PIIE PETERSON INSTITUTE FOR INTERNATIONAL ECONOMICS

POLICY BRIEF

26-7 Where Is AI in GDP Statistics?

Filling the Measurement Gap

Anton Korinek and Patrick McKelvey

May 2026

The artificial intelligence (AI) economy in the United States is growing at extraordinary rates of over 2,000 percent per year yet is leaving only a small mark in the nation's GDP figures. This is a measurement gap that, left unaddressed, will become a policy gap—because what cannot be measured cannot be steered.

We estimate in a companion PIIE Working Paper (Korinek and McKelvey 2026) that nominal AI compute spending grew by more than 140 percent per year each in 2024 and 2025, raw compute capacity by more than 200 percent per year, and quality-adjusted AI output by more than 2,000 percent per year. The divergence between this picture of the AI economy and the one drawn by conventional GDP statistics is itself an informative macroeconomic signal. Treating the AI sector as a coherent economic entity in its own right yields a preliminary estimate of nominal AI GDP of roughly $250 billion in 2025—comparable in size to the US scheduled passenger airline industry (Bureau of Transportation Statistics 2025)—yet growing at approximately 2,600 percent per year in quality-adjusted terms.

We argue that US statistical agencies and economic policymakers should start now to assemble better data on AI activity in AI satellite accounts—focused subsets of the national economic statistical accounts—in coordination with international counterparts, industry and researchers. They should begin to

Anton Korinek has been a nonresident senior fellow at the Peterson Institute for International Economics (PIIE) since February 2026 and head of Transformative AI Economic Studies at the Anthropic Institute since May 5, 2026. He is currently on educational leave from the University of Virginia, where he is professor of economics and faculty director of the Economics of Transformative AI (EconTAI) Initiative. This work was conducted in his capacity as a nonresident senior fellow at PIIE and professor at the University of Virginia.

Patrick McKelvey is a senior data scientist at the Bank of Canada.

The authors thank Kody Karmody and Dylan Ryfe for excellent research assistance, Leopold Brown and Yuval Rhymon for their contributions to early-stage research and data collection, Future Impact Group for their support, and Andrey Fradkin for generously sharing data on inference prices.

incorporate AI productive-capacity measures into medium-term projections and scenario analysis. Building this measurement infrastructure today, while the AI sector is still small in nominal terms, is needed preparation for a potential phase change in the future after which policymakers may need a strong statistical apparatus to make well-informed decisions about the fast-growing AI economy.

TWO PICTURES OF THE SAME ECONOMY

The question of where AI lies in the US GDP statistics has become a recurring mystery in economic commentary. Frontier AI capabilities are advancing at what industry observers consider a remarkable pace, with some seeing the possibility of artificial general intelligence within just a few years. Yet when one looks at the conventional national statistical accounts, the AI revolution registers only as upstream investment—a data-center capital investment boom—while the productive activity those data centers support remains nearly invisible. Overall US GDP growth remains moderate, productivity statistics have barely ticked up, and the disconnect between widely reported AI capability gains and their macroeconomic effects has become a source of puzzlement.

One natural explanation is that AI adoption takes time, and the kinds of broad productivity gains economists associate with general purpose technologies typically arrive years after the technology itself does. This is almost certainly part of what is going on. But there is a second explanation, complementary rather than competing, that has received much less attention: national accounts as currently constructed have a hard time seeing the AI sector at all.

The visibility problem is structural. The conceptual architecture of GDP measurement was developed in the mid-20th century to track an economy organized around manufacturing. That architecture has served well for a long time and continues to do so for the bulk of economic activity. But it presumes that the structure of the economy changes only slowly, and that quality improvements in any given sector unfold at a pace that statistical agencies can easily capture. AI strains that presumption. Quality is improving so quickly that conventional hedonic adjustment methods, designed for sectors where quality improvements occur at a more moderate pace, do not capture what is happening. Moreover, AI-related activity is dispersed across a long list of industries—cloud services, software publishing, data processing, professional services, and others—with no single category that captures the AI economy as a whole.

AI is the latest in a series of fast-moving technologies that have raised measurement concerns; semiconductors and the internet generated similar debates in their time. But there is a feature of the AI case that distinguishes it from those precedents and that may make the measurement question much more consequential. In the prior episodes, the rapidly improving technology was a *complement* to human labor at the aggregate level: better chips made workers and equipment more productive; free digital services raised welfare in ways that flowed through human time and attention. Because the gains had to pass through a human bottleneck, their macroeconomic footprint was bounded—part of why careful researchers ultimately judged the resulting mismeasurement to be modest. AI is the first plausible candidate for large-scale technological mismeasurement in which the rapidly improving sector may become a *substitute* for human labor. If that possibility is realized, the bound that disciplined prior episodes will not apply, and the measurement gap could become much more consequential than past gaps proved to be.

These challenges are with us now in modest form, but the case for confronting them rests on a forward-looking observation: statistical infrastructure takes years to build, and once measurement gaps become acute, the cost of not having addressed them earlier is difficult to recover.

WHAT DIRECT MEASUREMENT SHOWS

In our companion Working Paper (Korinek and McKelvey 2026), we develop direct estimates of US AI production by combining several data streams: data-center electricity usage and chip-stock characteristics, prevailing GPU rental rates, AI inference prices observed at fixed performance levels, and the pace of algorithmic progress in training. The methodology produces three layers of measurement, each capturing a different aspect of how the AI economy is expanding.

Nominal compute spending. US AI compute spending—measured on an imputed-rental basis using prevailing graphics processing unit (GPU) rates—rose from $37 billion in 2023 to $90 billion in 2024 and $219 billion in 2025, annual growth rates of roughly 145 percent in each of the past two years. A separate cross-check based on global chip-sales data, drawing on largely independent inputs, points to growth rates that are similar in order of magnitude.

Raw compute capacity. As chips became more efficient, each dollar of compute spending bought more physical computing capacity. Measured in H100-equivalent units, US AI computing capacity grew at more than 200 percent per year, outpacing nominal spending.

Quality-adjusted AI output. An even larger growth rate becomes visible once one accounts for algorithmic progress. Inference prices at fixed benchmark performance fell by roughly 94 percent per year over our sample, and Ho et al. (2024) estimate that the compute required to train a model at fixed performance falls by about two thirds annually. Combined, these efficiency gains imply that quality-adjusted AI output grew at roughly 2,290 percent in 2024 and 2,271 percent in 2025 (see table 1 below). Of the three exponentials behind these figures, algorithmic progress is the largest over the observed period.

Table 1

US AI production estimates, annual averages

Year	Nominal compute spending (billions of dollars)	Spending growth	Compute (H100-equivalent units)	Compute growth	Quality-adjusted output growth
2023	36.9	—	1.09M	—	—
2024	90.5	145%	3.41M	212%	2,290%
2025	219.2	142%	10.7M	214%	2,271%

Source: Korinek and McKelvey (2026).

What "quality-adjusted" means here

Quality-adjusted inference output measures the AI services delivered to users—chiefly inference tokens—adjusted so that growth reflects both more tokens produced and higher capability per token. Quality-adjusted training output measures investment in new AI model capital from compute used in research and development, adjusted for the fact that a dollar of training compute now produces more capable models than it did in prior years. Quality-adjusted AI output aggregates the two using their nominal shares, providing a single index of AI sector growth (see figure 1 below).

Figure 1

The cost of using AI models has declined at about 94 percent per year, reflecting falling prices for given AI capability levels

Chained Fisher price index (log scale) for AI inference token prices, 2023–26 (uniform weights across intelligence bins)

Source: Authors' calculations.

Two contributions, two ambitions. It is useful to be explicit about how these measures relate to standard national accounting practice, since they make two related but distinct contributions. The first contribution sits *within* the existing conceptual framework of national accounts. Tracking AI as a coherent sector—by aggregating nominal compute spending, raw compute capacity, and quality-adjusted output across the dispersed industry codes that AI activity falls under—requires no change to headline GDP methodology. It is the same kind of work that produces tradable-sector or energy-sector accounts. More aggressive hedonic adjustment is similarly continuous with existing practice: statistical agencies routinely apply quality adjustments to fast-moving technology sectors, and the question for AI is not whether to do this but at what magnitude. Byrne, Oliner, and Sichel (2018) found that even within the familiar semiconductor sector, official deflators substantially understated the pace of quality-adjusted price decline; the AI case may pose the same challenge in a more acute form. Both elements together support the case for an *AI satellite account* within the existing system of national accounts.

The second contribution is more ambitious and more speculative. To put the production estimates in context relative to the overall economy, the Working Paper develops a framework that treats the AI sector as a coherent quasi-economic entity in its own right—partitioning value creation according to whether it is more closely associated with AI computation or with human brain activity, and computing a corresponding *AI GDP*. Under simple assumptions about gross margins, services revenue, and labor inputs, we estimate nominal AI GDP at roughly $250 billion in 2025, growing at approximately 2,600 percent per year in quality-adjusted terms. We present this framework as preliminary rather than settled. It represents a more fundamental departure from current accounting conventions and is most informative as a forward-looking exercise: a way to track AI's net contribution to the economy on its own terms, useful in proportion to how much AI's economic role grows beyond the categories that contain it today.

Implications for total GDP. As an illustrative exercise, one can ask what aggregate US GDP growth would have looked like if the AI portion of the economy had been deflated using our quality-adjusted price index rather than the conventional deflator implicit in current statistics. The answer is that real GDP growth would have come in higher by roughly 2 percentage points in 2024 and roughly 4 percentage points in 2025. We present this as an upper bound rather than a point estimate, for two reasons. First, AI inference is overwhelmingly an intermediate input, not a final good, and the production function that maps improved inference quality into additional final output is not well understood—the observed sharp improvement in token quality did not translate into a proportional increase in spending that would have shown up in final GDP. Second, our quality adjustments are anchored to benchmark performance rather than to directly observed economic value. Even with these caveats, the exercise illustrates how sensitive headline statistics are to deflator methodology in the AI sector, and how rapidly the gap could widen if quality-adjusted output continues to grow at recent rates.

WHY IT IS HARD TO SEE AI IN OVERALL GDP

That AI's growing footprint is so faintly visible in national GDP statistics has, in part, a straightforward accounting explanation. Nominal AI revenues grow only moderately because per-unit prices for any given level of AI capability fall almost as fast as quality-adjusted output rises. In the semiconductor industry, this pattern played out for decades: each generation of chips was dramatically cheaper per unit of performance than the last, so the semiconductor share of GDP remained modest even as quality-adjusted output expanded enormously. AI today is, by this metric, an even more pronounced version of the same phenomenon.

A useful framing is what Bhagwati (1958) called *immiserizing growth*—a sector or country whose output expands rapidly but whose terms of trade deteriorate so quickly that the gains do not register in market value. AI services have been on something close to that trajectory, with massive increases in the volume and quality of output offset by collapsing per-unit prices. Whether this dynamic of declining prices continues is an open question; recent firmness in GPU rental pricing and the growth of large enterprise commitments to frontier AI services may attenuate it going forward.

A second source of two-sided uncertainty involves consumer surplus. As with prior generations of digital services—including free or ad-supported tools that the Bureau of Economic Analysis has begun to track within a Digital Economy Satellite Account, and as analyzed in the new-goods literature pioneered by Hausman and in Brynjolfsson et al. (2025) on "GDP-B"—significant value from AI may accrue to consumers in ways that recorded revenues do not capture. Conversely, the capability improvements that we attribute to producers as quality gains may pass through to consumers rather than register as additional output in any conventional sense. Both forces complicate the translation between our quality-adjusted production measures and standard welfare statistics.

It is worth recalling the history of measurement debates around prior fast-moving technology sectors here. A substantial literature has asked whether the productivity slowdown of the past two decades reflects, in part, statistical failures to capture the value created by the internet and adjacent digital services. Byrne, Fernald, and Reinsdorf (2016) and Syverson (2017) concluded, after careful examination, that mismeasurement in those cases was real but too small—and too uncorrelated with the cross-country patterns of the slowdown—to explain the puzzle. We take that conclusion seriously. Our argument here is not that AI mismeasurement today is large enough to materially shift headline numbers but that it may become so within a small number of years.

The reason we think the AI case may eventually diverge from the semiconductor and internet comes back to the complementarity point introduced earlier. The technologies whose mismeasurement was carefully assessed in those debates were ultimately complements to human labor and consumption, and their aggregate footprint was bounded by the size of the activities they enabled. AI is the first plausible candidate for large-scale technological mismeasurement in which the improving sector is a potential substitute for labor itself. If AI capabilities continue to broaden—first across cognitive work, and eventually, with more capable robotics, across physical work—the human bottleneck that disciplined prior episodes is precisely what begins to give way. As AI services become substitutable for tasks previously performed by people, the relevant price comparison shifts: not AI versus prior-generation AI (where prices have been falling rapidly), but AI versus prior human wages (which are considerably more expensive). At that point the rapid decline in inference prices may slow or reverse, and the productive capacity that has been almost invisible in nominal terms may begin to register in headline statistics, and potentially abruptly so. Building measurement infrastructure now is, among other things, preparation for the possibility of such phase changes.

FROM MEASUREMENT GAP TO POLICY GAP

The case for closing the measurement gap is sharpest in the medium-term planning horizons that fiscal and monetary authorities are responsible for. Two policy domains illustrate the stakes.

First, income and payroll taxes anchor the revenue base of every advanced economy. The central fiscal question raised by AI (Korinek and Lockwood 2026) is whether the wage tax base will erode and what, if anything, can replace it. Answering that question requires tracking the AI sector's productive capacity. Conventional statistics show a sector growing slowly in nominal terms; our measures show one whose underlying capacity is more than doubling annually. A finance ministry running ten-year revenue projections off the conventional data will materially underweight the probability of a labor-tax-base shock—and will be correspondingly unprepared to design responses such as tax system reforms, sovereign wealth funds, or other benefit-sharing schemes that such a shock may call for. A windfall that cannot be seen cannot be shared. The point generalizes beyond revenue forecasting: proposals to tax AI directly—whether through a compute tax, an excise on AI services, or a windfall levy on frontier-model providers—presupposes a quantitative picture of the base being taxed. Right now, official statistics cannot tell policymakers what a given AI tax instrument would raise. Even recent industrial-policy proposals from within the AI industry presuppose a quantitative picture of the AI economy that current statistics simply do not provide.

Second, economic policymakers shape policy on the basis of measured output gaps, productivity, and the natural rate of interest, all inferred from national accounts that struggle to see the AI sector. As AI investment absorbs an increasing share of real resources—electricity, capital expenditure, and skilled labor—the natural rate of interest is likely to be moving, but the signal in conventional statistics will be muted and lagged. The challenge sharpens further if and when AI becomes a close enough substitute for labor that its output prices stop falling at the current pace and the productive capacity that has been growing invisibly in nominal terms may begin to register in headline statistics, potentially abruptly. A policymaker who has been reading a low-growth, disinflationary economy may find itself behind an economic phase transition it had no statistical apparatus to anticipate. The cost of investing in measurement infrastructure now is small relative to the cost of being caught flat-footed at exactly the moment when getting policy right matters most.

THREE RECOMMENDATIONS

These observations suggest three priorities for the policy and statistical communities that we outline in the following. The right design will need to evolve as data improve and as the AI economy's contours become clearer.

First, statistical agencies, in coordination with the OECD, the IMF, and counterparts in other advanced economies, should begin developing AI satellite accounts. Several of the measures discussed above—nominal compute spending, raw compute capacity, and quality-adjusted output indices—could serve as initial components, providing structured visibility into AI activity without requiring changes to headline GDP methodology. Detailed architecture can be refined as data improve; what matters is that the work begin before the gap becomes acute. The capitalization of R&D as investment, eventually adopted in the 2008 System of National Accounts (SNA) revision, took decades from initial proposal to implementation. The growth rates we estimate suggest the window for AI is considerably shorter.

Second, better data on the AI economy could be assembled through structured collaboration between statistical agencies, AI firms, and academic researchers. Gaps in current measurement—including the allocation of compute between training and inference, AI providers' gross margins, and the relationship between benchmark performance and economic value—cannot be closed through external estimation alone. Targeted disclosure protocols, harmonized definitions across firms, and academic-industry data-sharing arrangements would each narrow these gaps materially. Other industries provide precedents for this kind of structured cooperation, and the case for adopting one for AI grows with the sector itself.

Third, economic policymakers should incorporate AI productive-capacity measures into their medium-term projections and scenario analysis, even before satellite accounts are fully constructed. Imperfect measures of AI productive capacity are far more informative than the implicit assumption embedded in conventional projections—that the AI sector's productive capacity is small and slow-growing. Fiscal authorities could use such measures to stress-test projections about the labor tax base; central banks could use them in assessing how rapidly AI investment is absorbing real resources and in scenario analysis covering the regime change discussed above.

LOOKING FORWARD

For now, standard national accounts continue to do well what they were designed to do: track the pace of improvement in human-experienced economic welfare. Direct measurement of the AI economy answers a different and increasingly important question—how rapidly the AI sector's productive capacity is expanding, and what that expansion may imply for the broader economy in the near future. Both questions matter. The divergence between the two answers may itself become one of the most informative macroeconomic indicators of the next several years.

The estimates presented here are a first approximation. The underlying assumptions can be refined as better data become available, and the AI GDP framework will require additional development before it is ready for routine policy use. But the case for treating AI measurement as a current priority—rather than something to revisit when the AI economy is "large enough to matter"—strengthens with each passing quarter. Building this measurement infrastructure today, while the AI sector is still small in nominal terms, is the most cost-effective path to having usable statistics at the moment they are most needed.

REFERENCES

Bhagwati, J. 1958. Immiserizing Growth: A Geometrical Note. *Review of Economic Studies* 25, no. 3: 201–205.

Bureau of Transportation Statistics. 2025. U.S. airlines net profit was $1.6 billion in third quarter 2025, a decrease in profit over third quarter 2024. Available at https://www.bts.gov/newsroom/us-airlines-net-profit-was-16-billion-third-quarter-2025-decrease-profit-over-third (accessed on April 27, 2026).

Brynjolfsson, E., A. Collis, W. E. Diewert, F. Eggers, and K. J. Fox. 2025. GDP-B: Accounting for the Value of New and Free Goods. *American Economic Journal: Macroeconomics* 17, no. 4: 312–44.

Byrne, D. M., J. G. Fernald, and M. B. Reinsdorf. 2016. Does the United States Have a Productivity Slowdown or a Measurement Problem? *Brookings Papers on Economic Activity* 2016, no. 1: 109–182.

Byrne, D. M., S. D. Oliner, and D. E. Sichel. 2018. How Fast Are Semiconductor Prices Falling? *Review of Income and Wealth* 64, no. 3: 679–702.

Demirer, M., A. Fradkin, B. Ifrach, N. Tadelis, and S. Peng. 2025. *The Emerging Market for Intelligence: Pricing, Supply, and Demand for LLMs*. NBER Working Paper 34608. Cambridge, MA: National Bureau of Economic Research.

Ho, A., T. Besiroglu, E. Erdil, D. Owen, R. Rahman, Z. C. Guo, D. Atkinson, N. Thompson, and J. Sevilla. 2024. Algorithmic Progress in Language Models. Available at https://arxiv.org/abs/2403.05812 (accessed on May 8, 2026).

Korinek, A., and P. McKelvey. 2026. *Measuring the AI Economy*. PIIE Working Paper 26-9. Washington: Peterson Institute for International Economics.

Korinek, A., and L. Lockwood. 2026. Public Finance in the Age of AI: A Primer. Forthcoming in Agrawal et al. (2026), *The Economics of Transformative AI*, NBER and University of Chicago Press.

McKelvey, P., L. Brown, Y. Rhymon, and A. Korinek. 2026. *Measuring the AI Economy: Conceptual Framework and Implementation*. Technical Appendix in PIIE Working Paper 26-9. Washington: Peterson Institute for International Economics.

Patel, D., D. Nishball, and J. Eliahou Ontiveros. 2024. AI Datacenter Energy Dilemma—Race for AI Datacenter Space. SemiAnalysis.

Sevilla, J., H. Petrovic, and A. Ho. 2026. Can AI Companies Become Profitable? Epoch AI Gradient Updates.

Syverson, C. 2017. Challenges to Mismeasurement Explanations for the US Productivity Slowdown. *Journal of Economic Perspectives* 31, no. 2: 165–186.

Trammell, P., and A. Korinek. 2026. Economic Growth Under Transformative AI. *Annual Review of Economics*, forthcoming.

PIIE PETERSON INSTITUTE FOR INTERNATIONAL ECONOMICS

The findings, interpretations, views, and conclusions expressed herein are solely those of the authors and do not necessarily represent those of the Bank of Canada, the Anthropic Institute, or the Peterson Institute for International Economics. The authors have used AI extensively at every stage of the research and writing process and have subjected all AI-generated output to careful human review.