

# AI Spend vs. Benefit:

## Marking our views to market two years later

---

**Revisiting June 2024's "Top of Mind: GenAI:  
too much spend, too little benefit?"**

April 2026

James Covello

Goldman, Sachs & Co.

1-212-902-1918

[james.covello@gs.com](mailto:james.covello@gs.com)

Goldman Sachs does and seeks to do business with companies covered in its research reports. As a result, investors should be aware that the firm may have a conflict of interest that could affect the objectivity of this report. Investors should consider this report as only a single factor in making their investment decision. For Reg AC certification and other important disclosures, see the Disclosure Appendix, or go to [www.gs.com/research/hedge.html](http://www.gs.com/research/hedge.html). Analysts employed by non-US affiliates are not registered/qualified as research analysts with FINRA in the U.S.

---

# What's Inside

## Table of contents

---

- 4 Where have our views proven to be wrong?
  - Consumer adoption has been much more robust than we expected.
  - Despite underperforming the market, hyperscalers have continued to raise capex.
  
- 7 Where do we have more conviction now?
  - Despite robust enterprise adoption, enterprise ROI is proving challenging.
  - The economics of AI continue to be very challenging for all but the semi companies.
  - There are jobs being replaced by AI but nowhere near the level that the media or many Tech CEOs predicted.
  - SLMs are having more impact in the enterprise than LLMs.
  
- 16 Where is value ultimately going to accrue in the AI supply chain?
  - Is the concentration of value within semis sustainable?
  - What are the most interesting companies we have met in the last two years working on?
  - What are the key components to unlock the economics in the enterprise?
  
- 21 What profit pools are actually being disrupted?
  - Is software really at as much risk as the market currently assumes?
  
- 27 What advice would we give to C-suites on AI?
  - Ensure FOMO isn't forcing your company to rush an AI solution that isn't the right long-term answer
  - Remember that in the internet era, the pioneers got the arrows and the settlers got the land.
  - How different is having hundreds of disparate SLMs in an enterprise than the SaaS applications that they were meant to replace?
  
- 29 How would we invest in AI from here?
  - Go long the hyperscalers vs. underweight semis.

Thanks to **Milos Raonic** for his contribution to this report.

---

# Revisiting the AI Spend vs. Benefit Debate

Marking our views to market two years after our original report

---

In June 2024, we collaborated with our research colleagues on a “Top of Mind” report led by Allison Nathan in which we posed the question, “Gen AI: Too Much Spend, Too Little Benefit?” In the nearly two years since, there’s been endless (and divided) debate about AI’s potential, with views ranging from many tech executives suggesting that it can replace a large portion of knowledge workers to some investors’ views that AI is a bubble headed for a dot-com-like collapse. Needless to say, there has been, and remains, a very large gap between those views.

In this report (which we are publishing in slide format), we update our thoughts on AI and the spend vs. benefit debate. We **mark to market our current views** vs. our original views and, in doing so, highlight **where we have been wrong** and **where we have even more conviction** than before.

Importantly, while we believe it’s important to mark the original views to market (especially where we have been wrong), we think it’s much more important to look forward to where the technology is going and what challenges need to be addressed for AI to be as successful as many predict.

The key question in our mind that needs to be answered for investors to understand how and where to invest going forward is “where is the value going to accrue in the supply chain?” Will value accrue to the enterprise layer, the model layer, the hyperscaler layer, or the chip layer? In this report, we discuss **an emerging layer that we believe will be critical to unlocking the economics of AI in the enterprise**. Models will continue to improve at a fast pace, but, right now, model capability is not what’s holding back successful enterprise use cases. We believe data structure and orchestration are critical factors to unlock AI in the enterprise.

To date, nearly all the value has accrued to the chip layer, which is both unprecedented and unsustainable. We covered semi companies for nearly 20 years, and the general idea is that chip companies are supposed to thrive when their customers thrive. In this cycle, the chip companies are thriving at the expense of everyone above them in the chain. Something has to change with this dynamic: either the companies higher in the chain need to start earning a return on investment or they will eventually need to spend less on the chips that are powering this build. While much has changed over the last two years, this dynamic hasn’t really changed at all.

In addition to trying to answer these key questions, we also **offer advice to the C-suite** on how to think about AI and conclude with some thoughts on **how we believe investors can make money** from here.



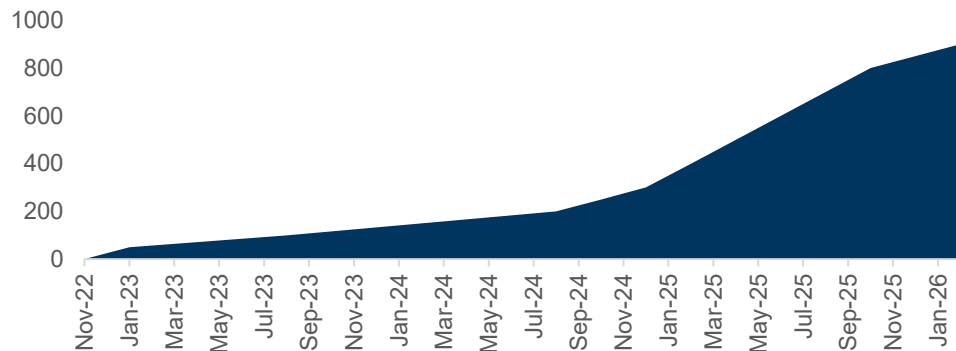
# Where Have Our Views Proven Wrong?

- Consumer adoption of AI has been much more robust than we expected
- Despite hyperscalers' stocks underperforming the market, those companies have raised their capex instead of cutting as we expected

# Consumer AI Adoption Has Been Much Greater Than Our Expectations

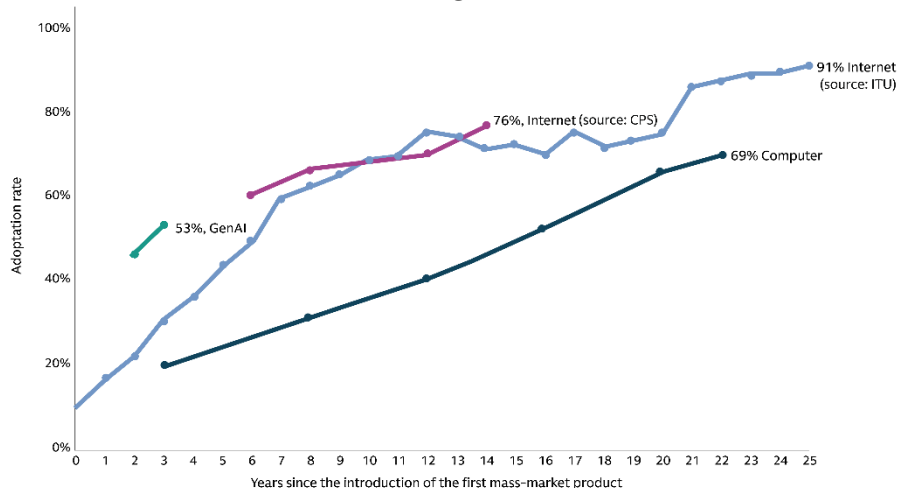
Heavy reliance on free tiers emphasizes need for enterprise success

## Open AI WAU (millions)



Source: Axios, Reuters, Business Today, Verge, Twitter, Forbes, TechCrunch, OpenAI<sup>1</sup>

## Speed of AI adoption by technology



Source: Generative AI Adoption Tracker

Per a recent Harmony study<sup>2</sup>, Americans spend an average of 5 hours and 16 minutes per day on their phone. Only 11% of survey respondents felt their phone time was productive. So, it has been surprising to us that a group that admittedly spends over 4 unproductive hours a day on their phones are so interested in saving approximately 30 minutes per day leveraging AI according to a recent Fortune article<sup>3</sup>.

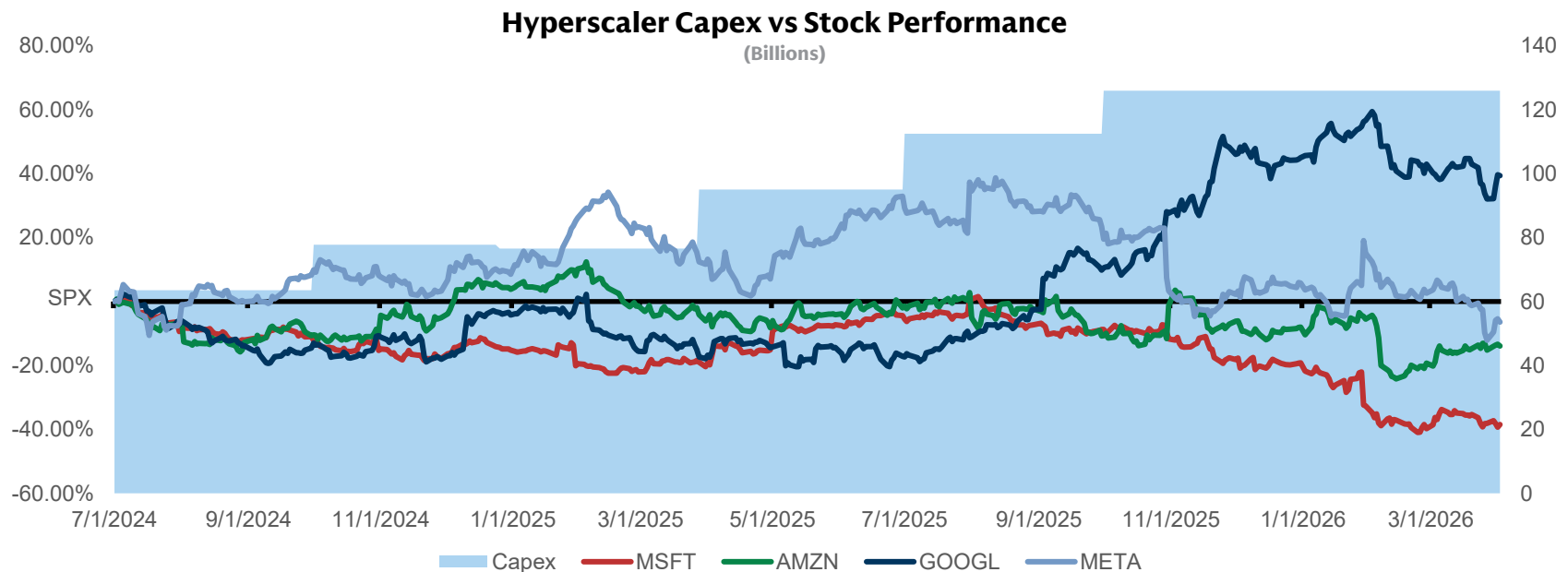
We have a hard time reconciling the excitement around the time saved vs. the unproductive time doom scrolling. Regardless, we got this part wrong as consumer adoption has been spectacular.

The Stanford Institute of Human-Centered AI (HAI)<sup>4</sup> showed how GenAI's adoption has been rapid in the years after its mass market introduction. Measured from the release of each technology's first widely available product, generative AI reached approximately 53% adoption within three years, well above the initial trajectories of the personal computer and the internet over comparable time frames.

While we were too conservative in our assumptions around consumer adoption, **the fact that 95% of weekly active users are using a free version of GPT**, as reported by The Information, **underscores a broader point we have made all along**. That is, enterprise adoption and success is the critical factor that will drive AI economics and ROI for the supply chain.

# Hyperscalers Surprised Us by Raising Spending Despite Stocks Lagging

We believe that deteriorating cash flow profiles will eventually force a slowdown



Source: Goldman Sachs Global Investment Research

**We had incorrectly assumed that if the hyperscaler stocks underperformed the market for a sustained period, they would decrease their spending on AI capex. Instead, they have significantly increased their spending despite relatively poor stock performance.**

FOMO has proven a stronger incentive than poor stock performance as hyperscalers have prioritized being involved in the AI arms race over their current shareholders. Obviously, the hyperscalers believe this spending is in the best interests of their long-term shareholders. We spend a good portion of this report exploring what has to happen in order for the hyperscalers to earn an ROI.

Eventually, we believe the reality that this capex spending is impacting the previously outstanding cash flow profile of these companies will drive lower spending; but for sure, two years later, we have been wrong on this dynamic.



Q.

# Where Do We Have More Conviction Now?

- Despite robust enterprise adoption, enterprise ROI is proving challenging
- The economics of AI continue to be very challenging for all but the semi companies
- There are jobs being replaced by AI but nowhere near the level that the media or many Tech CEOs predicted
- SLMs are having more impact in the enterprise than LLMs

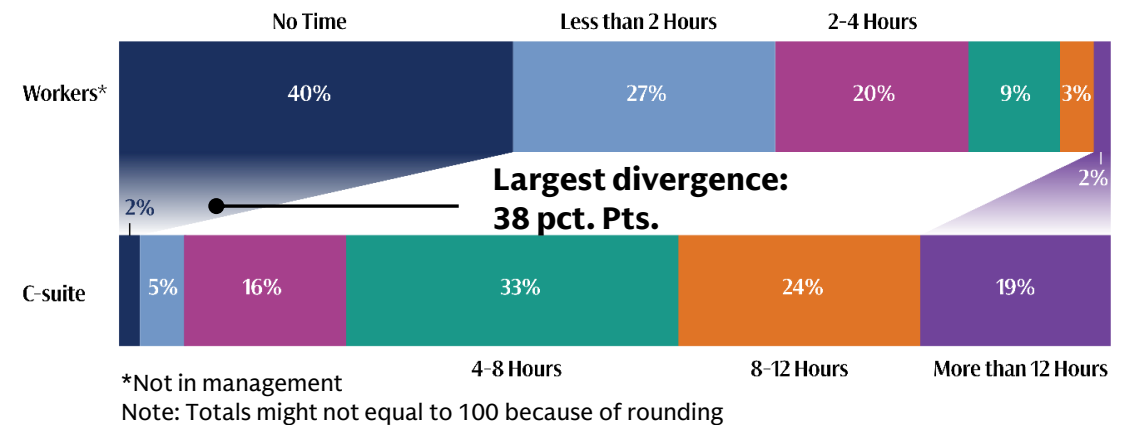
# Despite Significant Spending, Enterprise AI Results Are Mixed at Best

Line worker surveys do not match the enthusiasm of C-suites and the media

There are many concrete examples of successful use cases in the enterprise, particularly for coders and developers. There are also an emerging number of high value add SLM implementations which we discuss on slides 14-15. However, as we highlight below, in a number of recent studies, there is a very significant disconnect between what C-suites and Tech execs are saying about AI efficiencies and the “realities on the ground.” We discuss later in the report what needs to happen in order to unlock AI effectiveness and economics but it’s clear that currently, there is a mismatch between expectations and reality.

The Wall Street Journal published a survey showing the significant difference between how workers are perceiving the productivity boost from AI vs the C-suite.

**How much time do you think you are saving each week by using AI?**



Source: Lindsay Ellis, Stephanie Stamm, Wall Street Journal<sup>5</sup>

“ Mercor, an AI hiring startup, tested several AI agents powered by top-tier frontier models on 480 workplace tasks frequently carried out by human bankers, consultants, and lawyers. Every agent they tested failed to complete most of its duties. – Mercor<sup>6</sup>

“ 56% of Americans say that they use AI, yet 85% of the workforce does not have a value-driving AI use case.” – Section AI<sup>7</sup>

“ Our survey found that revenue growth has been the least important metric for AI investment to date.” – Economist Impact Report<sup>8</sup>

“ Less than a third of knowledge workers report saving 4+ hours a week with AI – when most organizations should be targeting a 10+ hour time savings per employee to generate ROI.” – Section AI

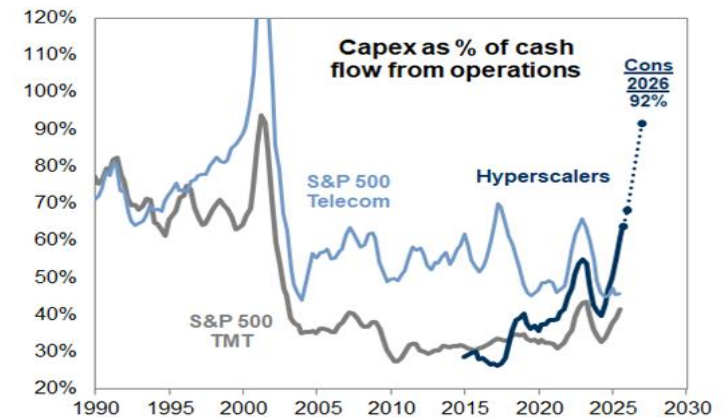
“ Employees reported spending an average of one hour and 56 minutes dealing with each instance of workslop. These incidents carry an invisible tax of \$186 per month. For an organization of 10,000 workers, given the estimated prevalence of workslop (41%), this yields over \$9 million per year in lost productivity.” – Harvard Business Review<sup>9</sup>

# AI Economics Are Challenging for Everyone Except the Semi Companies

Enterprise buyers, model companies and hyperscalers have yet to show returns on spending

- Most enterprises are losing money implementing AI:
  - An MIT Labs report<sup>10</sup> showed that despite \$30–40 billion in enterprise investment into GenAI, a surprising **95% of organizations are getting zero return**.
  - Axios reported that IT budgets are getting blown out as some companies spend more on AI than on employees' salaries... Maybe human labor will be more cost efficient after all.<sup>11</sup>
  - An Ernst & Young survey<sup>12</sup> shows that **almost every company in their survey (99%) reported financial losses from AI-related risks** (for example bias in outputs), and 64% experienced losses exceeding US\$1 million. On average, the financial loss to companies that have experienced risks is conservatively estimated at US\$4.4 million. That's an estimated total loss of US\$4.3 billion across the 975 respondents in their sample.
  - **IT spend to grow from \$5tn in 2024<sup>13</sup> to \$6.15tn in 2026<sup>14</sup>** according to Gartner (i.e., we aren't seeing budget savings on the aggregate level).
- The cash burn issues with the foundational model companies are well reported by the press and well understood by the markets.
- The hyperscaler companies have burned through all of their cash flow from operations and are now borrowing to fund the buildout:
  - In 2025 data center debt issuance doubled to \$182bn<sup>15</sup>
  - Original bull case on the AI buildout was “this is different than the internet buildout because the companies spending the money have it to burn.” That is no longer the case.

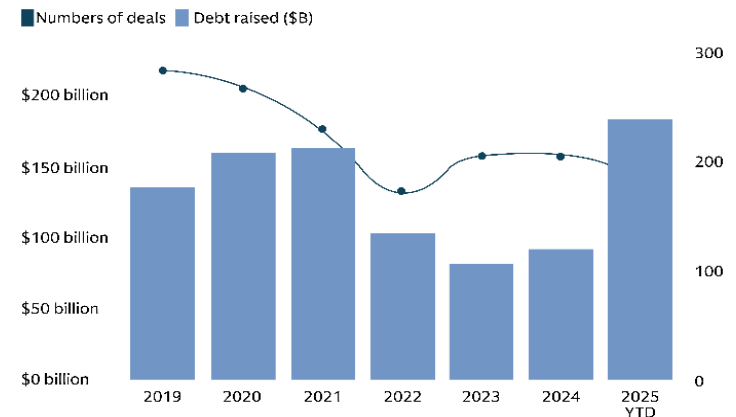
## Hyperscalers have burned through cash flow from operations



Source: Compustat, FactSet, Goldman Sachs Global Investment Research

## Data center debt issuance nearly doubles to \$182bn in 2025

Hyperscalers Meta and Google were among the most active issuers in 2025



Source: S&P Global Market Intelligence

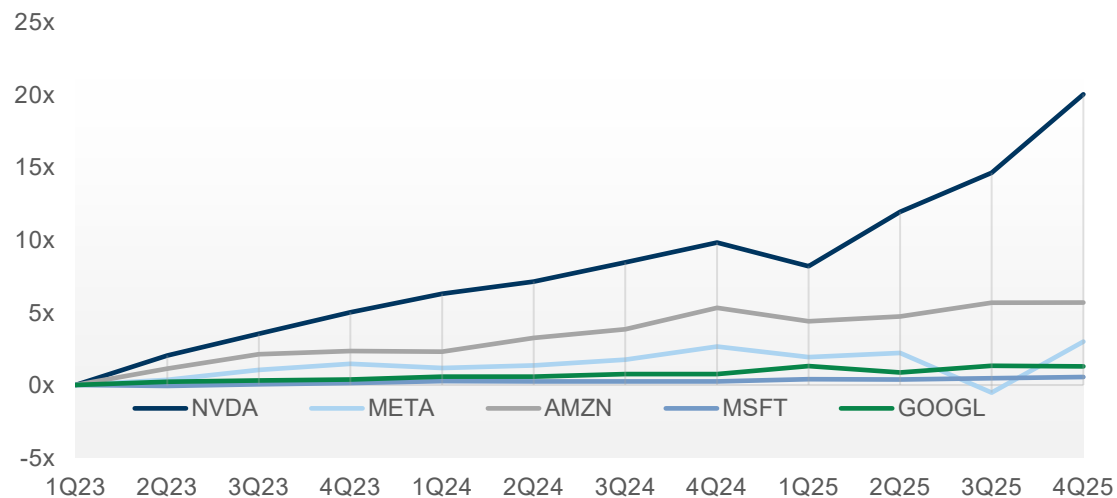
# The Concentration of Economic Value in Semis Is Not Sustainable

Historically, semi companies thrived when their customers were thriving – that’s not the case today

As discussed on the prior slide, most of the companies higher in the supply chain are losing money on AI while the majority of the economic value is accruing to the semiconductor companies. Enterprises are losing money implementing it, the model companies’ cash burn and need to raise capital has been well covered by the press and the hyperscalers have burned through all of their excess cash flow from operations.

Meanwhile, the semi supply chain is generating record revenues and profits. This dynamic is both unprecedented and unsustainable. We covered semi companies for nearly 20 years, and semi companies are meant to thrive when their customers are thriving. They are not supposed to be thriving at the expense of the companies higher in the chain. Something has to change with this dynamic – either the companies higher up in the chain need to start generating profits due to AI or they will eventually have to scale back on the spending.

**Net Income Growth since release of ChatGPT-3.5**

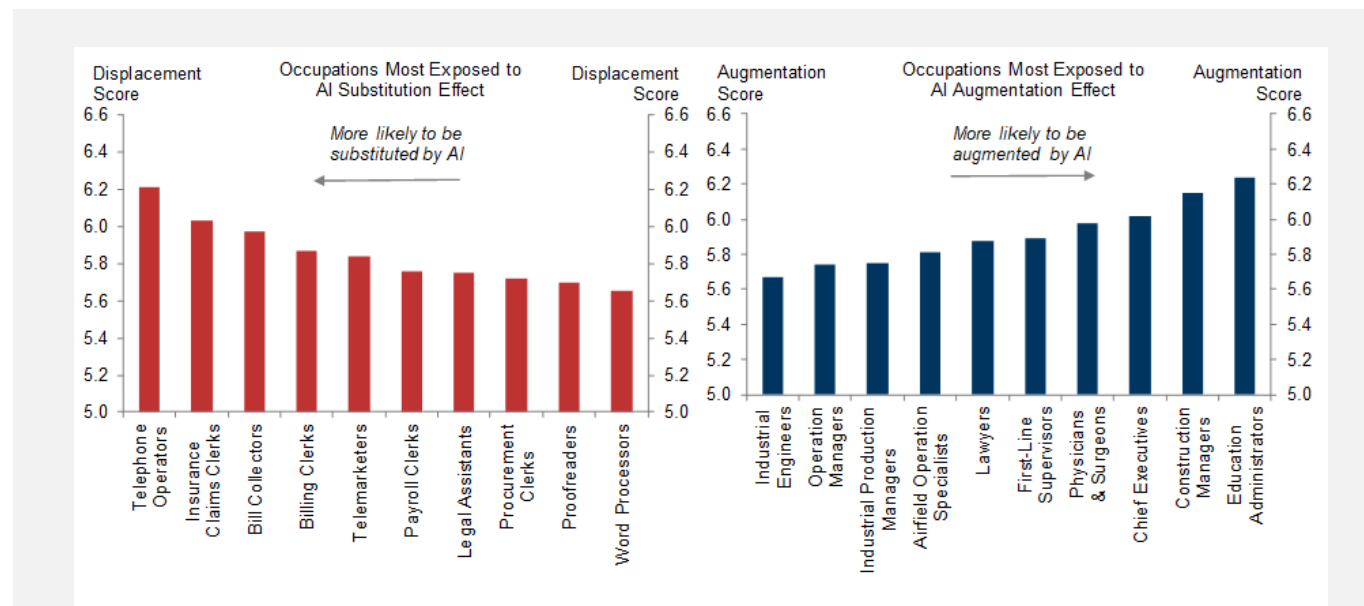


Source: Company Data, Goldman Sachs Global Investment Research

# There's a Significant Difference Between AI Replacing vs Augmenting Jobs

The market and media have overestimated job replacement and underestimated augmentation

We believe that one important distinction that is often lacking in the public discourse about AI and the jobs market is that there is a significant difference between AI replacing jobs (AI substitution) and jobs in which workers can increase efficiency by leveraging AI tools (AI augmentation). In a report led by Elsie Peng<sup>16</sup>, our Goldman Sachs Macro team did outstanding work putting analytical rigor around these distinctions in a manner which has been lacking in most of the media commentary about AI and the job market. The bottom-line in their findings is that there has been some AI-related impact on job losses (i.e., AI substitution) but there is just as much AI augmentation as there is AI substitution. The “AI will replace 50% of jobs” headlines will most likely persist as that drives clicks and views—but the analytics tell a different story.

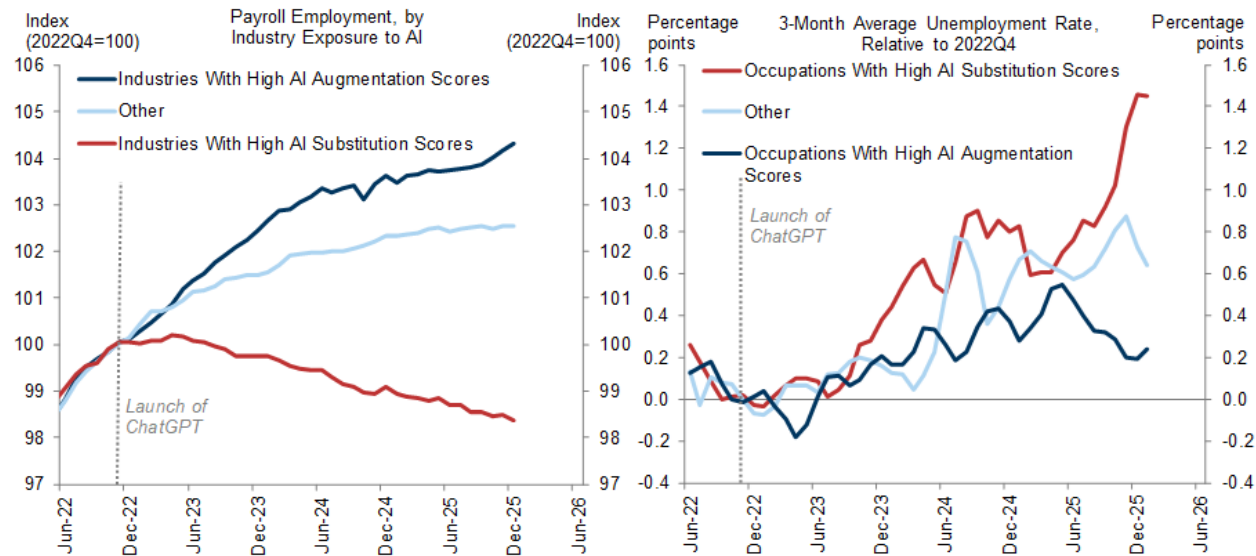


Source: Goldman Sachs Global Investment Research

- We assign a high *substitution* score to occupations where the AI exposure is high but complementarity is low, and we assign a high *augmentation* score to occupations where both AI exposure and complementarity are high.
- According to these scores, **telephone operators, insurance claims clerks, and bill collectors face the highest AI substitution risk** (left panel), while **education workers, judges, and construction managers face the highest AI augmentation potential** (right panel).

# No Significant Job Replacement Yet

Benefits in some industries have helped offset losses in others



Note: We exclude healthcare, education services, and government industries, which saw strong job growth due to post-pandemic catch-up effects.

Source: Goldman Sachs Global Investment Research

AI has lowered hiring in occupations where it substitutes for human labor but increased it slightly in those where it augments human labor.

- Industries and occupations facing high AI substitution risk have experienced a larger decline in employment and a larger increase in unemployment rates than less AI-exposed industries.
- Industries and occupations more exposed to AI augmentation have seen above-average employment growth and a below-average increase in unemployment, partially offsetting the substitution effects.
- We estimate that the net effect of AI substitution and augmentation has been a 16k drag on monthly payroll growth and a 0.1pp boost to the unemployment rate.

# AI's Long-Term Impact on Jobs Depends on Adoption, Displacement

Our baseline is for a modest peak increase in unemployment as workers transition

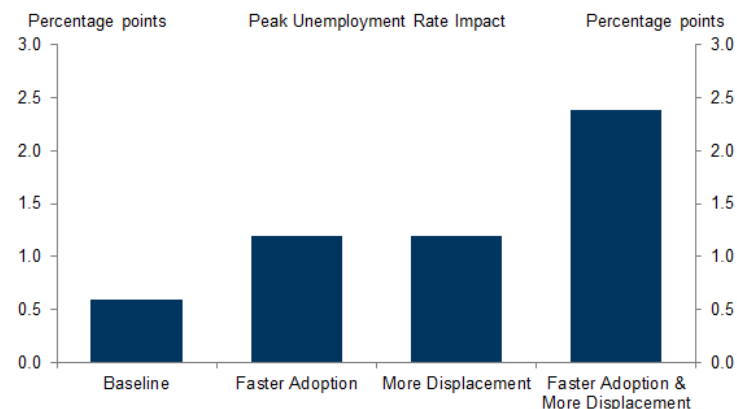
Jan Hatzius, Joseph Briggs, Sarah Dong, and our Global Macro Team have done tremendous work on AI's potential impact to the workforce over the next 10 years<sup>17</sup>. The baseline estimate is that AI can potentially replace 25% of all work hours in the economy, using Department of Labor's O\*NET database to estimate the share of work tasks that can potentially be automated following the full adoption of GenAI.

Combining our baseline forecast for a 15% AI-driven labor productivity uplift with the historical relationship between technological-driven productivity gains and job loss implies that **6-7% of jobs could be displaced** as AI is widely adopted.

While AI-driven displacement could create significant real income losses and economic headwinds for affected workers, history strongly suggests that AI will also create new jobs. Most immediately, the infrastructure buildup necessary to facilitate the AI transition is providing a near-term employment boost. Job growth for electricians, HVAC installers, commercial building contractors, and utility construction workers has outpaced both broader construction and economy-wide job gains since the ChatGPT moment in late 2022.

How much could the unemployment rate rise as displaced workers are transitioning to new roles? Under our baseline assumptions that 1) the adoption timeline is 10 years at the firm level, 2) 6-7% of workers are displaced during this transition period, and 3) most workers take less than a year to find a new job, **we estimate a peak gross unemployment rate increase of around 0.6pp**. That said, the actual increase in the unemployment rate would likely be somewhat lower if the Fed moves to lower rates to offset the weaker labor market.

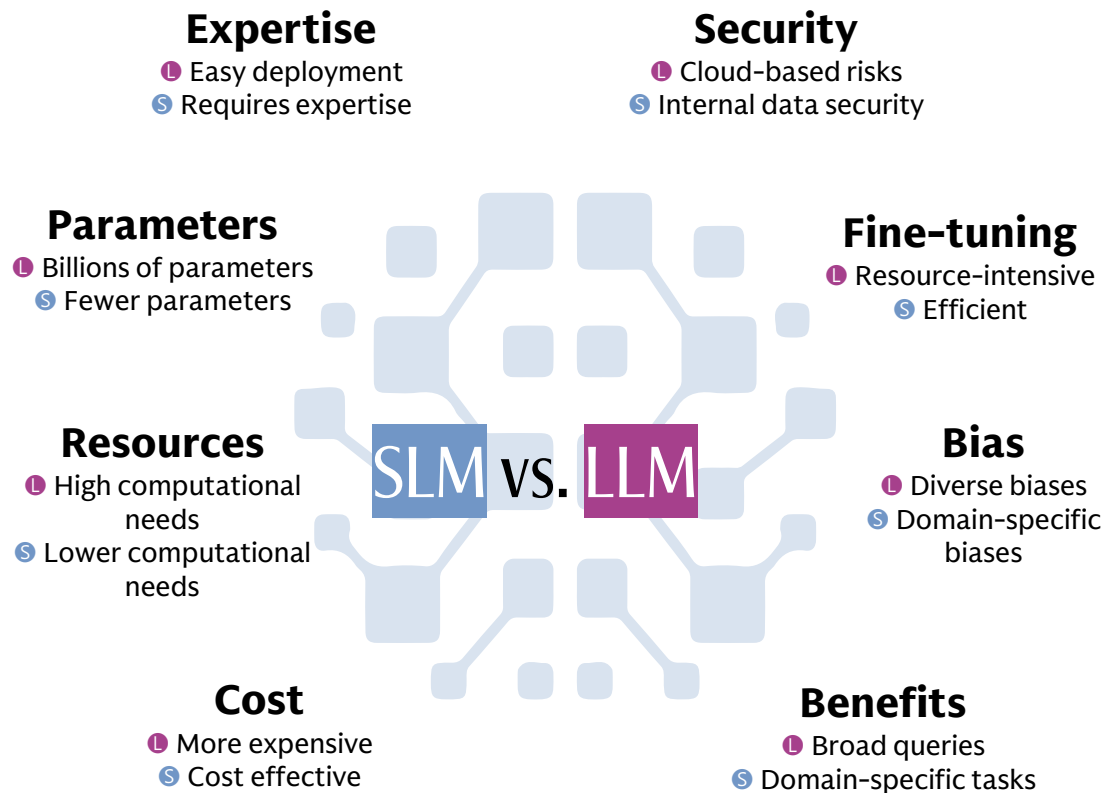
However, both the displacement rate and the transition period are uncertain, and we cannot rule out larger increases in the unemployment rate. If AI-driven labor displacement turns out to be twice as large as our baseline (corresponding to a scenario where AI fully automates work tasks in the 50% most-exposed occupations) or AI adoption occurs twice as quickly (i.e., over a 5-year period with 3-year intra-firm adoption lags), the peak unemployment rate impact would double to 1.2pp (2mn unemployed workers). And in a scenario where AI displacement is both larger and occurs more quickly, the peak unemployment rate impact would rise to 2.4pp (4.1mn unemployed workers). We see further upside risks from “congestion externalities”—where a sharp rise in unemployment leads to a less efficient matching process—or if AI-related job losses front-run AI-related productivity gains.



Source: Goldman Sachs Global Investment Research

# SLMs vs. LLMs

Task-specific small models offer significant advantages over broad LLMs



## Speed and real-time needs<sup>19</sup>

SLMs process data quickly, making them ideal for applications where every second counts, like voice commands in drones or instant customer support.

## Cost savings

Training and running LLMs can cost millions, requiring powerful GPUs. SLMs, with their lower resource needs, are much cheaper, making them perfect for startups, small businesses and task specific needs.

## Resource constraints

SLMs, because they have fewer parameters and computational needs, can run on edge devices like smartphones or IoT gadgets, while LLMs need cloud infrastructure, which isn't always practical.

## Privacy and control

Keeping data local with SLMs avoids cloud risks and the ease of fine-tuning give the models a big advantage in sensitive fields like healthcare or finance.

Source: Edureka<sup>18</sup>

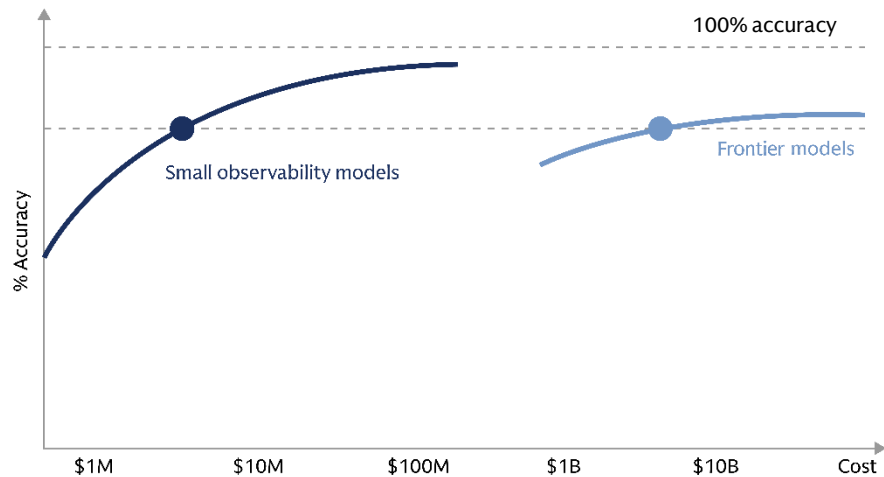
# SLMs Are a Better Value Proposition for Many Use Cases

## Examples where small models saved costs and improved results

At Datadog's Feb. 12 analyst day, Datadog's cofounders elegantly showed the importance of context by illustrating that its SLMs trained on its internal data are able to deliver better accuracy at lower cost than frontier models because of its domain experience. We believe similar illustrations from app software companies also demonstrate the value of SLMs.<sup>20</sup>

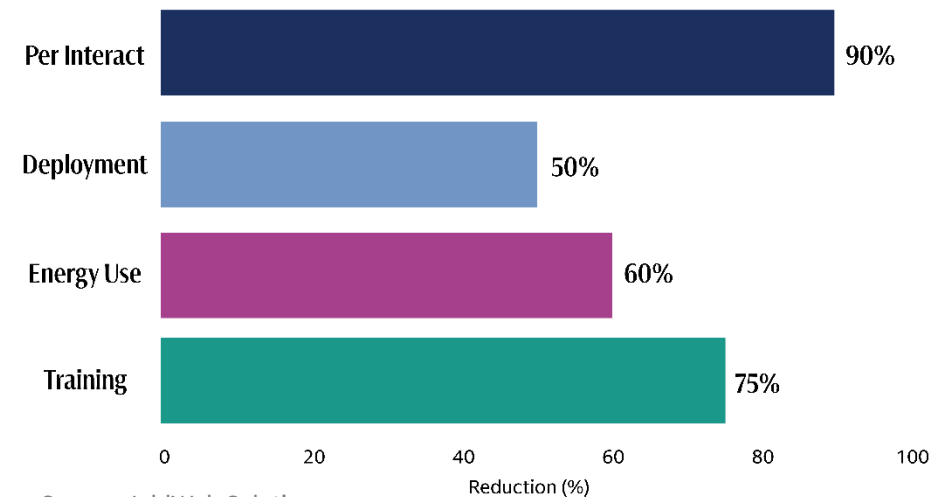
- A supply chain company reduced response latency by 47% and cut costs by 50% after switching from LLM to specialized SLM.<sup>21</sup>
- An enterprise SaaS provider replaced their generic LLM chatbot with an SLM fine-tuned on internal documentation. Internal chatbot usage increased 62% within three months, with users reporting higher satisfaction due to more accurate, contextually relevant responses.

### SLMs Vs LLMs



Source: Datadog

### Cost efficiency: SLMs Vs LLMs



Source: AddWeb Solutions



Q.

# Where Is the Value Ultimately Going to Accrue in the Supply Chain?

What are the most interesting companies we have met in the last two years working on?

# What Has to Happen to Unlock the Value of AI for the Enterprise?

Economic benefits need to spread across the supply chain

As we previously discussed in the section on AI economics, the majority of value in the supply chain is currently accruing to the semi companies.

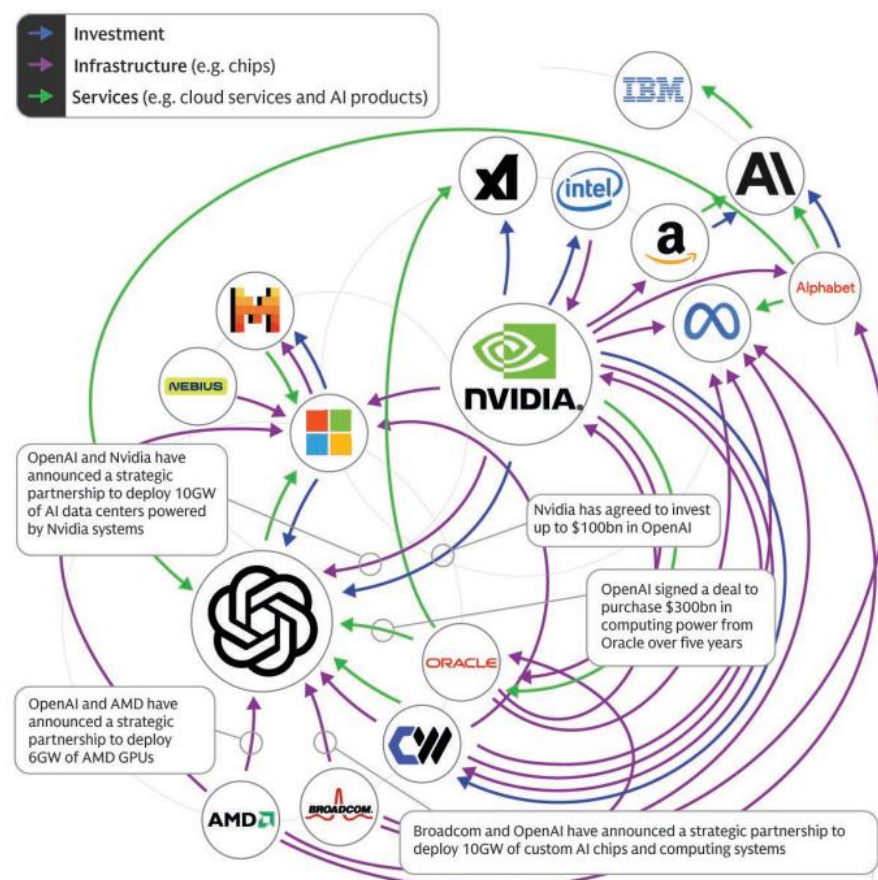
In order for AI to be successful, enterprises have to derive enough economic value from it such that they can pay the model companies enough of a margin so the model companies can pay the hyperscalers and the hyperscalers can continue to pay the semi companies. Now, the actual payments are much more complicated than the simple chain we describe above as there are economic arrows pointing in every direction in this supply chain. However, the point holds that the enterprises at the top of the chain have to begin to derive economic value from AI in order for the spending throughout the chain to be sustainable.

This leads us to the important question of...what has to happen to unlock the economic value of AI for the enterprise, and which companies are helping to unlock this value?

**We believe 1) data structure, and 2) orchestration and deployment are two critical factors that need to be solved to unlock the value of AI in the enterprise.** We ultimately believe many of the questions we receive about “tokenomics” will be addressed by companies providing solutions in these layers.

Building agents on top of data that isn't structured properly will lead to the kind of continued disappointing results in the enterprise that we referenced on p. 8. In addition, not orchestrating workloads properly will lead to the continued issue of costs that far exceed budgets because expensive models are too frequently being used to address commodity workloads.

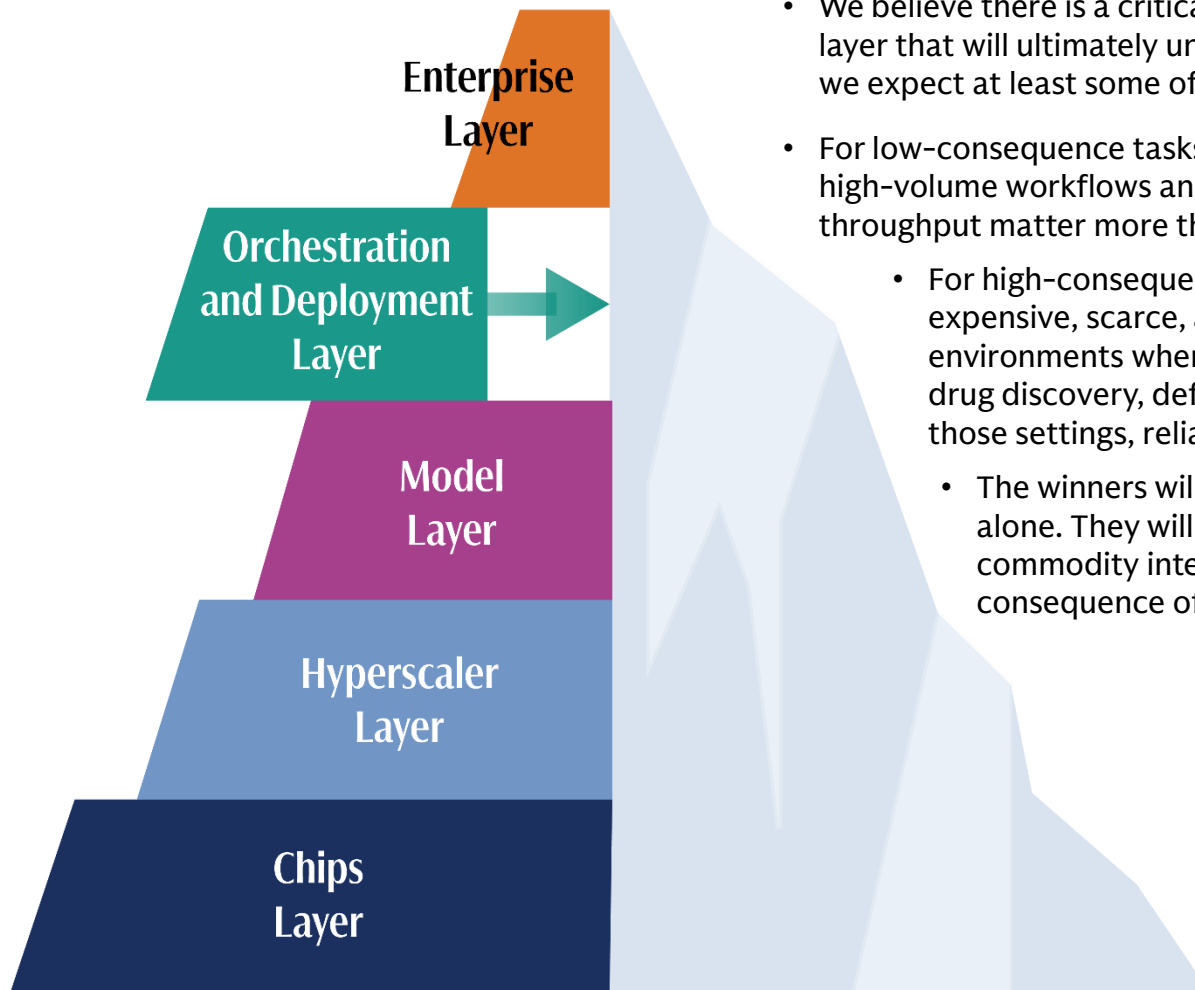
## Circular Financing in AI



Note: Exhibit does not represent an exhaustive view of the AI ecosystem but rather the most prominent AI model companies, infrastructure providers, and hyperscalers. Source: Company announcements, various news sources, compiled by Goldman Sachs GIR.

# Agentic AI Demands a New Orchestration and Data Layer

Ensuring data is structured properly and workflows are directed appropriately



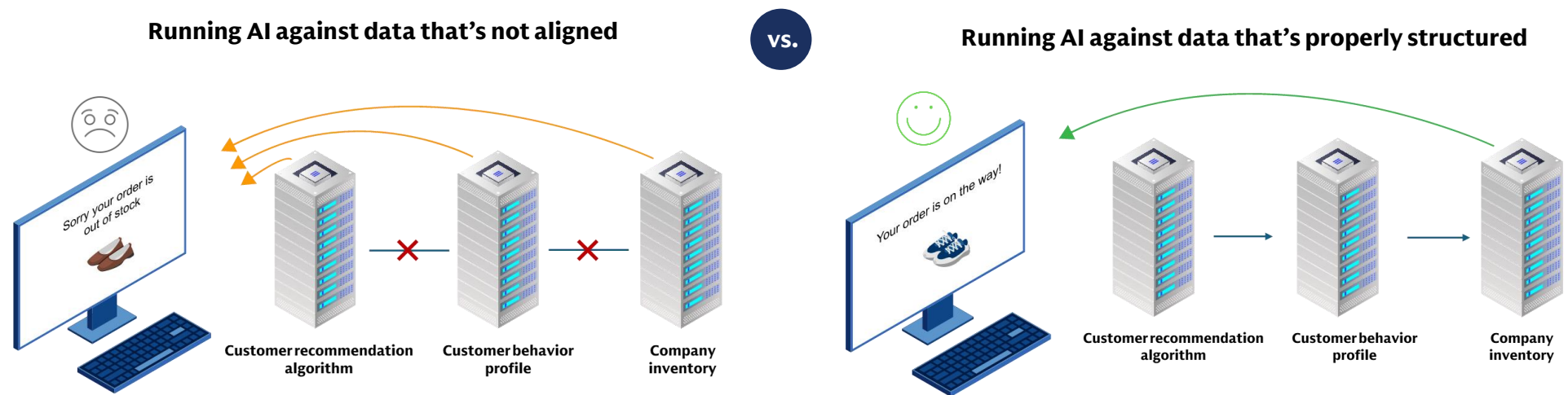
- We believe there is a critical layer between the model layer and the enterprise layer that will ultimately unlock the value of AI in the enterprise and this is where we expect at least some of the value to accrue.
  - For low-consequence tasks, open-source and lightweight models will handle high-volume workflows and everyday use cases where cost, speed, and throughput matter more than precision.
    - For high-consequence tasks, the most advanced models will remain expensive, scarce, and strategically important. They will be reserved for environments where the cost of failure is too high to absorb, whether in drug discovery, defense, or high-stakes financial decision-making. In those settings, reliability is key.
    - The winners will not be defined by access to one class of model alone. They will be the companies that can orchestrate both: commodity intelligence at scale, and premium intelligence where the consequence of error is too high to tolerate failure.
- Agentic AI will only be successful once data is structured properly and workflow is directed properly. This is the function we think the emerging Orchestration and Deployment layer will serve.

# Data Structuring

## Non-technical practical example: AI Agent on a retail website

In lieu of technical schematics which can be tremendously confusing for non-engineers (us included), below we provide a non-technical example of how an organization needs to structure data to enable models to most effectively meet customer needs. In addition, we provide a non-technical example of how workflows need to be orchestrated to manage token usage and improve AI economics. The simplicity of these explanations will likely offend any technical readers but we are presenting the information this way because we have gotten so many questions from investors about how this layer practically works.

We hear all the time that, “good AI requires better data structure” but what does that actually mean? Let’s take a retail company that has different databases. One database might have customer recommendation algorithms, another database has the company’s current inventory and yet another stores the customer behavior profiles. Those databases may sit in different silos within the organization. Unless all of that data is organized and aligned, a company’s AI can’t give the customer a suggestion that will always be implementable. For example, what if the databases currently can’t talk to one another and a customer uses a company’s AI agent on the website to search for a recommendation but the product the AI recommends is out of stock, that’s not a good customer experience. This is a concrete example of how a company needs to ensure that the databases are organized and aligned in a way that the AI agent can take full advantage of the data that the company has. In many cases today, agents are being built on top of data that isn’t aligned yet and therefore, the customers aren’t getting the best user experience. There are many new startups, including companies like AI One, that are offering solutions to help enterprises reengineer their data in a way that is efficient and effective.



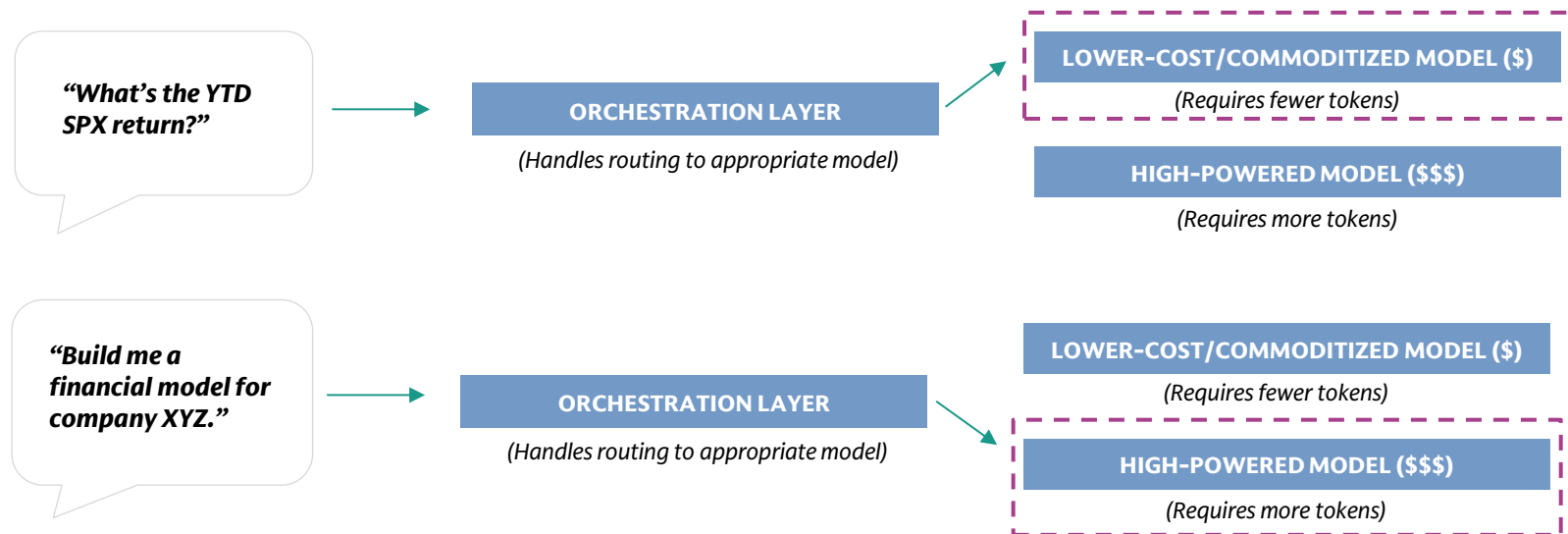
# The Orchestration and Deployment Layer

## Non-technical practical examples: Hedge fund PM queries

Another concrete (albeit non-technical) example of the importance of the orchestration layer can be seen in a portfolio manager at a hedge fund. Let's assume a fund which provides its entire employee base with access to one of the most powerful LLMs. There are certain queries that demand the use of a high-powered model but other queries that are essentially glorified web searches and those workloads should be routed to a lower-end model. If an analyst wants to build a new model from scratch or do some complex valuation analysis across industries or different periods of the market, then that workload needs to be routed to the high powered LLM. However, if the analyst or PM simply wants to know the multiple of a given stock or how a certain index has performed in the last week, that's a workload that should be sent to a cheaper model that doesn't use as many tokens per query. Again, while a very simple explanation, we hope this demonstrates the concept of how orchestration can impact AI economics in an organization. In many organizations today, this orchestration layer doesn't exist and so the low consequence queries are getting addressed by the highest end models which is one of the reasons we see so many stories of companies' AI budgets being so elevated.

We believe organizations getting these building blocks in place will be key to unlocking AI economics in the enterprise.

### Illustrative example: Orchestration in action for a hedge fund PM





Q.

# What Profit Pools Could Be Disrupted?

- AI potential to disrupt advertising
- Will AI eat software?
- Is AI a feature or a bug in cybersecurity?
- AI in transportation case study

---

# Looking Beyond Labor for ROI Opportunities

Advertising, software, cybersecurity and transportation have profit pools that could be disrupted

---

Whenever we analyze new technologies or new technology companies, the first question we always ask ourselves is, “**what profit pool will the new technology disrupt?**”

Amazon disrupted brick and mortar retail, Google disrupted print advertising, Tesla disrupted the internal combustion engine, Nvidia disrupted Intel’s data center CPUs, Uber disrupted taxis, Cloud disrupted on-prem, etc., etc. In each of these cases, one company’s profits became the disrupting companies’ revenue. That is the dynamic we have been on the lookout for since AI first came onto the scene.

What profit pool could AI disrupt to allow the supply chain to see ROI? The market has focused on jobs replacement and we demonstrated earlier that this isn’t happening yet. That doesn’t mean it won’t happen, but we think the market needs to look beyond jobs to feel comfortable that there will eventually be ROI.

Our equity analysts at GS have done an outstanding job exploring which profit pools could be ripe for disruption in the years ahead. To be clear, we aren’t seeing the disruption to justify the AI spending yet, but if we are going to see enough profit pool disruption to justify the spending, it’s possible that it comes from one of the areas that we examine in the case studies below.

## GS Research’s AI Disruption Case Studies

Advertising



Led by Eric Sheridan

Software



Led by Gabriela Borges

Cybersecurity



Led by Gabriela Borges

Transportation



Led by Mark Delaney

# AI's Potential to Disrupt the Advertising Profit Pool<sup>22</sup>

Opportunities in content creation, improved effectiveness and expanding the TAM

We see several multi-billion-dollar advertising profit pools that AI has the potential to disrupt in the coming years. These include:

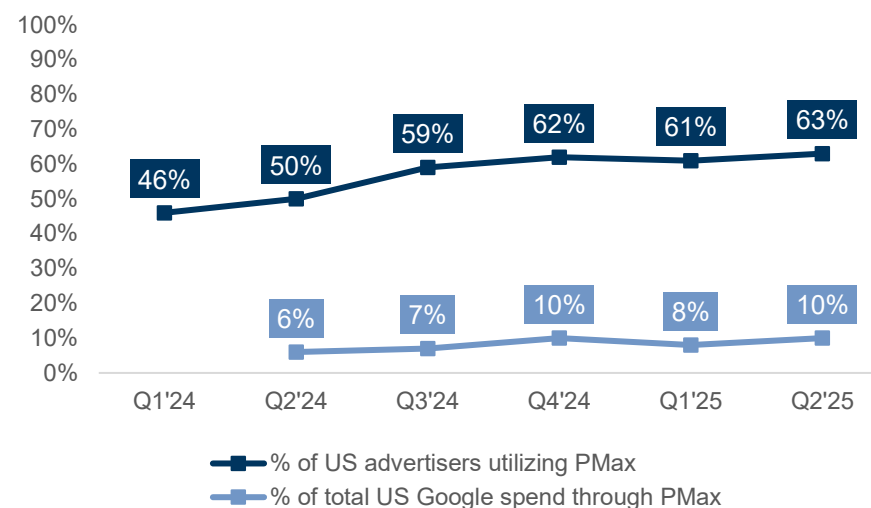
- Accelerating the shift of ad spend from traditional to digital channels (~\$170 bn)
- Automating the ad creative generation process (~\$114 bn)
- Consolidation of ad tech “Intermediaries” (~\$25 bn)
- Potential impact to the ad agency ecosystem (~\$161 bn)

We see AI being implemented across various areas/use cases including: Content Creation, AI Assistants & Agents, Content Recommendations, Optimization Engines, & Campaign Execution and Modeled Conversions.

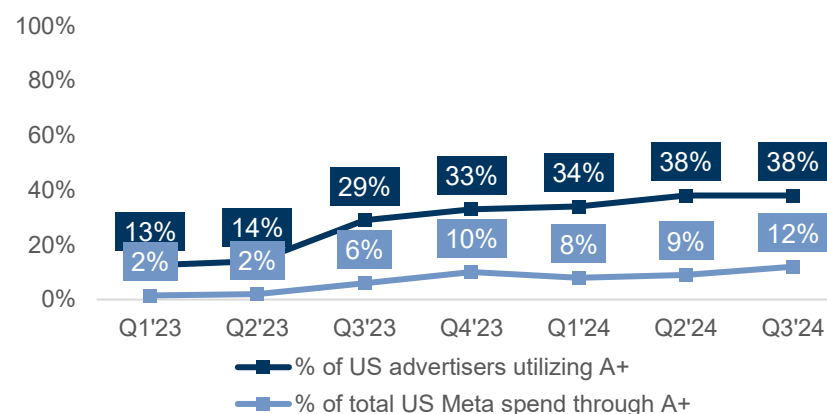
Google's (GOOGL) Performance Max and Meta's (META) Advantage+ are thus far the most successful examples of most of these AI use cases being combined into a single product. Both products offer various levels of automated campaign execution (across different ad formats & media), personalization and granular user targeting and, increasingly, generative AI tools for content creation.

We also see a number of new **“TAM expansionary” growth opportunities** that AI may facilitate in the future. These include: (1) improving return on ad spend (“ROAS”) of existing channels; (2) “democratization” of advertising expanding the addressable advertiser base; & (3) introduction of new ad formats.

## Adoption of Google's Performance Max (Pmax)



## Adoption of Meta's Advantage+ (A+)



Source: Goldman Sachs Global Investment Research

# Will AI Eat Software?<sup>23</sup>

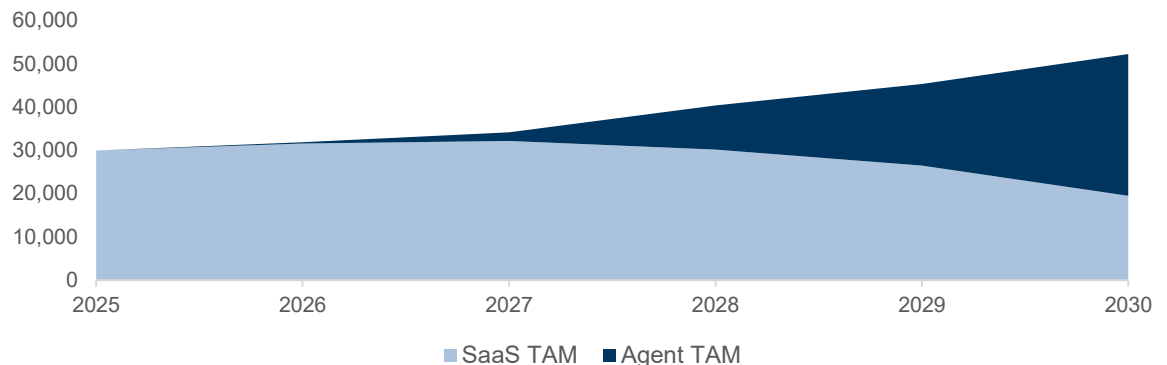
We see TAM growing, not shrinking, as domain expertise drives better AI outcomes

**We are more convicted on the Software TAM getting bigger.** Companies are increasingly positioning their AI workflows as selling a unit of labor or a unit of productivity, which allows them to tap into larger deal sizes and new budget.

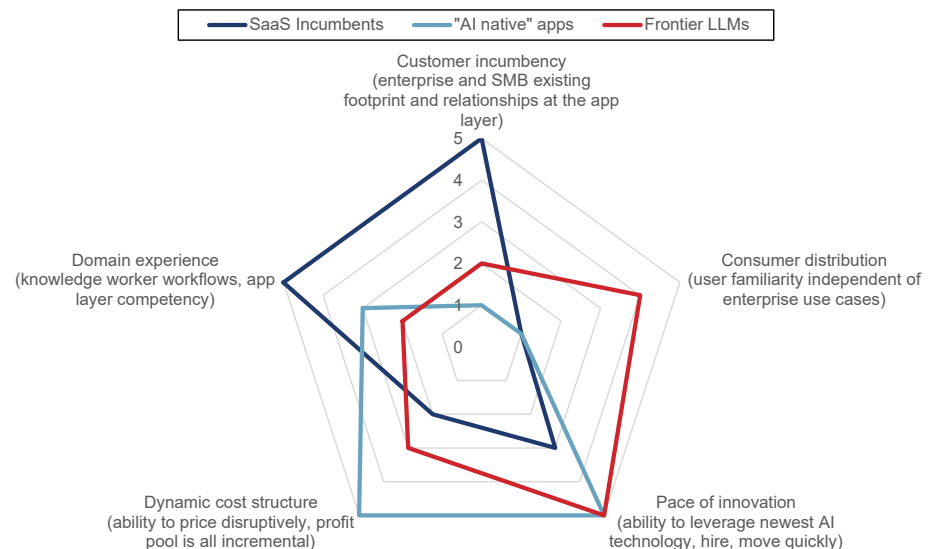
The risk from AI to today's incumbents in our view is less on rip and replace, but more on value abstraction. **The biggest advantage that incumbents have is their domain experience and context.** Domain experience lends itself to SLMs that are higher accuracy and lower cost than general purpose models. This dynamic is one of the reasons why enterprises will likely engage with an ecosystem of software providers that are independent of any one model provider: the enterprise benefits from more efficient routing and lower inference costs.

**The software leaderboard in 2030 will likely include select incumbents and several next-gen AI startups.** We expect select incumbents will be able to prove to customers that this domain experience drives better AI outcomes, provided they have been thoughtful about cleaning up technical debt and innovating as fast followers. However, not all companies are well positioned to execute through intensifying competition, which lends itself to stock-specific opportunities.

**Our TAM math on the Customer Service subsector shows that the TAM is ultimately bigger with a shift from SaaS to agents because of realized productivity gains; today's Software leaders need to capture part of the upside to avoid being overly exposed to the downside**



**We see three cohorts of competition at the app software layer with relative strengths and weaknesses**



# AI in Cybersecurity: Net Positive or Net Negative? <sup>24</sup>

We expect a divide between "base-level" and "protection" security

We expect over time the LLM providers will focus on two runways:

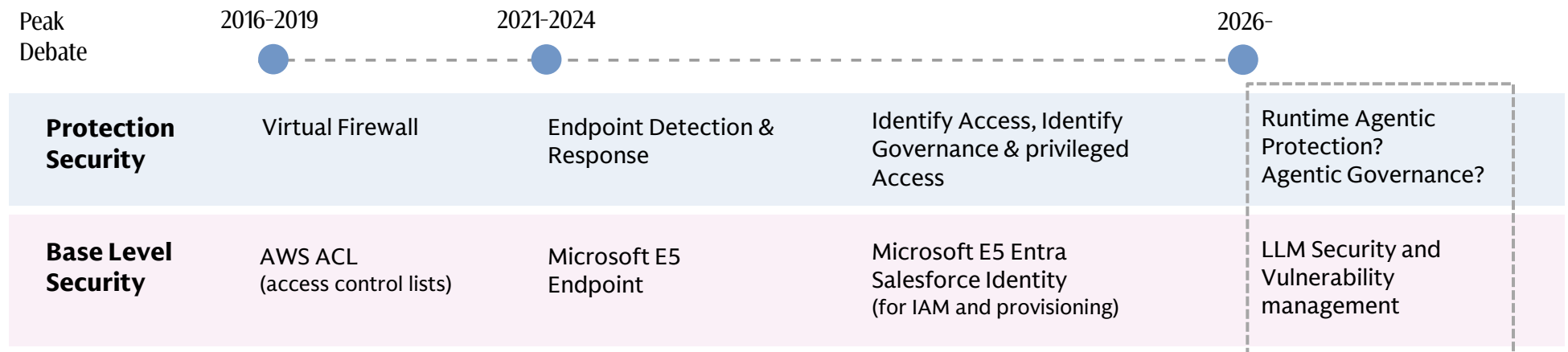
1. Providing a base level of security within their product that limits any friction to adoption.
2. Security use cases that could be solved with very strong anomaly detection engines in spite of limited security-specific domain experience.

We think one way to frame AI-native security capabilities is through the lens of "base-level security" vs. "protection security." Platform providers often embed baseline security controls directly into their core products to remove friction to adoption (identity, access control, basic threat detection), while customers still layer independent vendors on top for deeper protection and visibility.

The LLM providers are incentivized to provide a base level of protection in their core products to remove friction points for adoption. This base level of protection may limit growth in select subsegments of Security for less differentiated vendors. However, these security solutions will likely not be comprehensive enough to displace large parts of the Security TAM.

We expect to see three product cycles: 1) a near-term push to improve the quality of existing infrastructure in partnership with the leading LLMs for vulnerability management (e.g. Project Glasswing with Anthropic and Trusted Access with OpenAI); 2) an upgrade in tooling in the security operations center (where analysts monitor, investigate and resolve breaches), including for run time agentic protection; and 3) identity and governance tools for agents.<sup>25</sup>

**We believe the construct of "base level" and "protection" security can be helpful: the former removes friction to adopting the product; the latter offers best-in-breed security**

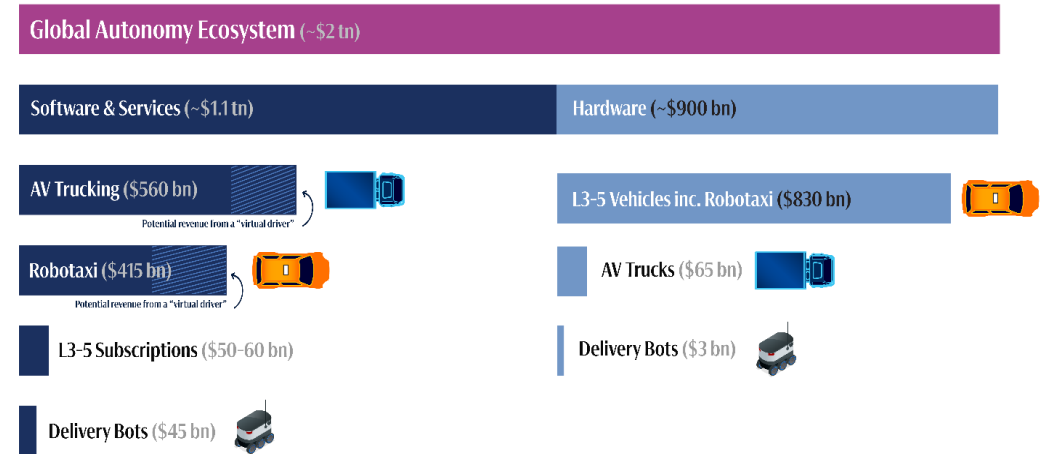


# AI in Transportation<sup>26</sup>

Robotaxis and autonomous heavy duty trucking are sizable end markets

Our Autos team, led by Mark Delaney, has done a lot of work examining AI's potential impact in the transportation industry. Considering the market more broadly for AV hardware and software & services, **we estimate that revenue associated with the industry could reach approximately \$2 trillion in 2035** comprised of hardware sales and software/digital services. Recognizing that a meaningful portion of this \$2 trillion of revenue associated with autonomy is coming from business for goods/services that isn't incremental, **when focusing more specifically on the piece of this market attributable to AI, we estimate that the market will be approximately \$300 bn in 2035.** While a portion of AV volumes will likely be incremental demand, we believe autonomy could also disrupt existing markets in the long-term. We frame risks and scenarios including to human-operated rideshare/taxis, trucking, and light vehicle unit sales in the United States. **We estimate the disruption of AI to existing markets could potentially be ~\$440 bn in the US.**

## Potential global autonomy ecosystem market size 2035



Source: Goldman Sachs Global Investment Research

- We now estimate that the AV robotaxi market in the US will reach \$19 bn in 2030 and \$48 bn in 2035. When including global markets, we estimate that the robotaxi market in 2035 could be ~\$415 bn. We assume that the gross margins of a vertically integrated operator could be 30%-50%, implying gross profit for the market of ~\$150 bn in 2035 and cumulative gross profit of ~\$440 bn over the next decade for robotaxis.
- Class 8 trucking (heavy duty commercial vehicles with a gross vehicle weight rating >33,000 pounds) is also growing, and we think the market could reach \$105 bn in the US and ~\$560 bn globally in 2035. We estimate that gross profit for AV trucking will be ~\$135 bn in 2035, and about \$300 bn cumulatively over the next decade.



Q.

# What Advice Would We Give to the C-Suite on AI from Here?

- Slow down: Is FOMO driving enterprises to adopt tech that is not really ready?
- Build vs. Buy
- SLMs vs SaaS: What's the difference in the long-run?

---

# Focus on AI Building Blocks and a Rational End Game

Ensure that FOMO isn't rushing an AI solution that isn't the right long-term answer

---

- **Slow down now so you can speed up later!** We understand that there is immense pressure from investors to articulate an AI strategy to the market. FOMO is dictating that companies need to talk about how far along they are with their AI implementations or risk their shareholders moving on to another company that is articulating a better AI story.
  - The problem with this dynamic is that there are still some very important issues that need to be resolved to make AI both effective and affordable in the enterprise. We believe C-suites should play the long game and get all the building blocks necessary for long-term AI success (which we detailed in the previous slides) in place before rushing to build agents on top of data that isn't ready for an agent or before there is an orchestration layer that is ready to steer workloads to the appropriate models.
  - Until these building blocks are ready, there's a good chance a lot of spending that is being done is going to have to be redone.
  - Remember the lessons of the internet: the pioneers got the arrows and the settlers got the land. Uber built a life-changing app on top of the capacity that some other companies went bankrupt building, but it didn't happen until 20 years later. Companies should strive to be the Uber in this analogy – not the companies that rushed to build the capacity because they wanted to be the first to articulate a good story to the Street.
- **Make sure the end game makes sense.** While LLMs have proven challenging to successfully implement, SLMs have been more successful. That's the good news. The question that C-suites should be asking in light of this is, "what's the difference between the SaaS applications I am trying to replace and an enterprise that is running hundreds of discrete SLMs?" If 3 years from now, companies are running hundreds of SLMs that are doing a great job addressing task-specific functions but a) those SLMs are expensive to build and maintain, and b) none of those SLMs talk to one another, how different is that from the SaaS applications that they were meant to replace?
- **Build vs. buy dynamics may be shifting toward buy.** One of the original aspects about AI that appealed to organizations was that AI could be built internally, which would allow less reliance on expensive-to-maintain software licenses. Given how expensive and difficult the AI build is proving to be, there has been increasing momentum toward "buying" as customers recognize that success requires far more than a point-in-time agent. As our software team has noted, operating production-grade conversational AI agents demands ongoing investment in testing, safety, model upgrades, and orchestration, which requires engineering teams to work on the project indefinitely. Customers are now increasingly questioning whether bespoke agents are truly core IP and whether allocating top engineering talent is justified, as opposed to partnering with platforms that have already solved this complexity at scale. C-suites should definitely be pushing on the build vs. buy strategy within their corporations.



Q.

# Stock Implementation Idea

We recommend being long the hyperscalers vs. underweight semis

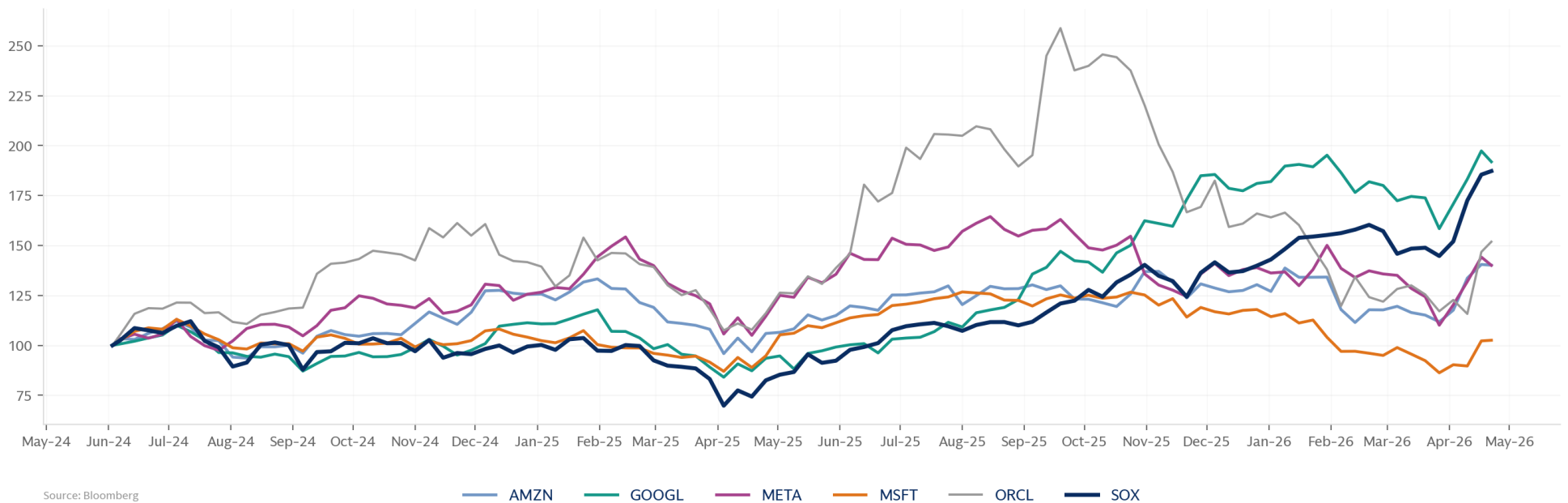
# Relative Value Trade Between Hyperscalers and Semis

The "picks and shovels" phase may be fully priced in

In our original piece in June '24, we suggested that investors focus on the "picks and shovel" providers in the early days of the AI buildout. That has played out as the semi and semi cap stocks have significantly outperformed the market while the hyperscalers (other than Alphabet) have underperformed the market. From here, we think investors should be relatively overweight a basket of the hyperscalers and underweight a basket of semi and semi cap stocks against the long positions. We see two scenarios where investors will get paid on this relative value trade and one scenario in which this will work against investors.

The starting point for this idea is that the market has priced in a fair degree of skepticism on the ROI for the hyperscalers, as reflected in significantly compressed multiples for that group.

**SOX has significantly outperformed the hyperscalers**



---

# Assessing Scenarios in Which This Trade Wins or Loses

Path to profit if enterprises start showing ROI or if hyperscalers pare back spending

---

Here are the 3 scenarios we envision and how we would expect the trade to play out:

1. Enterprises begin to demonstrate a positive ROI in which case investors will be less skeptical of the spend by the hyperscalers and those stocks should recover a significant portion of the multiple compression while the semi stocks rise less as the market has been generally more positive on that group.
2. Enterprise ROI remains challenging and the hyperscalers decide to pull back on their AI capex spend. We believe this is the best case scenario for this trading idea as we believe the hyperscalers would see a significant relief rally on better cash flow prospects and the semi stocks would get hit significantly as the lower hyperscaler capex would negatively impact their revenues.
3. The status quo persists (and investors lose money on this trade). That is, the hyperscalers continue to increase their capex despite continued enterprise ROI challenges. In this scenario, all of the value would continue to accrue to the semi companies while the hyperscalers see even worse cash flow dynamics.

---

# Further Reading

---

[Top of Mind: Gen AI: too much spend, too little benefit?](#)

25 Jun 2024

[Top of Mind: Will AI eat software?](#)

9 Mar 2026

[Analyzing the impact of AI on profit pools - Part II - A Transportation Case Study](#)

16 Apr 2026

[Global Economics Analyst: How Concerned Should We Be About a Job Apocalypse?](#)

09 Jan 2026

# Bibliography

- (1) Singh, Shubham. (Mar 26). ChatGPT Statistics (2026) – Active Users & Growth Data. DemandSage
- (2) (Dec 2023). Are You Addicted to Your Phone? American Phone Usage & Screen Time Statistics. Harmony
- (3) Rosa Royle, Orianna. (Mar 2026). Workers are using AI to sneak out for spin classes and skip lunch meetings – and new research shows they're clawing back 30 minutes a day. Fortune.
- (4) Sha Sajadieh, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Lapo Santarlasci, Juan Pava, Nestor Maslej, Russ Altman, Erik Brynjolfsson, Carla Brodley, Jack Clark, Virginia Dignum, Vipin Kumar, James Landay, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Elham Tabassi, Russell Wald, Toby Walsh, Dan Weld. "The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026.
- (5) Ellis, Lindsay, Stamm, Stephanie. (Jan 2026). CEOs Say AI Is Making Work More Efficient. Employees Tell a Different Story. The Wall Street Journal.
- (6) Vidgen, Bertie, Mann, Austin, Fennelly, Abby, Wright Stanly, John, Rothman, Lucas, Burstein, Marco, Benchek, Julien, Ostrofsky, David, Ravichandran, Anirudh, Sur, Debnil, Venugopal, Neel, Hsia, Alannah, Robinson, Isaac, Huang, Calix, Varones, Olivia, Khan, Daniyal, Haines, Michael, Bridges, Austin, Boyle, Jesse, Trist, Koby, Richards, Zach, Mahapatra, Chirag, Foody, Brendan, Nitski, Osvald. (Feb 2026) APEX-Agents. Mercor Research
- (7) (Jan 2026). The AI Proficiency Report. Section AI
- (8) Unlocking enterprise AI: opportunities and strategies. The Economist Impact
- (9) Niederhoffer, K., Kellerman, G. R., Lee, A., Liebscher, A., Rapuano, K. and Hancock, J. T. (Sept 2025). AI-Generated "Workshop" Is Destroying Productivity. Harvard Business Review.
- (10) Challapally, Aditya, Pease, Chris, Raskar, Ramesh, Chari, Pradyumna. (Jul 2025). The GenAI Divide, State of AI in Business 2025. MIT Nanda
- (11) Mills, Madison. (Apr 2026). 1 big thing: AI can cost more than human workers now. Axios
- (12) Depa, Joe. (Oct 2025). How can responsible AI bridge the gap between investment and impact? EY
- (13) (Jul 2025) Gartner Forecasts Worldwide IT Spending to Grow 7.9% in 2025. Gartner
- (14) (Feb 2025) Gartner Forecasts Worldwide IT Spending to Grow 10.8% in 2026, Totaling \$6.15 Trillion. Gartner
- (15) Roach, April. (Dec 2025). Data center deals hit record \$61 billion in 2025 amid construction frenzy. CNBC
- (16) Peng, Elsie. (April 2026) US Daily: Breaking Down the Impact of AI on the Labor Market: Substitution vs. Augmentation. Goldman Sachs Global Investment Research
- (17) Briggs, Joseph, Dong, Sarah (Jan 2026) How Concerned Should We Be About a Job Apocalypse? Goldman Sachs Global Investment Research
- (18) Mishra, Anupam (Dec 2025) LLM vs SLM: What's the Difference in 2026. edureka
- (19) (Oct 2025). LLMs vs. SLMs: The Rise of Smaller Models Outperforming Larger Ones. ARTIBA
- (20) Borges, Gabriela, Hotchkiss, Adam, Martino, Matthew, Naparst, Noah, Hager, Maura, Sklba, Greyson, Valenti, Callie, Zhang, Selina, Gamperl, Max, Arora, Praachi. (Feb 2026). Revisiting Moats Part I: Exploring AI Steel Men Arguments. Goldman Sachs Global Investment Research
- (21) Dhariwal, Saurabh (Dec 2025) How Small Language Models (SLMs) Are Democratizing AI for Mid-Market Companies. AddWeb Solutions
- (22) Covello, James, Sheridan, Eric, Vegliante, Alex, Sachdeva, Aarshiya. (Jun 2025) Analyzing the Impact of AI on Industry Profit Pools – Part I (Advertising Case Study). Goldman Sachs Global Investment Strategy Group
- (23) Borges, Gabriela, Hotchkiss, Adam, Martino, Matthew, Naparst, Noah, Hager, Maura, Sklba, Greyson, Valenti, Callie, Zhang, Selina, Gamperl, Max, Arora, Praachi. (Feb 2026). Revisiting Moats Part I: Exploring AI Steel Men Arguments. Goldman Sachs Global Investment Strategy Group
- (24) Borges, Gabriela, Gamperl, Max, Arora, Praachi. (Mar 2026). Revisiting Moats III: Cybersecurity. Goldman Sachs Global Investment Research
- (25) Borges, Gabriela, Gamperl, Max. (Apr 2026). Updated thoughts on LLM innovation and Security. Goldman Sach Global Investment Research
- (26) Delaney, Mark, Sheridan, Eric, Chang, Allen, Costa, Daniela, Frenes, Christian, Keung, Ronald, Yuzawa, Kota, Kong, Lincoln, Jeng, Verena, Hou, Tina, Zheng, Nick, Hyoung Kim, Do, Muthiah, Chandramouli, Bryant, Will, Gupta, Aman, Fein-Ashley, Julia, Menting Liu, Monika, Yang, Meihan, Zhang, Xuan, Ghose, Ayush, Huang, Emma. (April 2026). Analyzing the impact of AI on profit pools- Part II- A Transportation Case Study. Goldman Sachs Global Investment Strategy Group