Key AI debates: our take

Global Equity Strategy

Equity Strategy
Global

- We remain bullish on the AI trade, but the next leg depends on more evidence that monetization can keep pace with capex
- We provide our takes on the seven key global AI debates we have with investors
- Report includes 50+ of the most important AI-related charts

The AI trade has entered a more challenging phase as questions over monetization and returns intensify. In this report, we provide our take on the seven key debates we have with investors.

1) AI monetization: how big, how fast?: Monetization has significant runway. Reported annualised revenue run-rates are already large (Anthropic USD65bn, OpenAI USD40bn). Corporate adoption looks early: median spend is USD12/employee/month vs USD650 for the top decile (both rising quickly). Capability gains (e.g., longer task completion) should expand willingness to pay.

2) Token costs + competition: Token prices are deflationary, but offset by rising volumes (Jevons paradox dynamics). Chinese models are nearing ~50% of token volume, but US frontier players still capture ~80% of token spend. Also, "cost per token" matters less than "cost per task" as complexity rises.

3) Hyperscaler capex: overspend risk and ROIC sustainability: Capex is enormous (USD770bn in 2026e, USD1.1trn in 2027e, per consensus) but demand signals remain strong: hyperscaler backlog reached ~USD2.35trn in Q2-26 (vs USD815bn a year earlier), and GPU rental prices have stayed firm – suggesting scarce usable compute.

4) Capex vs monetization timing: Timing mismatch is real – FCF has turned negative and financing is rising: ~USD250bn debt issuance so far in 2026, plus >USD1trn of uncommenced lease commitments.

5) Korea memory: when to re-enter after the correction? Still bullish. Stocks are ~35% off peak; market expectation of earnings cycle duration has been compressed; positioning has been cleansed; fundamentals remain supported by capex-driven memory demand; shareholder returns are rising.

6) Mainland China AI: where does value accrue? We prefer domestically focused AI semis/hardware stack (inference growth + localisation + policy support). China remains supply-constrained but less efficient local chips can actually broaden demand for packaging/networking/power.

7) How to position for AI within a global portfolio: We tilt to areas with strongest demand/monetization and less-stretched valuations – especially EM (such as Korean memory, Taiwan semis) and mainland China (semi equipment, power infrastructure). In the US, our preference is semis > hyperscalers > software.

Alastair Pinder, CFA
Head EM and Global Equity Strategist
HSBC Securities (USA) Inc.
alastair.pinder@us.hsbc.com
+1 212 525 5972

Pankaj Agarwala*, CFA
EM and Global Equity Strategist
HSBC Securities and Capital Markets (India) Private Limited
pankajagarwala@hsbc.co.in
+91 0 80 30013716

** Employed by a non-US affiliate of HSBC Securities (USA) Inc, and is not registered/ qualified pursuant to FINRA regulations*

Disclosures & Disclaimer
This report must be read with the disclosures and the analyst certifications in the Disclosure appendix, and with the Disclaimer, which forms part of it.

Issuer of report: HSBC Securities (USA) Inc.

View HSBC Global Investment Research at:
https://www.research.hsbc.com

Debate #1: AI monetization: how big, how fast?

The debate: *Key questions remain around the pace of AI monetization and the ultimate size of the revenue opportunity, and whether corporates will truly pay for frontier models.*

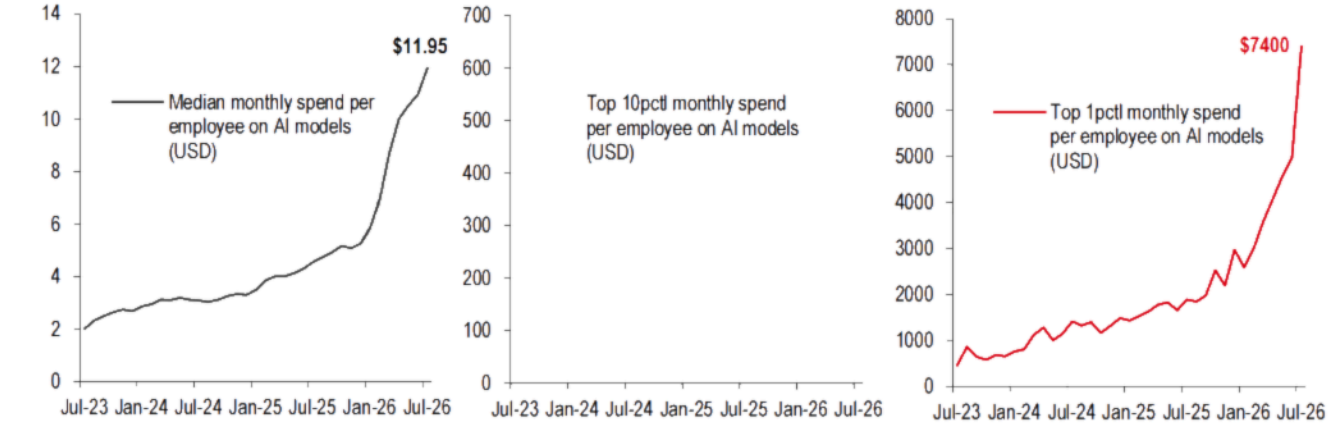
Our take: We continue to see substantial room for AI monetization. Demand for frontier AI models remains extraordinarily strong. Recent reports suggest Anthropic reached an annualized revenue run rate of USD65bn as of July, while OpenAI hit USD40bn (Bloomberg, 13 August 2026 and Reuters, 17 August 2026, Chart 2), highlighting that commercial adoption remains far from saturated. Ramp AI data shows the median company spends just USD12 per month per employee on AI, while the top 10% spends USD650 per month (Chart 1) . Both figures are accelerating sharply. Spending is also broadening across industries, with AI spend in Financials and Manufacturing sectors year-to-date up 3.5x and 2.5x, respectively. The hyperscalers are seeing the same broadening in enterprise adoption. Microsoft reported more than 30m paid M365 Copilot seats in Q2-26, with net paid seat additions more than doubling sequentially. Alphabet said nearly 90% of the Fortune 100 now use Gemini Enterprise, while more than 2,000 enterprises each consumed over 100bn tokens during the past 12 months.

Corporate budgets also appear large enough to support materially higher AI spending. If the average S&P 500 company spent USD650 per employee per month, annual spending would reach USD250bn, equivalent to just 5% S&P 500 of EBITDA, 8% of wages and 27% of R&D (Chart 3). According to Cledara, AI currently accounts for just 20% of software budgets, but that is materially shifting with 50% of new IT/software budgets being allocated to AI (Chart 4). Current budget constraints therefore appear manageable if companies continue shifting existing IT expenditure towards AI (Chart 5).

Improving model capability should also expand the amount companies are willing to spend (Charts 6 and 7). Alongside improving intelligence scores and accuracy rates, the ability to complete longer and more complex tasks should allow AI to move from assisting with individual prompts to executing end-to-end workflows, expanding both the addressable market and the value against which companies can justify spending. METR data show the Claude Mythos model has an 80% success rate at tasks lasting over three hours (Chart 11). Claude Opus, released in May 2025, could achieve a similar success rate only on tasks lasting 20 minutes.

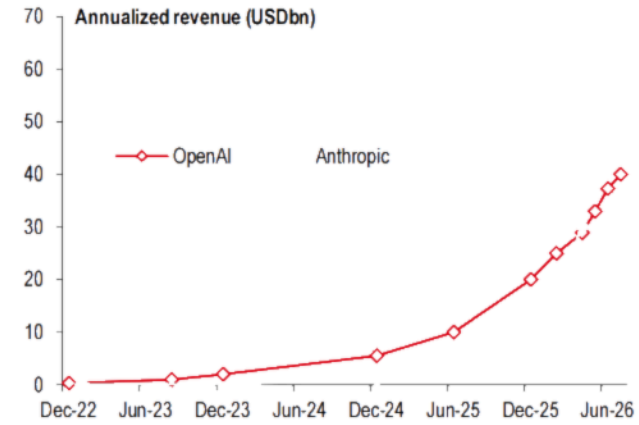
The Anthropic S-1, which is now reported to be made pubic in late September (Reuter, 4 September 2026), will provide one of the first major tests of the AI monetization thesis. Customer concentration, retention, gross margins after inference costs, compute commitments and cash conversion will likely matter more than headline revenue growth. Strong unit economics would demonstrate that rising AI usage can support continued infrastructure investment. Weak cash conversion would revive concerns that demand is being financed ahead of sustainable returns. Reports have suggested that Anthropic has already made positive adjusted operating income as of Q2-26 (Bloomberg, 14 August 2026).

1. Corporate spending on AI models per employee per month is growing exponentially



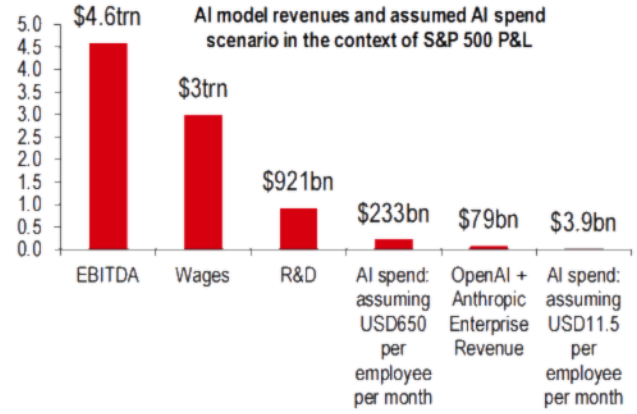
Source: Ramp AI, HSBC

2. OpenAI and Anthropic AI revenue trends



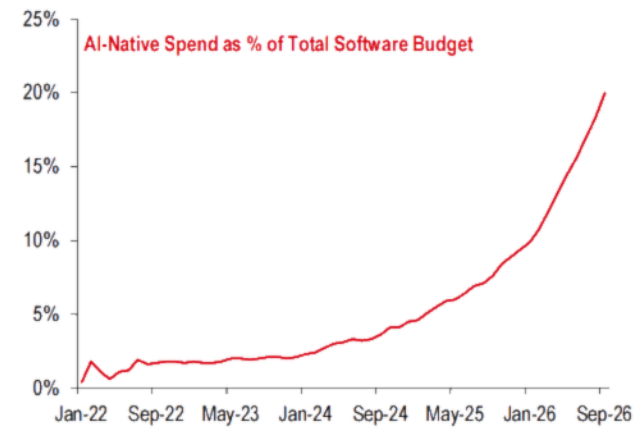
Source: Epoch AI, company reports, HSBC

3. AI spend is still an very small share of total budgets



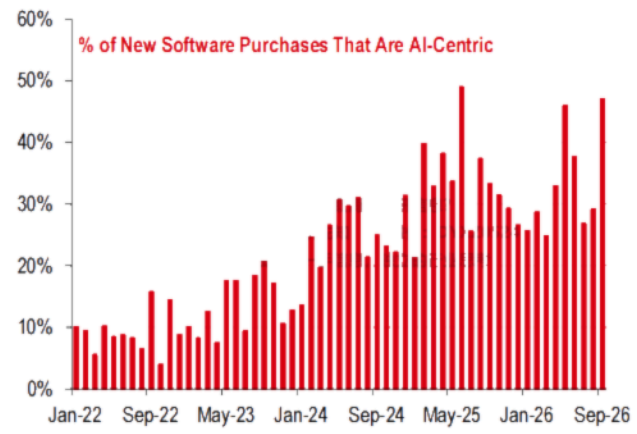
Source: Ramp AI, Factset, HSBC

4. AI spend as % of total software budgets…



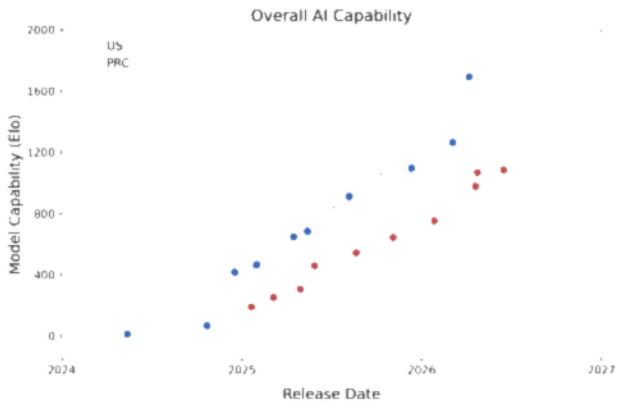
Source: Cledara, HSBC

5. … and as % of new software purchases



Source: Cledara, HSBC

6. US vs mainland China frontier model capability



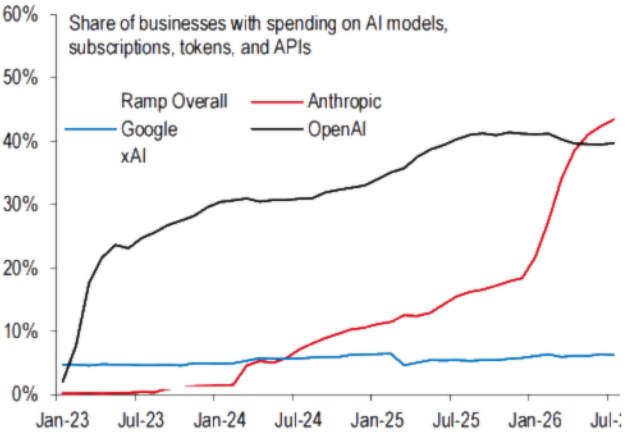
Source: CAISI, HSBC

7. Pace of improvement in AI models has increased further over the last couple of years



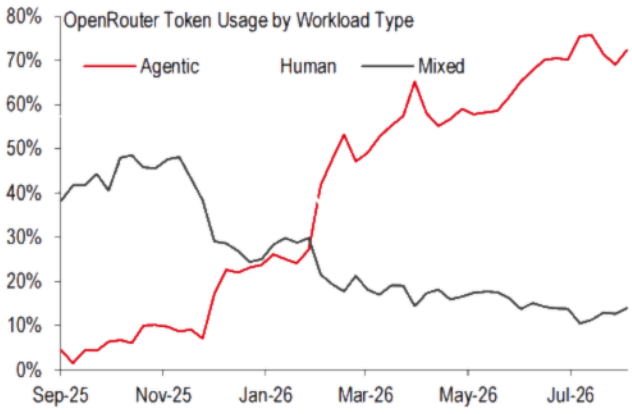
Source: Epoch AI, HSBC

8. US corporate AI adoption rate, split by model



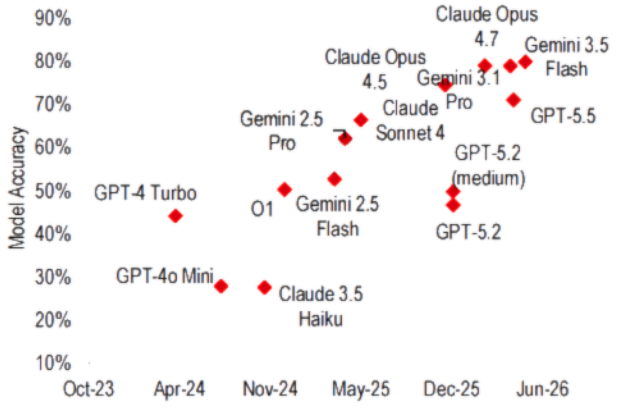
Source: Ramp AI, HSBC

9. Agentic models taking a greater share of AI token volume



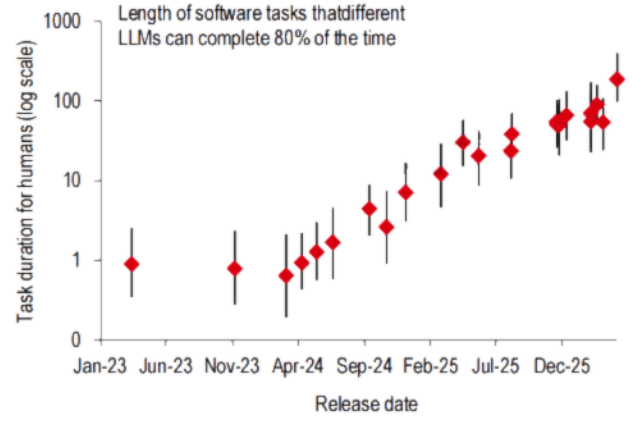
Source: OpenRouter, HSBC

10. Latest models are becoming more and more accurate…



Source: HAL, HSBC

11. … and capable of performing longer and more complex tasks



Source: METR, HSBC

Debate #2: Token costs and competition from Chinese / open-weight models

The debate: *Falling token prices and the growing use of Chinese and open-weight models have increased concerns over AI pricing power. According to the Silicon Data Token Expenditure Index, the average price paid for tokens has fallen 50% since its May 2026 peak, largely as users have migrated towards cheaper models. OpenRouter data also show Chinese models approaching 50% of total token volume.*

Our take: Both trends are deflationary for price per token, but neither is necessarily deflationary for aggregate AI spending.

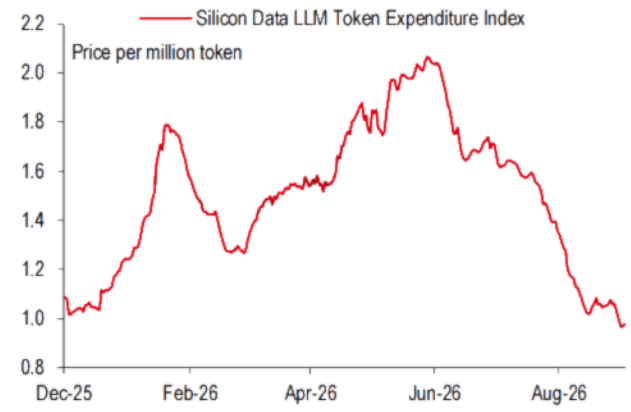
First, Jevons paradox appears to be in full swing: token prices are down, but token volumes have surged almost 3x since June (Chart 13). OpenRouter's recent discount experiment provides a useful test of this elasticity. Daily token usage for OpenAI's Terra and Luna models rose 5.6x and 13.8x, respectively, during the discount window, compared with just 1.1x for the undiscounted Sol model, while around 32% of users continued to use the discounted models after the promotion ended (OpenRouter, 25 August 2026). Lower inference costs can therefore create additional demand that persists after pricing normalizes. Second, token-volume share should not be confused with token-spend share. While Chinese models have gained a larger share of token volume, they have not captured a significantly larger share of token spending. Anthropic and OpenAI, for example, have consistently taken around 80% of token spend since 2025, according to data from Vercel (Chart 15).

We would also challenge how people view the "costs" between US and Chinese models. Price per token has becomes less useful as workloads grow more complex. OpenAI's latest Astra model costs USD7.7 per 1m tokens, almost 6.5x the price of Kimi3, while its cost per task is only 2x higher because of its greater capability (Chart 16). For agentic applications, repeated attempts, additional token consumption and human supervision all increase the cost of completing a task. Higher-priced frontier models can therefore remain competitive where better reliability lowers the total cost of execution. Meanwhile, the aggressive pricing of Chinese models may prove less durable than many assume. DeepSeek has already announced 'significant' price increases (Bloomberg, 5 August 2026). Moonshot AI is reportedly negotiating with Microsoft Azure, Amazon Web Services and Google Cloud for as much as a 30% revenue share when those platforms host its Kimi K3 model in order to monetize adoption beyond current headline token prices (Reuters, 26 August 2026).

Crucially, the other point to make is that broader open-weight adoption does not necessarily disintermediate the hyperscalers. They can monetize these workloads through managed API / cloud access, accelerator rental, storage and database usage. Hyperscalers can therefore capture infrastructure revenue even where customers choose a third-party or open-weight model.

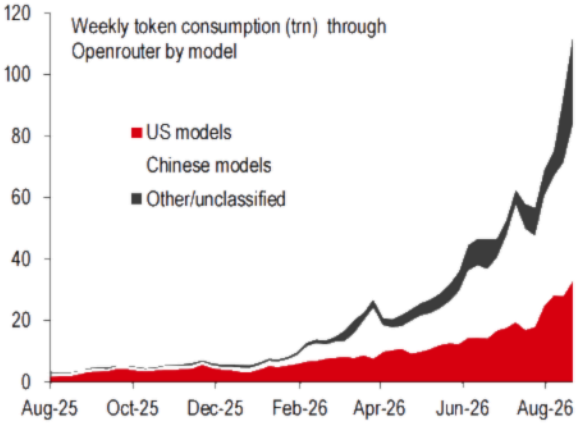
We foresee a world where US frontier models and Chinese or open-weight alternatives all coexist, accommodating the needs of different customers and use cases. Frontier US models should remain best positioned in complex reasoning and agentic applications where capability and reliability have greater value. Chinese and open-weight models should gain more share in high-volume, price-sensitive and specialized workloads. Enterprises can use several models and route tasks according to performance and cost, increasing competitive pressure at the model layer while supporting continued growth in aggregate inference demand.

12. Average price paid per token is falling



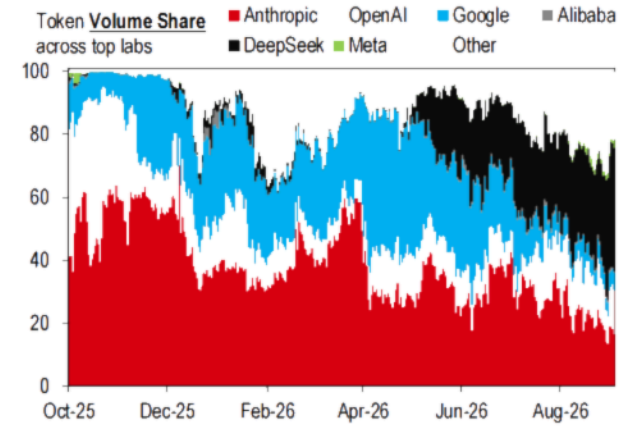
Source: Bloomberg, HSBC

13. Chinese AI models are accounting for a greater share of token volume



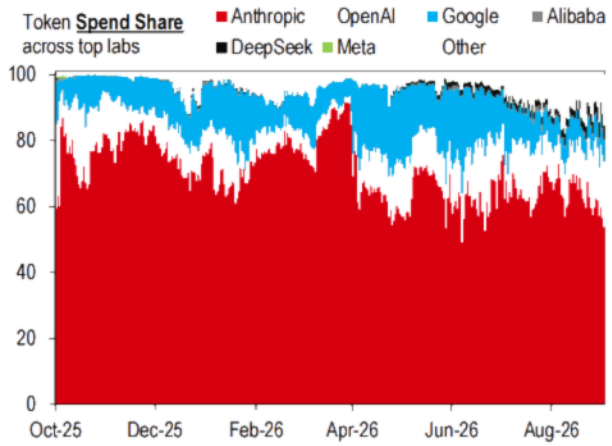
Source: OpenRouter, HSBC

14. Although Anthropic and OpenAI are losing market share in token volumes…



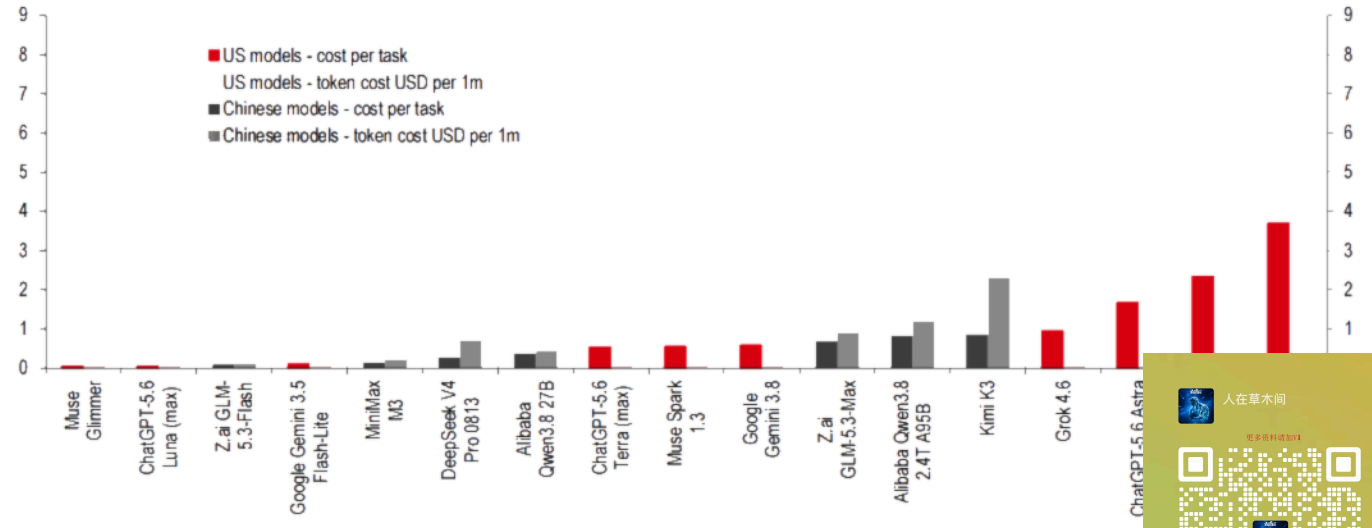
Source: Vercel, HSBC

15. … their share in of token spend has remained more consistent



Source: Vercel, HSBC

16. AI models cost comparison



Source: Artificial Intelligence, HSBC

Debate #3: Hyperscaler capex: overspend risk and ROIC sustainability

The debate: *The scale and continued acceleration of hyperscaler capex has intensified concerns around overinvestment, particularly whether demand can absorb the capacity being built and generate sufficient returns on the capital deployed.*

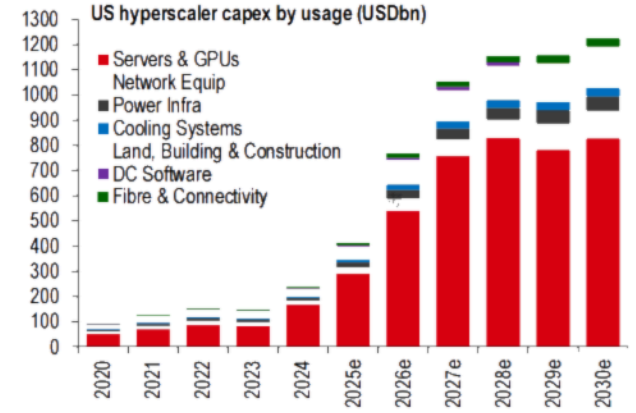
Our take: Consensus expects US hyperscaler capex to reach USD770bn in 2026 and USD1.1tm in 2027 (Chart 17). These estimates have been revised up aggressively since the start of the Q2 earnings season, by USD170bn and USD260bn, respectively (Chart 18). Even these numbers understate the full spending commitment: outstanding purchase commitments and lease commitments not yet commenced total around USD2.5trn. S&P Global Ratings estimates that the US hyperscalers will spend more than USD7tm on data centers and AI-related capex between 2025 and 2030 (S&P Global Ratings, 4 September 2026). The risk is that infrastructure supply eventually grows faster than the revenue available to support it.

There is little evidence of that imbalance yet. Remaining performance obligations (RPOs) across Microsoft, Amazon, Alphabet and Oracle reached approximately USD2.35trn in Q2-26, up from USD815bn one year earlier (Chart 26), suggesting contracted demand has so far risen alongside capex. The composition of that backlog is also improving. We estimate around 35% is linked to frontier-model companies, but Microsoft's commercial RPO still increased 25% excluding OpenAI and customers outside frontier-model companies accounted for all sequential RPO growth during the quarter. Alphabet's Cloud backlog reached USD514bn, with the majority coming from standard Google Cloud Platform (GCP) contracts across a broader customer base and just over half expected to convert into revenue within 24 months. The expansion of demand beyond the major AI laboratories should reduce one of the biggest concerns around the sustainability of the current investment cycle.

Pricing in the compute market gives a similar signal. Rental prices for NVIDIA B200 and H100 GPUs have continued to rise despite the rapid increase in capacity (Chart 23), suggesting usable compute remains scarce. This matters because announced capex does not translate immediately into available supply: power constraints, construction and commissioning can delay the point at which new infrastructure actually enters the market. HSBC's technology team also expects the AI ecosystem to generate approximately USD920bn of annual revenue by 2030, with enterprises contributing 75% of the total[1], while HSBC's analysts also suggest ROIC will fall, but remain sufficiently high at 17% by 2028[2] (Charts 27 and 28). We therefore see limited evidence of aggregate overcapacity today. The risk becomes more material if the current pace of capex continues after compute pricing softens, backlog growth slows, or utilization begins to fall.

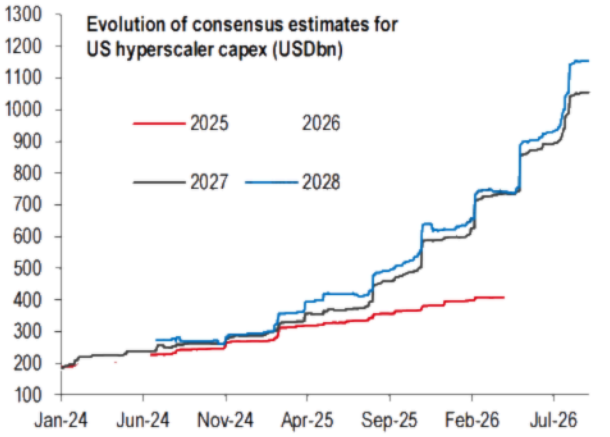
1 *Global Tech Platform: B2B up, B2C down: our new AI TAM forecasts to 2030*, Nicolas Cote-Colisson, 9 June 2026
2 *Global Technology: Sustainable CSP capex to support AI trend through FY28e*, Frank Lee, 12 August 2026

17. Consensus estimates US hyperscaler capex spend by usage



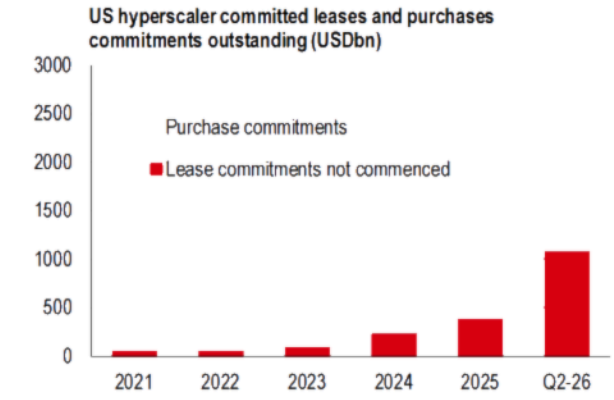
Source: Factset, HSBC

18. Evolution of consensus hyperscaler capex estimates



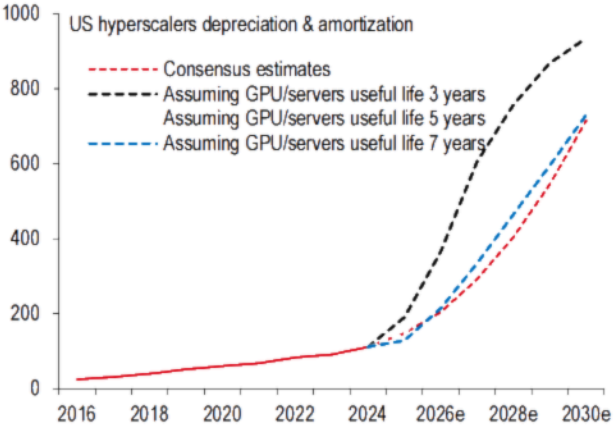
Source: Factset, HSBC

19. US hyperscaler committed leases and purchases outstanding



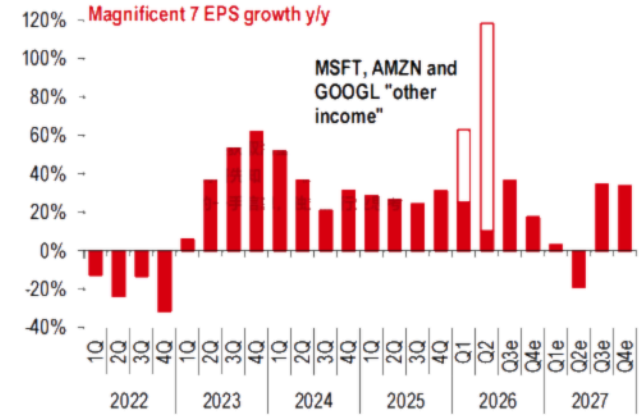
Source: Company reports, HSBC

20. US hyperscaler depreciation & amortization set to surge (USDbn)



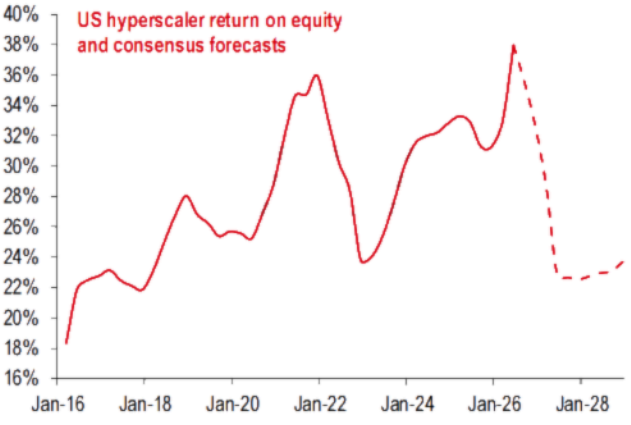
Source: Factset, HSBC

21. Magnificent 7 consensus EPS growth



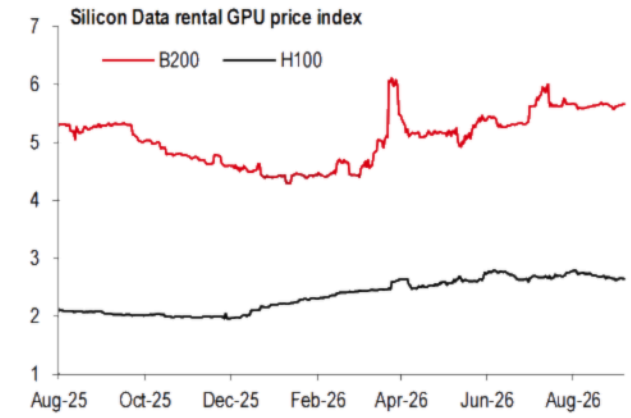
Source: Factset, HSBC

22. US hyperscaler return on equity set to decline



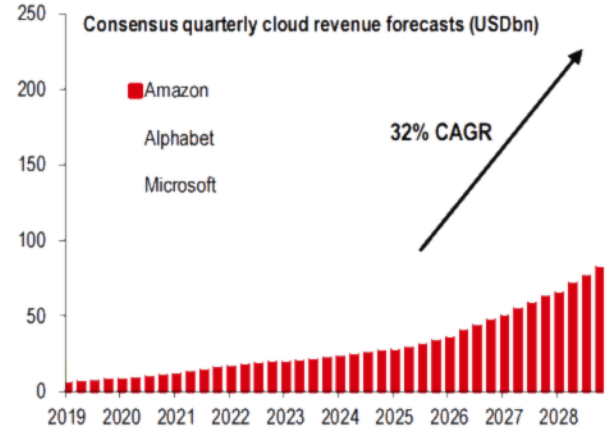
Source: Factset, HSBC

23. Rental GPU prices have remained firm



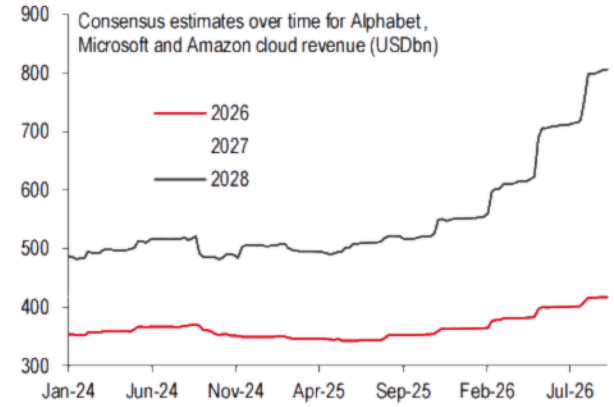
Source: Factset, HSBC

24. US cloud revenues are expected to surge...



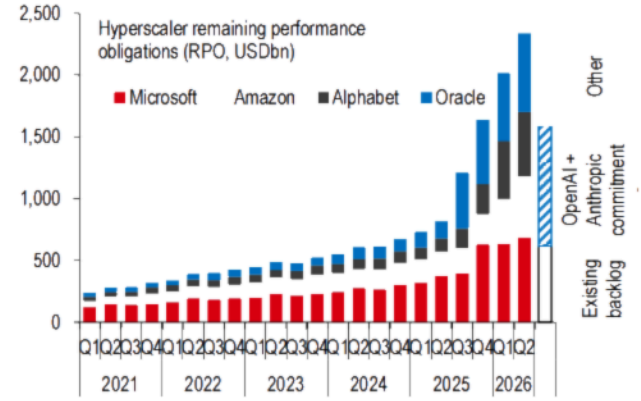
Source: Factset, HSBC

25. … and consensus estimates have been revised up



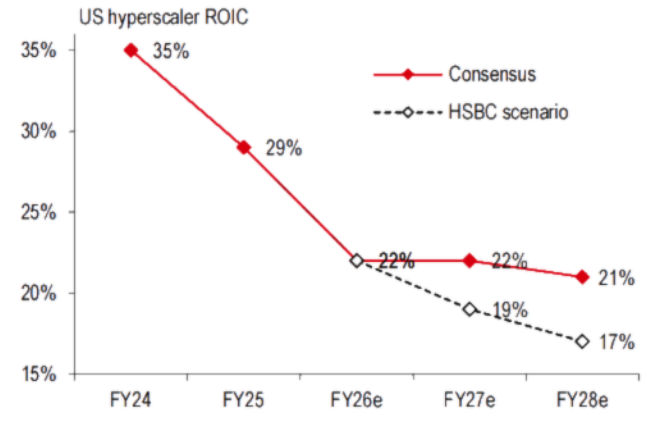
Source: Factset, HSBC

26. Hyperscaler backlog is extremely strong



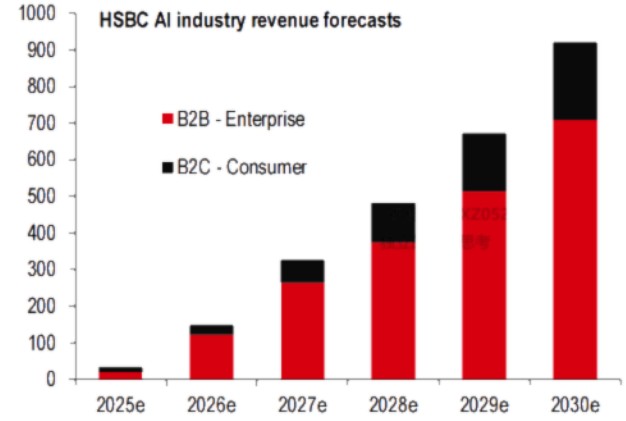
Source: Company reports, HSBC

27. US hyperscaler ROIC is expected to fall but remain sufficiently high



Source: Visible Alpha consensus estimates, HSBC estimates. See *Global Technology: Sustainable CSP capex to support AI trend through FY28e*, Frank Lee, 12 August 2026

28. HSBC's AI industry revenue forecasts (USDbn)



Source: HSBC estimates. See: *Global Tech Platform: B2B up, B2C down: our new AI TAM forecasts to 2030*, Nicolas Cote-Colisson, 9 June 2026

Debate #4: Capex vs monetization timing mismatch

The debate: *Even if the longer-term demand ultimately justifies the capex, concerns remain around the mismatch between the timing of investment today and the revenues expected to materialize over the coming years.*

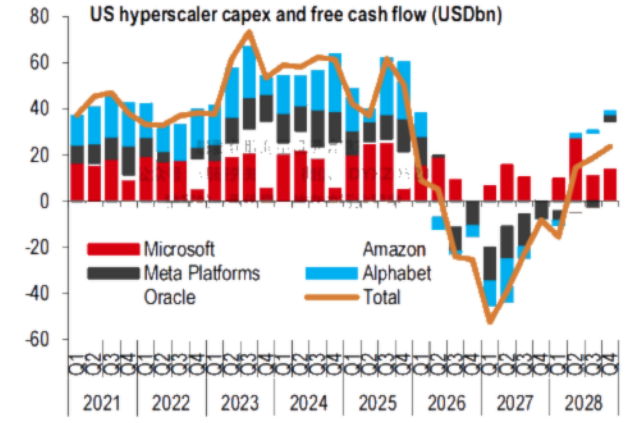
Our take: Free cash flow has already turned negative and is expected to remain under pressure until the end of 2027 as spending is incurred several years before the full revenue opportunity expected by 2030. Hyperscalers are increasingly bridging that gap through external financing, with debt issuance reaching nearly cUSD250bn so far in 2026 and borrowers expanding issuance outside the USD market (Chart 32). Leases provide another large source of funding. As discussed in detail by the HSBC credit team[3], the five hyperscalers and Coreweave collectively have more than USD1trn of lease commitments that have not yet commenced. They estimate that these commitments could support approximately USD400-500bn of off-balance-sheet borrowing, compared with just over USD500bn of bonds currently outstanding (Chart 31). Private credit can therefore finance a substantial part of the buildout without requiring the hyperscalers to issue the full amount directly in public debt markets. These structures also allow financing to be matched more closely with individual data-centre assets. A project can be funded according to its expected cash flows and useful life rather than relying solely on conventional 30y corporate debt.

The pressure is also beginning to affect capital allocation. Hyperscalers have grounded their buybacks programs, while Alphabet has issued USD10bn of equity, with an additional USD40bn available. Balance sheets still have significant financing capacity, but credit markets are becoming more sensitive to the scale of the commitments. Bond yields and CDS spreads have risen (Chart 33 and 34), while Oracle's downgrade to BBB- in July illustrates the pressure that can emerge when negative free cash flow, higher leverage and concentrated customer exposure occur together (S&P Global Ratings, 9 July 2026).

For equity investors, the more important consequence is the impact on returns. We expect GPU/AI servers to account for approximately 70% of hyperscaler capex in 2026, meaning a growing share of investment is going into assets with relatively short useful lives and significant depreciation. Depreciation and amortization could rise from USD110bn in 2024 to USD730bn-USD930bn by 2030 depending on the useful lives assigned to GPUs, while consensus expects hyperscaler ROE to decline from its 38% peak in Q2-26 to 23% by Q4-28 (Chart 20 and 22). The current capex cycle can therefore prove economically rational over the long term while still creating several years of weaker returns for shareholders. How quickly AI revenues scale against depreciation, financing costs and the larger asset base will determine the depth and duration of that ROE decline.

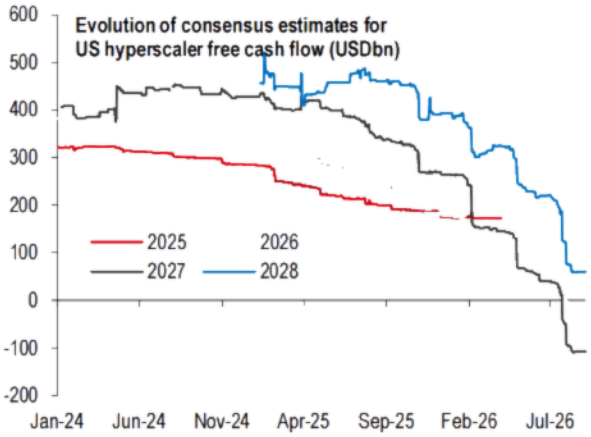
3 *The Credit Corner: AI financing – the other half trillion*, Song Jin Lee, 2 September 2026

29. Consensus expects US hyperscaler free cash flow to turn negative in 2026 and 2027



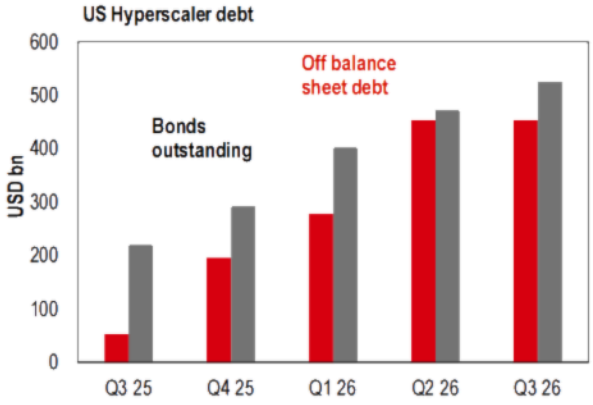
Source: Factset, HSBC

30. Consensus estimate revision for US hyperscaler FCF



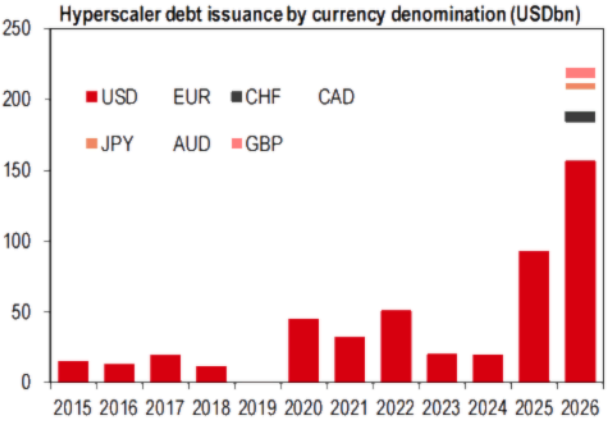
Source: Factset, HSBC

31. Hyperscaler bonds outstanding and off-balance sheet debt



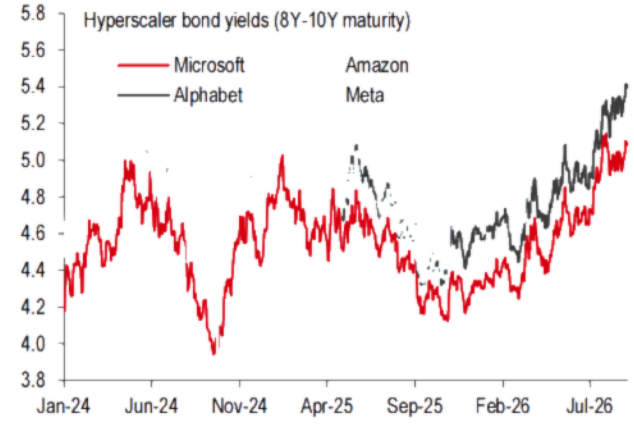
Source: HSBC estimates, Company Accounts, Bloomberg, Moody's. Note: Bonds of all currencies and estimated debt supported by leases from Microsoft, Alphabet, Meta, Amazon, Amazon and Coreweave. See *The Credit Corner: AI financing – the other half trillion*, Song Jin Lee, 2 September 2026

32. Hyperscaler debt issuance by currency



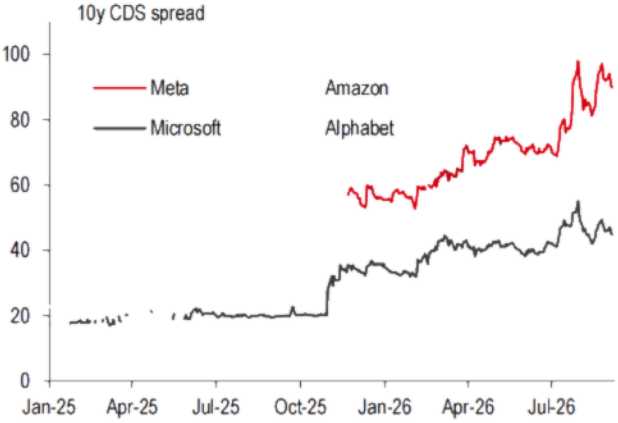
Source: Bloomberg, HSBC . Note: 2026 is year-to-date

33. Hyperscaler bond yields are rising…



Source: Bloomberg, HSBC

34. … along with their CDS spreads



Source: Bloomberg, HSBC

Debate #5: Korea memory: when to re-enter after the correction?

The debate: *Heightened volatility and the sharp correction in Korean memory has shifted the debate towards when and if it's the right time to re-enter the sector and whether enough optimism has been sufficiently priced out.*

Our take: We remain bullish on the Korean memory sector. The stocks are down c35% below their peak, and the correction has created a substantially better combination of valuations, positioning and earnings risk than existed at the June peak. For example, our market-implied earnings model[4] suggests that investors have compressed the expected duration of the earnings cycle for Samsung Electronics and SK Hynix from around six years to just 2.5 (Charts 35 and 36). This suggests that the market is pricing overly pessimistic assumptions about the AI investment cycle and/or the risk that competition from mainland Chinese suppliers drives DRAM prices lower. Positioning has also been substantially cleansed. Foreign selling across EM's tech trident has reached approximately USD150bn year-to-date, including USD60bn since June, while assets in leveraged single-stock ETFs have fallen from USD37bn at their peak to USD10bn (Charts 37 and 52). However, we are seeing signs that most of the selling pressure has now subsided. This should help reduce the volatility of the memory names which would then help lower the cost of equity and support valuations throughout the sector. (For example: Samsung Electronics and SK Hynix 90-day realized vol is 96% and 114%, respectively, vs 28% for the median stock in the FTSE EM index.)

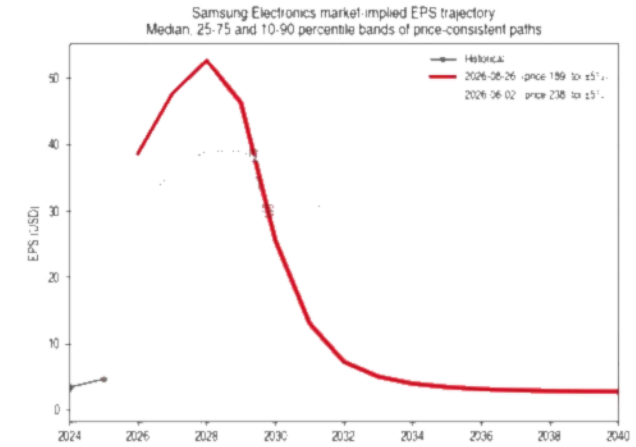
The fundamental backdrop, meanwhile, remains supportive. As mentioned above, we believe hyperscaler capex estimates can continue to move higher into year-end. This should drive further upside for Korea memory revenue/earnings estimates: we estimate that c70% of US hyperscaler capex is spent on GPUs and servers, of which nearly 30% is spent on memory. The HBM4 ramp-up, broader server DRAM and eSSD demand, and continued supply tightness should provide positive catalysts.

The main downside risk, is the prospect of greater memory supply from Chinese producers, notably ChangXin Memory Technologies (CXMT). CXMT mainly produces and competes in conventional DRAM, not HBM, yet that segment still accounts for about half of Korean memory revenue. CXMT's capacity is expanding quickly and it plans to increase production by c200k wafers a month to 300k by the start of next year. As a result, CXMT could reach up to 18% of DRAM market share by 2028, from <8% today, according to TrendForce. However, the shift toward long-term agreements (LTAs) should provide some insulation from this threat for the memory sector. For example, SK Hynix disclosed on its earnings calls that it completed multi-year agreements with c10 customers, while Samsung targets two-thirds of memory output under contracts lasting at least five years. Importantly, these agreements include deposits and price floors.

Shareholder returns is the next big story, with free cash flow for Samsung Electronics (005930 KP, KRW270,000, Buy, covered by Ricky Seo) and SK Hynix (000660 KP, KRW1,783,000, Buy, covered by Ricky Seo) likely to reach nearly USD300bn in 2026, according to consensus estimates, up from USD40bn in 2025. SK Hynix announced a KRW40trn buyback and cancellation, equivalent to 3.3% of shares, and lifted its 2025-27 payout target from "within 50%" to "over 50%" of cumulative free cash flow[5]. Samsung announced that it expects KRW90-110trn of shareholder returns in 2026, including KRW30trn of Q3 dividends, with further dividends and buybacks to be finalized in January (Yahoo Finance, 21 August 2026).

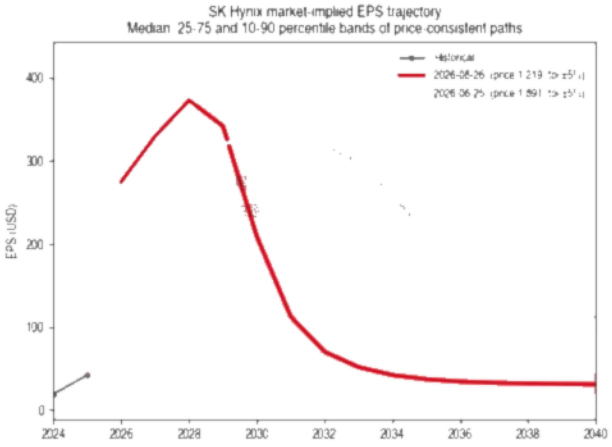
4 *AI capitulation: what's now priced? GEMs Equity Strategy*, Alastair Pinder, 3 August 2026
5 *Asia Memory: Cash dividend not good? Improvement in total shareholder return is on track in Korea memory*, Ricky Seo, 26 August 2026

35. Samsung Electronics: market-implied EPS trajectory today vs market peak



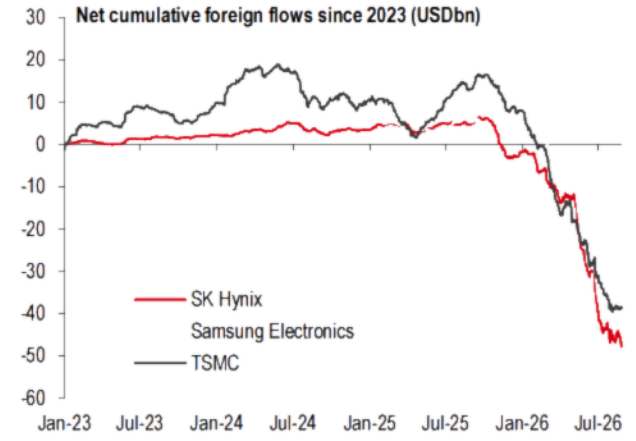
Source: Factset, HSBC

36. SK Hynix: market-implied EPS trajectory today vs market peak



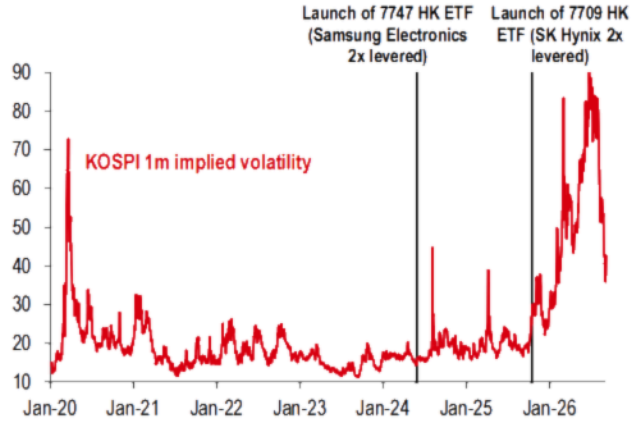
Source: Factset, HSBC

37. Foreign outflows from EM's tech trident have totalled USD150bn year-to-date



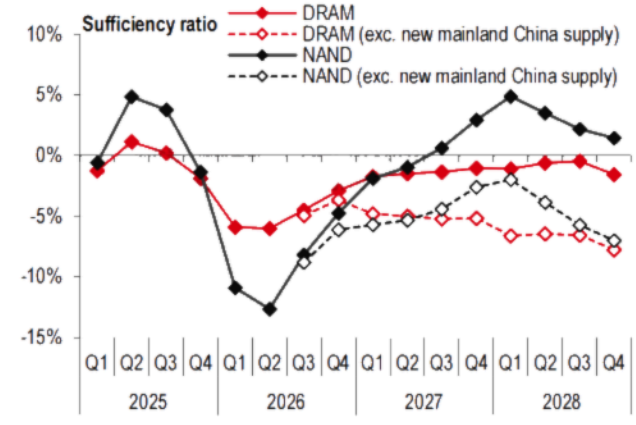
Source: Bloomberg, KRX, HSBC

38. Kospi implied vol has fallen sharply from peak levels



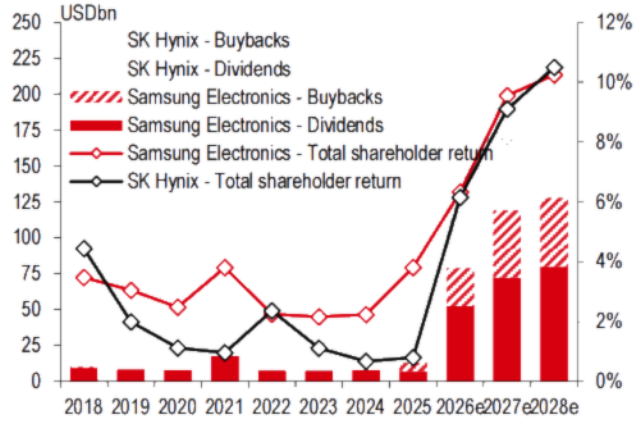
Source: Bloomberg, HSBC

39. DRAM and NAND sufficiency ratio



Source: Bloomberg, TrendForce, Gartner, HSBC

40. Korea memory total shareholder returns



Source: Factset, HSBC estimates. See *Asia Memory: Cash dividend not good? Improvement in total shareholder return is on track in Korea memory*, Ricky Seo, 26 August 2026

Debate #6: Mainland China AI: where does value accrue?

The debate: *China's AI trade raises a different set of questions from the US: whether domestic substitution can create a durable earnings cycle, how quickly local hardware can close the technology gap, and where value accrues within the AI supply chain.*

Our take: We prefer mainland China's domestically focused AI semiconductor and hardware stack where inference growth, localization and policy support are combining to drive demand. China's AI infrastructure cycle is broadening from model development into a domestic hardware buildout, as rapid inference growth increases compute requirements and restrictions on foreign GPUs accelerate localization. This also provides some diversification from Taiwan and Korea, where earnings remain more directly exposed to the US hyperscaler capex cycle.

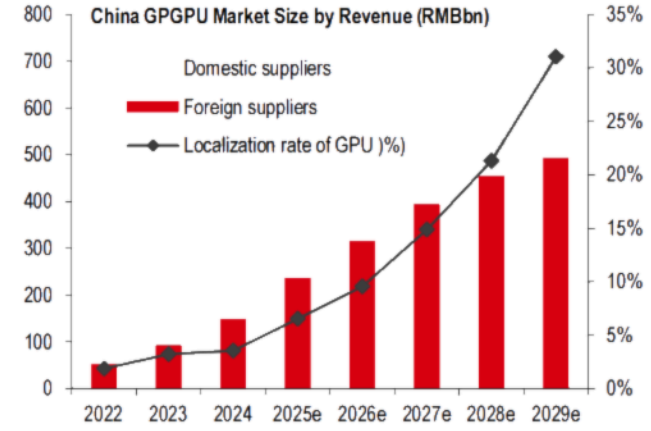
Mainland China's constraint remains supply rather than demand. The country is one of the world's largest markets for memory and compute, yet domestic fabs supply <15% of demand (3 May 2026, IndexBox). Bringing additional capacity online therefore allows domestic suppliers to capture demand currently served from overseas. Policy reinforces that opportunity, with semiconductor self-sufficiency remaining central to China's technology strategy as US export controls restrict access to advanced equipment and GPUs.

The technology gap also increases the amount of domestic hardware required. Chinese chips remain less efficient and interconnected than NVIDIA's leading products, requiring more chips, packaging, networking and power to deliver equivalent usable compute. This broadens the beneficiaries beyond accelerators into semiconductor equipment, servers, cooling, data centres and power infrastructure. Lithography remains the largest constraint. China has been excluded from state-of-the-art EUV equipment since 2019 and is unlikely to produce a domestic EUV machine before 2028 (24 December 2025, FinancialContent), forcing manufacturers to use less efficient processes at the leading edge.

Recent earnings suggest this localization cycle is already translating into stronger fundamentals. In their latest Q2 earnings results SMIC's AI-peripheral-chip revenue increased c40% q/q, utilization approached 94%, and wafer ASPs rose 5.7% q/q. Demand is also spreading into power-management, memory-controller, connectivity, MCU and industrial chips, creating a broader semiconductor multiplier. Accelerating domestic equipment orders should benefit suppliers including Naura, AMEC, and ACM Research as production moves onshore.

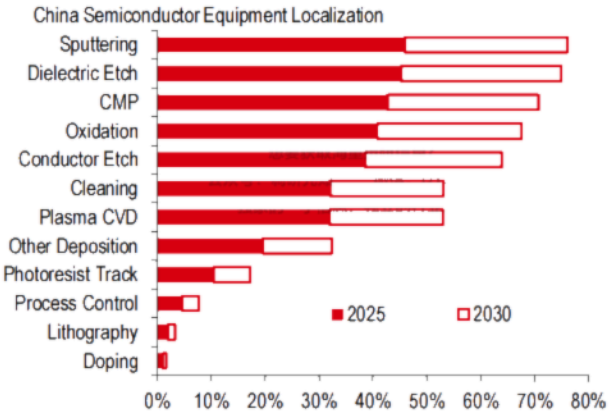
We are more selective elsewhere in China's AI stack. Optical module companies with greater US revenue exposure, including Innolight and Eoptolink, face structural pressure from FCC regulatory restrictions (source: Bloomberg, 5 August 2026). Monetization among internet companies also remains mixed. Alibaba's external cloud revenue growth accelerated to 45% y/y in Q2-26 and AI products recorded a twelfth consecutive quarter of triple-digit growth. Tencent shows the cost of that investment: new AI initiatives reduced operating profit by cRMB10.5bn, while capex reached RMB52.8bn, up 176% y/y and pushing free cash flow negative as of Q2. For now, the clearer earnings opportunity remains in supplying China's AI infrastructure buildout rather than monetizing the applications running on it.

41. China's push for localization should support the tech stack…



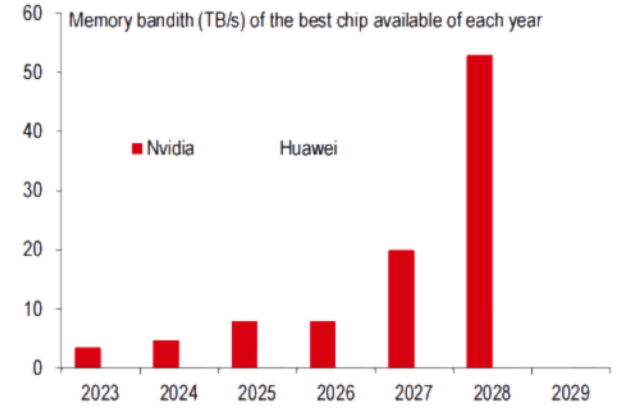
Source: Frost & Sullivan, HSBC

42. … particularly semiconductor equipment



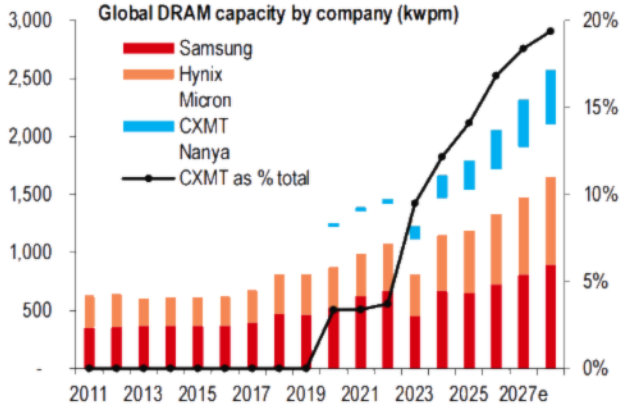
Source: Bloomberg estimates, HSBC

43. Nvidia vs Huawei memory bandwidth of best chips



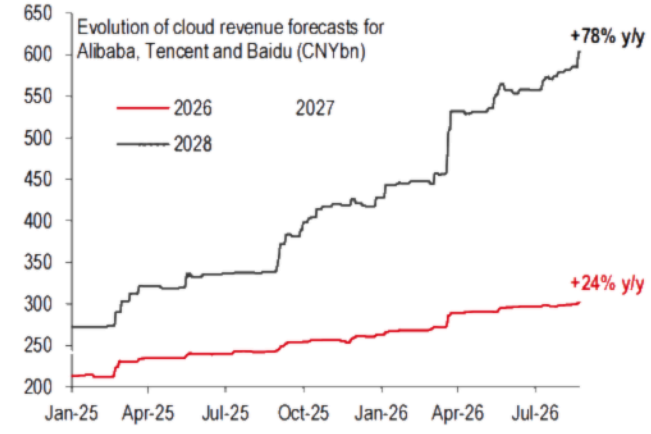
Source: Nvidia Roadmap, Huawei Roadmap, SemiAnalysis, CFR, HSBC

44. CXMT expected to take a larger market share of DRAM market



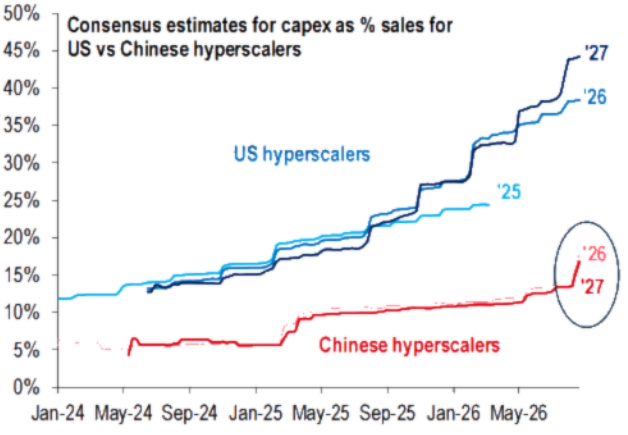
Source: TrendForce estimates, Company data

45. Chinese cloud revenue consensus forecasts have been upgraded significantly…



Source: Factset, HSBC

46. … but capex estimates for the Chinese internet names were also raised meaningfully recently



Source: Factset, HSBC

Debate #7: how to position for AI within a global portfolio

The debate: *Amid volatility and sharp rotation within the AI space, the critical question is where to be positioned now.*

Our take: Our preferences within the AI space is for areas with the strongest demand and monetization outlook and where valuations are preferably not stretched.

That keep us increasingly tilted towards EM, particularly Korean memory, Taiwan semiconductors and advanced packaging, and mainland China's semiconductor equipment and power infrastructure. These markets provide exposure to areas where demand remains strong and capacity remains constrained, while valuations are generally less stretched than across parts of the US AI complex. We outline our bullish case for Korea and mainland China in the previous sections.

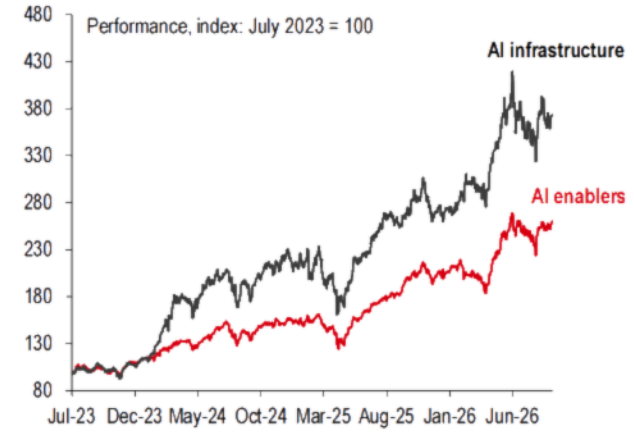
In Taiwan, the next leg of the AI trade is broadening beyond GPUs as hyperscalers increase their use of custom ASICs. More chip designs and greater complexity increase demand for advanced packaging, shifting one of the principal supply bottlenecks further downstream. Major Asian OSAT companies have raised 2026 capex by more than 80%, yet CoWoS equipment lead times can extend to 12 months and companies have indicated that capacity will remain tight into at least 2027-28. Persistent capacity constraints should support both pricing and utilisation even as packaging investment accelerates.

Within the US, we prefer semis > hyperscalers > software. Semiconductor and cloud infrastructure companies retain direct exposure to rising training and inference volumes, with revenues less dependent on which individual model or application ultimately wins. We would fade the recent rotation into software, where we see the move as driven more by short covering than a material improvement in underlying fundamentals.

In Europe, we also focus on semis. HSBC's analysts note that recent datapoints suggest the analog upcycle is strengthening, with multiple analog peers signalling further price increases[6]. We see this as consistent with demand surprising positively (with price rises likely to flow through at different cadences by end-market). Higher volumes should help underutilisation charges normalise faster than expected, supporting a quicker recovery in profitability.

6 *European Semiconductors: IFX/STM: From trough to throttle*, Adithya Metuku, 29 June 2026

47. US AI performance by theme



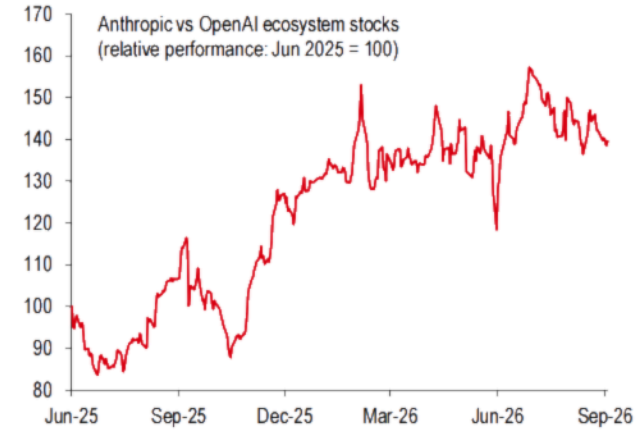
Source: Bloomberg, HSBC

48. Magnificent 7 valuations have de-rated significantly



Source: Factset, HSBC

49. Anthropic vs Open AI ecosystem performance



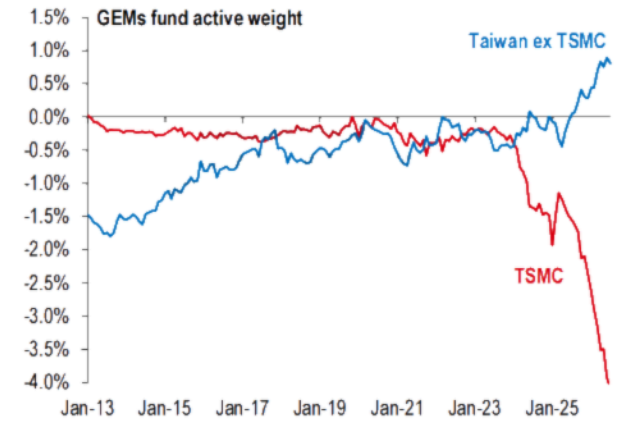
Source: Bloomberg, HSBC

50. China internet vs China tech hardware and semis performance



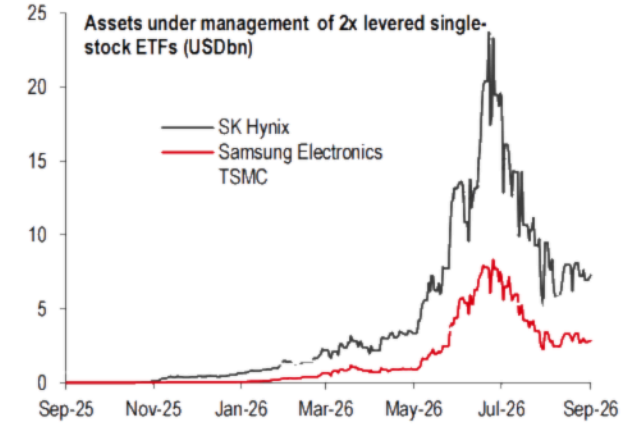
Source: Bloomberg, HSBC

51. GEMs funds are rotating deeper into Taiwan's supply-chain names



Source: Factset, HSBC

52. Leveraged ETF AUM has declined significantly



Source: Bloomberg, HSBC