AI as a non-separable value driver" in the sense of Section 4.1.1; the residual Cognitive Services are narrow-ML products whose disruption under a US frontier AI sanction would be marginal.

Second, the $13 billion figure is difficult to reconcile with the reported revenues of OpenAI, the company whose models AOAI resells. At December 2024, OpenAI's direct-customer ARR was $5.5 billion. If AOAI alone accounted for, say, $8 billion of the $13 billion "AI business" figure, the implied total OpenAI-model end-customer revenue (OpenAI direct plus AOAI) would be roughly $13.5 billion and the Microsoft-channel share would exceed the direct channel at a stage of the partnership when direct-to-customer sales were generally understood to dominate. Analyst decompositions (Perspectives Plus, 2026) place standalone AOAI revenue at December 2024 in the $2–4 billion range, consistent with a combined OpenAI-model end-customer revenue near $7–9 billion rather than above $18 billion.

I therefore abandon the $13 billion headline as the anchor and rebuild the broad-definition Microsoft AI perimeter bottom-up from AOAI, Microsoft 365 Copilot, and GitHub Copilot separately.

## B.2 Component-Level December 2024 Anchors

**Azure OpenAI Service.** Microsoft does not disclose standalone AOAI revenue. I adopt $3 billion as the December 2024 central estimate–the midpoint of the $2–4 billion analyst range in Perspectives Plus (2026), consistent with fiscal-Q2 2025 call commentary that Azure AI services contributed materially to Azure's growth at a scale compatible with this figure within the broader Azure AI services line. AOAI is a direct resale of OpenAI model inference; its growth is mechanically coupled to OpenAI's, a fact I exploit for the forward projection below.

**Microsoft 365 Copilot.** I anchor December 2024 revenue using the seat trajectory Microsoft subsequently disclosed. At December 2025 the company reported 15 million paid Microsoft 365 Copilot seats, up 160 percent year-on-year (Microsoft Corporation, 2026), implying December 2024 seats of approximately 5.8 million. At the list price of $30/month and an assumed 50 percent average effective discount–consistent with the 40–60 percent discounting documented by Citi and J.P. Morgan for competitive enterprise deployments, and with implied ARR estimates in the $1.5–2.5 billion range at the December 2025 15 million-seat level–this delivers a December 2024 Microsoft 365 Copilot ARR of approximately $1.0 billion. For December 2025 I take the 15 million-seat disclosure at the same effective ASP, delivering $2.6 billion.

**GitHub Copilot.** At December 2025 Microsoft reported 4.7 million paid GitHub Copilot subscribers (Microsoft Corporation, 2026), roughly doubling year-on-year; I adopt 2.5 million as the December 2024 anchor. At an average effective ASP of $15/month (reflecting the mix of the Individual tier at $10, Business at $19, and Enterprise at $39, weighted toward the lower two), this delivers a December 2024 GitHub Copilot ARR of approximately $0.45 billion and a December 2025 ARR of $0.85 billion.

**Aggregate December 2024 anchor.** The three components sum to approximately $4.5 billion: $3.0B (AOAI) + $1.0B (M365 Copilot) + $0.45B (GitHub Copilot). This figure is substantially below the $13 billion aggregate "AI business" ARR Microsoft reported–the gap being Cognitive Services, AI-related Azure infrastructure, and related non-frontier lines–and substantially below the $5.5 billion OpenAI direct ARR at the same date. The latter property is the central reconciliation that the $13 billion headline fails: "Microsoft reselling OpenAI plus Copilots" is *smaller* than OpenAI's direct channel, consistent with direct-to-customer sales dominating the partnership's early years.

## B.3 Forward Projection: OpenAI-Anchored Growth Rates

A projection of Microsoft AI revenues through 2028 requires a growth trajectory. Microsoft discontinued aggregate AI business ARR disclosure after fiscal-Q2 2025 (Perspectives Plus, 2026) and has not subsequently provided forward

guidance that maps cleanly onto the three components of the broad-definition perimeter. In the absence of Microsoft-specific forecasts, I anchor the projection to OpenAI's ARR trajectory constructed in Section 4.1.1. The logic is two-fold. First, AOAI is a direct resale of OpenAI inference; its growth is mechanically coupled to OpenAI's. Second, for the two Copilot products, OpenAI's ARR growth rates are the best available comparator: both products compete in the same frontier AI application layer, scale through enterprise-software sales cycles of comparable character, and target broadly overlapping customer bases. I return to the limits of this assumption in Section B.9.

OpenAI's ARR trajectory from Table 3 implies year-on-year growth rates of +282% (December 2024 → December 2025), +100% (2025 → 2026), +90% (2026 → 2027), and +44% (2027 → 2028). For 2025, I have Microsoft-specific evidence for two of the three components: disclosed seat growth of +160% for Microsoft 365 Copilot and approximately +89% for GitHub Copilot (implied by the 2.5M to 4.7M trajectory). I apply these Microsoft-specific rates in 2025 and fall back on OpenAI's +282% only for AOAI (where no Microsoft-specific evidence exists). For 2026 through 2028, in the absence of any Microsoft-specific data, I apply OpenAI's rates uniformly across all three components. Table 10 collects the resulting trajectory.

**Table 10:** Microsoft AI Revenue Components, 2024–2028 (Year-End ARR, $B)

| Component | 2024 | 2025 | 2026 | 2027 | 2028 |
|---|---|---|---|---|---|
| Azure OpenAI Service | 3.0 | 11.5 | 22.9 | 43.6 | 62.7 |
| Microsoft 365 Copilot | 1.0 | 2.6 | 5.2 | 9.9 | 14.2 |
| GitHub Copilot | 0.45 | 0.85 | 1.70 | 3.24 | 4.66 |
| **Microsoft AI total (ARR)** | **∼4.5** | **∼15** | **∼30** | **∼57** | **∼82** |

*Notes:* 2024 anchors: AOAI from the midpoint of the $2–4B analyst range in Perspectives Plus (2026); Microsoft 365 Copilot from 5.8M back-calculated seats at $30/month list price with a 50% effective discount; GitHub Copilot from 2.5M back-calculated subscribers at $15/month effective ASP. 2025 growth rates: Microsoft 365 Copilot +160% and GitHub Copilot +89% from disclosed seat trajectories (Microsoft Corporation, 2026); AOAI +282% inherited from OpenAI's ARR growth in Table 3. 2026–2028 growth rates for all three components: OpenAI's disclosed ARR multipliers (The Information, 2025e)–implied year-on-year growth of +100%, +90%, and +44% respectively. Totals are rounded to aid reading; the underlying components carry one more decimal.

For the actual-revenue (Panel B) equivalent, I apply the linear-interpolation formula used throughout Section 4.1.1, actual ≈ (start-ARR + end-ARR)/2, with a December 2023 Microsoft AI ARR anchor of approximately $1.2 billion: $1.0B for AOAI (implied by roughly tripling to $3B over 2024, comparable to OpenAI's own 2024 growth from a $2B to a $5.5B ARR), $0.1B for Microsoft 365 Copilot (the product launched commercially in November 2023), and $0.1B for GitHub Copilot (approximately 1 million subscribers at $12/month effective). The resulting Microsoft AI actual-revenue trajectory is $3 billion (2024), $10 billion

(2025), $22 billion (2026), $43 billion (2027), and $69 billion (2028).

## B.4 Broad-Definition Revenue Totals

Table 11 reproduces the narrow-definition subtotals from Table 3, adds the Microsoft AI trajectory constructed above, and reports the broad-definition totals. By 2028 the broad-definition global revenues reach $452 billion in year-end ARR and $335 billion in full-year actual revenues, compared with $370 billion and $266 billion under the narrow definition.

**Table 11:** Broad-Definition Frontier AI Revenue, 2024–2028 ($B)

| | 2024 | 2025 | 2026 | 2027 | 2028 |
|---|---|---|---|---|---|
| *Panel A: Year-end ARR (central estimates)* | | | | | |
| Narrow total (OpenAI + Anthropic + Google) | ∼6.5 | ∼36 | ∼150 | ∼226 | ∼370 |
| Microsoft AI (Table 10) | 4.5 | 15 | 30 | 57 | 82 |
| **Broad total** | **∼11** | **∼51** | **∼180** | **∼283** | **∼452** |
| *Panel B: Full-year actual customer revenues (central estimates)* | | | | | |
| Narrow total (OpenAI + Anthropic + Google) | 5.6 | 20.5 | 92 | 184 | 266 |
| Microsoft AI | 3 | 10 | 22 | 43 | 69 |
| **Broad total** | **∼8.6** | **∼30.5** | **∼114** | **∼227** | **∼335** |

*Notes:* Narrow subtotals reproduced from Table 3; Microsoft AI row from Table 10. Both panels exclude advertising revenue and investor funding. Panel B is derived from Panel A via linear interpolation (start-ARR + end-ARR)/2 applied at the component level before aggregation.

## B.5 Broad-Definition Domar weight

Applying the same per-country allocation as in Section 4.1.2 to the broad-definition global totals — equivalently, scaling the narrow basket Domar weight by the broad-to-narrow revenue ratio ($452B/$370B = 1.22 under Panel A and $335B/$266B = 1.26 under Panel B) — delivers a broad basket Domar weight of $\lambda_{AI}^{\text{broad}} = 0.69\%$ (Panel A) and 0.51% (Panel B). Table 12 summarizes.

The broad-definition Domar weight is approximately 20–25 percent larger than its narrow-definition counterpart ($\lambda_{AI} \in [0.41\%, 0.57\%]$ from Table 4), reflecting the Microsoft AI contribution's share of the broad-definition total.

## B.6 Broad-Definition First-Order GDP Response

The Hicks-neutral productivity shock estimators of Section 4.2 apply to the broad definition without recalibration. Both estimators measure the capability ratio $\kappa$ between the US frontier and the best non-US alternative, and this ratio is a property of model quality independent of the revenue perimeter under which the Domar weight is computed. Microsoft AI's three products are built on US frontier models–OpenAI's GPT line for AOAI and the bulk of Microsoft 365

**Table 12:** Broad-Definition Domar weight of AI Services for the Representative Basket by 2028

| | Panel A (ARR) | Panel B (actual) |
|---|---|---|
| Global broad AI revenue ($B) | 452 | 335 |
| $\times (1 - 0.47)$: non-US ($B) | 239.6 | 177.6 |
| Allocated across the ten basket economies and averaged equally | | |
| **Broad Domar weight $\lambda_{AI}^{\text{broad}}$** | **0.69%** | **0.51%** |

*Notes:* Same per-country allocation as Table 4, applied to the broad-definition global totals from Table 11; equivalently, the narrow basket weight scaled by the broad-to-narrow revenue ratio. The uniform decomposition assumes Microsoft AI's customer base is distributed geographically in roughly the same proportions as the three frontier providers'; Microsoft's disclosed US revenue share sits near 50 percent, comparable to the 47 percent aggregate used throughout.

Copilot, with Anthropic's Claude powering part of GitHub Copilot–so a sanction cutting off US AI would reduce Microsoft AI's delivered capability by the same ratio as it would reduce the three frontier providers'. The four values of $\Delta \log A_{\text{AI}}$ (three scenarios of Estimator A and the single Estimator B reading) therefore transfer directly.[23]

Applying equation (4) with the broad-definition Domar weights from Table 12 delivers Table 13. The broad-definition envelope runs from $-0.13$ to $-0.43$ percent of GDP, versus $-0.11$ to $-0.36$ percent under the narrow definition of Table 7; the centrals span $-0.13$ to $-0.37$ percent under the broad definition versus $-0.11$ to $-0.31$ percent under the narrow. The widening is on the order of 20–25 percent across cells and does not alter any qualitative conclusion of the main-text exercise.

**Table 13:** Broad-Definition First-Order GDP Response, $\Delta \log Y^{(1)} = \lambda_{AI}^{\text{broad}} \cdot \Delta \log A_{\text{AI}}$, in percent of pre-shock GDP

| | Estimator A | | | Estimator B |
|---|---|---|---|---|
| | Conv. | Central | Div. | |
| $\Delta \log A_{\text{AI}}$ | $-0.29$ | $-0.54$ | $-0.63$ | $-0.26$ |
| Panel A ($\lambda_{AI}^{\text{broad}} = 0.69\%$) | $-0.20$ | $-0.37$ | $-0.43$ | $-0.18$ |
| Panel B ($\lambda_{AI}^{\text{broad}} = 0.51\%$) | $-0.15$ | $-0.28$ | $-0.32$ | $-0.13$ |

*Notes:* Cells report $\Delta \log Y^{(1)}$ in percent of pre-shock GDP. Panel A uses broad-definition year-end ARR; Panel B uses broad-definition full-year actual revenues. Estimators A and B as in Table 7. For comparison, the narrow-definition envelope of Table 7 runs from $-0.11$ to $-0.36$ percent.

[23] A task-mix reweighting toward Microsoft AI's product mix (heavier in office-productivity workflows, somewhat heavier in code) could shift Estimator B's central by a few log points in either direction–the sign is not obvious *ex ante*–but I do not attempt such a reweighting because no analogue of the Anthropic Economic Index task-composition disclosure exists for Microsoft AI.

## B.7 Broad-Definition Second-Order Response

The second-order machinery of Section 5.2 carries over directly to the broad definition: the logit update equation (12) is a function of the pre-shock Domar weight and the structural elasticity $\rho$, and swapping $\lambda_{AI}$ for $\lambda_{AI}^{\text{broad}}$ is the only change required. I adopt the same $\rho$ calibration as in the main text ($\rho \in [0.2, 0.6]$, baseline $\rho = 0.4$) and report the resulting post-shock broad Domar weights in Table 14.

**Table 14:** Broad-Definition Post-Shock Domar weight for the Representative Basket $\lambda_{AI}^{\text{broad}}(A + \Delta)$, in Percent of GDP, under Central Shocks

| | $\rho = 0.6$ | $\rho = 0.4$ (baseline) | $\rho = 0.2$ |
|---|---|---|---|
| Panel A (ARR) × Estimator A ($\Delta = -0.54$) | 0.99 | 1.54 | 5.68 |
| Panel A (ARR) × Estimator B ($\Delta = -0.26$) | 0.82 | 1.02 | 1.93 |
| Panel B (actual) × Estimator A ($\Delta = -0.54$) | 0.73 | 1.14 | 4.26 |
| Panel B (actual) × Estimator B ($\Delta = -0.26$) | 0.61 | 0.75 | 1.43 |

*Notes:* Logit update (12) applied with pre-shock broad Domar weights from Table 12 in place of the narrow weights of Table 4. The pre-shock anchors are $\lambda_{AI}^{\text{broad}}(A) = 0.69\%$ (Panel A) and 0.51% (Panel B).

The broad post-shock weights sit systematically 20–25 percent above their narrow counterparts in Table 6. At the $\rho = 0.2$ corner of the central grid (Panel A × Estimator A) the broad weight rises from 0.69 to 5.68 percent of GDP, versus the narrow counterpart rising from 0.57 to 4.74 percent. Applying the full Törnqvist formula (9) then delivers Table 15.

**Table 15:** Broad-Definition Total GDP Loss under the Törnqvist Approximation, Central Shocks (percent of pre-shock GDP)

| | $\rho = 0.6$ | $\rho = 0.4$ | $\rho = 0.2$ |
|---|---|---|---|
| Panel A × Estimator A | −0.45 | −0.60 | −1.72 |
| Panel A × Estimator B | −0.20 | −0.22 | −0.34 |
| Panel B × Estimator A | −0.33 | −0.45 | −1.29 |
| Panel B × Estimator B | −0.15 | −0.16 | −0.25 |

*Notes:* Törnqvist total log output loss, $\Delta \log Y = \frac{1}{2}[\lambda_{AI}^{\text{broad}}(A) + \lambda_{AI}^{\text{broad}}(A + \Delta)] \cdot \Delta \log A_{AI}$, applied with the pre- and post-shock Domar weights from Tables 12 and 14.

The broad-definition central envelope runs from −0.15 to −1.72 percent of GDP, against the narrow envelope of Table 8 of −0.12 to −1.43 percent. The widening is 20–25 percent across the twelve cells, inherited mechanically from the Domar-weight widening at the pre-shock anchor. The amplification ratio (Törnqvist over Hulten) is a function only of $\rho$ and $\Delta \log A$ through the logit update and is therefore *identical* across the narrow and broad definitions: every cell in Table 15 is its Table 8 counterpart scaled by $\lambda^{\text{broad}}/\lambda^{\text{narrow}} = 1.21$ (Panel A) and 1.24 (Panel B).

## B.8 Broad-Definition Divergence Corner

The divergence corner of the broad-definition exercise is constructed in direct parallel to the narrow case. At the divergence Estimator A shock $\Delta = -0.63$ with $\rho = 0.2$, the broad Panel A Törnqvist number is

$$\Delta \log Y^{\text{Törn}} \;=\; \tfrac{1}{2}(0.0069 + 0.0795) \cdot (-0.63) \;=\; -0.0272 \;=\; -2.72\%,$$

where the post-shock weight $\lambda_{AI}^{\text{broad}}(A + \Delta) = 7.95\%$ is obtained from equation (12) at the divergence shock. I take the broad-definition framework-internal divergence corner to be $-2.72$ percent of GDP, versus $-2.27$ percent under the narrow definition.

Stacking the three layers of the exercise under the broad definition, exactly mirroring Section 5.3:

1. *Hulten first-order floor* (Table 13): $-0.13$ to $-0.37$ percent at central shocks; $-0.13$ to $-0.43$ percent across the full grid.
2. *Törnqvist second-order envelope* (Table 15): $-0.15$ to $-1.72$ percent across the central $\Delta \log A$ shocks and $\rho \in [0.2, 0.6]$.
3. *Divergence corner (Törnqvist)*: $-2.72$ percent of GDP, at Panel A, divergence Estimator A shock $\Delta = -0.63$, and $\rho = 0.2$.

Each layer lies roughly 20 percent above its narrow counterpart. The framework-internal envelope widens from the narrow $[-0.12, -2.27]$ percent range to the broad $[-0.15, -2.72]$ percent range, a factor-of-19 spread inside each definition. The qualitative picture in Section 5.3–a framework-internal envelope that runs from a moderate commodity-export shock at the low end to a magnitude broadly comparable to Baqaee and Farhi (2019)'s 1973 oil counterfactual at the divergence corner–carries over without any change in character. The broad-definition calibration delivers a numerically larger but qualitatively indistinguishable answer, consistent with the interpretation that the narrow definition is a conservative lower bound on the exercise's headline.

## B.9 Caveats and Sensitivity

Three caveats qualify the broad-definition calibration.

First, the AOAI December 2024 anchor is the least-disciplined input of the three components. Microsoft does not disclose standalone AOAI revenue, and the \$3 billion central sits within a plausible range of at least \$2–4 billion. The broad-definition Domar weight and first-order response scale nearly proportionally with the AOAI anchor: a \$2 billion (resp. \$4 billion) December 2024 AOAI value, propagated with the same growth rates, delivers a 2028 Microsoft AI ARR of roughly \$61 billion (resp. \$103 billion) versus the \$82 billion central, and a Panel A broad-definition Domar weight of approximately 0.66 percent (resp. 0.73 percent) versus the 0.69 percent central.

Second, the use of OpenAI's ARR growth rates as the forward-projection driver embeds the assumption that Microsoft AI–and AOAI in particular–grows at rates comparable to OpenAI's direct channel. Enterprise cloud-channel adoption is conventionally slower than direct-to-customer channels, which would push AOAI's trajectory below the OpenAI-anchored projection. The broad-definition figures reported here therefore plausibly carry an upward bias of some magnitude. A conservative counter-specification that halves OpenAI's forward growth rates for AOAI alone (leaving the two Copilot projections unchanged) delivers a 2028 Microsoft AI ARR of roughly $38 billion (versus $82 billion), a 2028 broad-definition Panel A total of $408 billion (versus $452 billion), and a broad-definition Panel A Domar weight of approximately 0.63 percent (versus 0.69 percent)–still above the narrow-definition 0.57 percent by a modest margin.

Third, the broad definition as operationalized here excludes the balance of Microsoft's disclosed $13 billion AI business perimeter–generic Azure AI infrastructure, Cognitive Services, and related narrow-ML offerings–on the argument that these are not frontier AI products in the sense of Section 4.1.1. A more inclusive perimeter that treated the entire Microsoft AI business as sanctionable would add roughly $7–9 billion to the December 2024 Microsoft AI baseline, with correspondingly larger figures in 2028 under the same projection method.

None of the three caveats alters the qualitative conclusion that the broad-definition Domar weight, the first-order GDP response, the Törnqvist second-order total, and the Törnqvist divergence corner all lie within a factor of roughly 1.2 of their narrow-definition counterparts. The substantive headline of Section 5–a framework-internal envelope spanning from a few hundredths of a percent of GDP at the convergence end to roughly 2.27 percent at the divergence corner–is robust to the choice between the two perimeters.

# C Sensitivity: The Estimator A Frontier Anchor

The reliability-curve framework of Section 4.2 requires a pair $(h_{50}^F, h_{80}^F)$ to locate the US frontier on the task-length axis, after which the substitute curve is placed a fixed doubling lag behind. The body fixes the anchor at the median of METR Time-Horizon 1.1 scores (METR, 2026), $h_{50}^F = 6.40$ hours and $h_{80}^F = 1.16$ hours, and the doubling lag at the measured Claude Opus 4.5 vs. Kimi K2 Thinking gap of 2.27 (METR, 2025b). The doubling-lag sensitivity is reported in Section 4.2 as the convergence–divergence interval $\Delta \log A_{\text{AI}} \in [-0.71, -0.18]$. This appendix reports the orthogonal anchor sensitivity — the dependence of the headline shock on the choice of $(h_{50}^F, h_{80}^F)$ with the doubling lag held fixed — which Section 4.2 does not separately disclose. The anchor is the single highest-leverage discretionary choice in Estimator A: its variation alone moves $\Delta \log A_{\text{AI}}$ across an interval of width 0.27, comparable to roughly half the lag-based convergence–divergence band.

## C.1 The Counterintuitive Anchor Mechanism

A naive reading of the reliability-curve framework predicts that raising the frontier anchor deepens the shock: a more capable frontier appears to leave a larger gap to lose. The opposite is true. With the doubling lag held fixed at 2.27, raising $h_{50}^F$ shifts the substitute's $h_{50}^N$ proportionally rightward along the task-length axis, since $h_{50}^N = h_{50}^F \cdot 2^{-2.27}$. The frontier's coverage $\theta_F$ rises only modestly under this shift, because frontier reliability is already close to one on the short-task end of the AEI 1P API task-hour distribution (Anthropic, 2026a). The substitute's coverage $\theta_N$ rises faster, because the shift moves the substitute curve squarely into the one-to-ten-hour band where the AEI distribution carries the bulk of its mass. Since $\kappa_A = \theta_N/\theta_F$ is the ratio of the two, the substitute's faster gain compresses the gap and $|\Delta \log A_{\text{AI}}|$ falls.

The shock magnitude is therefore governed by where the frontier and substitute curves sit relative to the (fixed) AEI task-hour distribution, not by the absolute level of frontier capability alone. Anchoring at the most recent Opus reading softens the headline shock; anchoring at a lower point on the cluster of plausible frontier values deepens it.

## C.2 Four Anchor Candidates

Table 16 reports four anchor candidates evaluated under the same calibration used in the body, with the doubling lag held at the measured central of 2.27 and the coverage integrals taken against the same AEI 1P API task-hour distribution. The first candidate is the measured pair: Claude Opus 4.5 at $h_{50}^F = 4.63$ hours, the METR v1.0 reading used to identify the 2.27-doubling lag against Kimi K2 Thinking. The second is the chosen cluster median, $h_{50}^F = 6.40$ hours, the v1.1 median of the top US cluster. The third and fourth use Claude Opus 4.6 as the single best US model, at its METR Time-Horizon 1.1 score ($h_{50}^F = 11.98$ hours) and at its latest rescore ($h_{50}^F = 14.50$ hours).

The full anchor sensitivity range is $\Delta \log A_{\text{AI}} \in [-0.59, -0.32]$, of width 0.27. This is roughly half the width of the doubling-lag convergence–divergence interval $[-0.71, -0.18]$ reported in Section 4.2, and a non-trivial fraction of the central shock itself.

## C.3 The Case for Each Candidate

**Opus 4.5 (measured pair).** This is the internally consistent anchor: the 2.27-doubling lag was identified using Opus 4.5 ($h_{50} = 277.7$ minutes) against Kimi K2 Thinking ($h_{50} = 57.6$ minutes), both METR v1.0, November 2025 (METR, 2025b). Anchoring the frontier at Opus 4.5 keeps frontier and substitute as the exact pair from which the lag was measured, with no cross-version extrapolation — the cross-version hazard explicitly flagged in the footnote to Section 4.2. The cost of the choice is that it delivers the deepest shock in the candidate set ($-0.59$), and that anchoring on a single model leaves the calibra-

**Table 16:** Estimator A: Anchor Sensitivity, Doubling Lag Fixed at 2.27

| Anchor | $h_{50}^F$ (hours) | $h_{80}^F$ (hours) | $h_{50}^N$ (hours) | $\theta_F$ | $\theta_N$ | $\kappa_A$ | $\Delta \log A_{\text{AI}}$ |
|---|---|---|---|---|---|---|---|
| Opus 4.5 (measured pair, v1.0) | 4.63 | 0.78 | 0.96 | 0.624 | 0.347 | 0.555 | −0.59 |
| Cluster median (chosen) | 6.40 | 1.16 | 1.33 | 0.684 | 0.398 | 0.582 | −0.54 |
| Opus 4.6 (TH1.1) | 11.98 | 1.16 | 2.49 | 0.720 | 0.511 | 0.710 | −0.34 |
| Opus 4.6 (latest rescore) | 14.50 | 1.40 | 3.01 | 0.741 | 0.538 | 0.726 | −0.32 |

*Notes:* Each row reports the reliability-curve calibration under a single anchor choice, with the doubling lag fixed at the measured central of 2.27 (METR, 2025b). The implied substitute anchor is $h_{50}^N = h_{50}^F \cdot 2^{-2.27}$. Coverage values $\theta_F$ and $\theta_N$ are evaluated against the same AEI 1P API task-hour distribution used in Section 4.2 (Anthropic, 2026a). The cluster median is the central row carried into the body. The Opus 4.6 latest-rescore $h_{80}^F$ is approximated; the other three are read directly off METR releases (METR, 2026). Full anchor sensitivity range: $\Delta \log A_{\text{AI}} \in [-0.59, -0.32]$.

tion sensitive to that one model's estimate.

**Cluster median (chosen).** The median of the top US cluster on METR Time-Horizon 1.1, $h_{50}^F \approx 6.40$ hours, is robust to single-estimate noise and is the level the body uses. It avoids the question of which single model to privilege, and — as Table 16 shows — sits at the deep end of the candidate range, deeper than both Opus 4.6 candidates. It is the more conservative-on-the-headline choice relative to the two Opus 4.6 anchors. The cost is that it mixes a v1.1 anchor with a v1.0 lag; in this respect the cluster median is the less internally consistent of the two leading candidates, but the implied difference in $\Delta \log A_{\text{AI}}$ relative to the v1.0 alternative is small (0.05 log points).

**Opus 4.6 (TH1.1).** The current single best US model on METR Time-Horizon 1.1, at $h_{50}^F = 11.98$ hours. Two issues. First, the model's $h_{50}$ carries a wide METR 95% confidence interval, roughly 5.7 to 97 hours, because the task-length suite contains too few long tasks to identify a model at the frontier of capability with precision (METR, 2026). Second, anchoring here under the fixed 2.27 lag implicitly places the substitute at $h_{50}^N \approx 2.49$ hours — materially more capable than today's best open-weight model (Kimi K2 Thinking at $h_{50} \approx 1$ hour), so the choice silently extrapolates substitute capability forward as well.

**Opus 4.6 (latest rescore).** The highest single-model anchor in the candidate set, at $h_{50}^F = 14.50$ hours. It carries the same noise-sensitivity concern as the TH1.1 candidate, places the substitute even further forward ($h_{50}^N \approx 3.01$ hours), and produces the mildest shock in the set (−0.32). The associated $h_{80}^F$ is approximated rather than directly measured.

## C.4 Decision

The body retains the cluster median, $h_{50}^{F} = 6.40$ hours and $h_{80}^{F} = 1.16$ hours, delivering $\Delta \log A_{\text{AI}} = -0.54$. The choice is robust to single-estimate noise in any one model and sits at the deep end of the anchor sensitivity range $[-0.59, -0.32]$ — deeper than three of the four candidates — so the headline cannot be characterized as a soft reading. Relative to the internally consistent Opus 4.5 alternative ($\Delta \log A_{\text{AI}} = -0.59$), the cluster median is the more conservative-on-the-headline choice; the gap between them (0.05 log points) is small relative to the doubling-lag uncertainty reported in Section 4.2.

The headline result of Section 4.2 is reported under the chosen anchor; the figures of this appendix are the bracket within which the headline sits when the anchor is left free. Anchor variation moves $\Delta \log A_{\text{AI}}$ across $[-0.59, -0.32]$, comparable to roughly half the doubling-lag convergence–divergence interval, and is the single highest-leverage discretionary choice in Estimator A.