

NAVIGATING CHINA INTERNET/AI MODELS

Focuses on harness & model performance vs. price; multi-modal highlights; top AI/apps tracker

Navigating China Internet / AI models

Explore >



From cost efficiency, rise in model intelligence to new frontier/increasing pricing power of Chinese AI models (our [LLM primer](#)), we now estimate Chinese AI model Annualized Recurring Revenues (ARRs), in aggregate, to reach US\$13bn by year-end 2026 (prior: US\$10bn) on higher demand for Chinese models/faster ARR ramp up with our lifted Zhipu API/MiniMax group year-end ARR at US\$2.5bn/US\$1bn. We expect competition for the best performance-to-price balance to intensify amongst Chinese AI models, evident from MiniMax's H3's open-weight approach priced at ~1/3 of peers' pricing, and DeepSeek's latest V4 Flash (at 284bn parameters, reaching close to GLM5.2 in front-end coding) and Alibaba Qwen3.8 Max (at 2.4T parameters) official launches, both reached new performance-per-dollar efficient frontiers (as per Arena.ai for front-end coding). We believe Zhipu's GLM-5.5 and MiniMax's M3 next update/M3 Pro models will be next to watch.

We highlight the latest developments around **1) multi-modal**: Launch of MiniMax's H3 on July 31 where full modal-capabilities, open-weight approach and significantly lower pricing/costs are key highlights, while Bytedance's Seedance 2.5 (API channel launch on July 31) continues to lead in long-duration generation (30s max length) and consistency (50 multi-modal reference inputs); expanding agentic functionalities will drive further content creation adoption/penetration in our view; **2) our anticipation of rise in competition of higher-end coding/agentic-focused models**, at the 2-5 trillion parameter size class over the next few months as scaling continues, where Alibaba Qwen3.8 Max launched at 2.4T total parameter size moved up in Arena.ai's overall Code Arena and front-end coding rankings; **3) coding/agentic performance uplift in the smaller-sized/agentic-focused segment**, with DeepSeek V4 Flash official API launch, outperforming its earlier V4 Pro Preview; and Tencent's Hy3 marked performance uplift vs. Hy3 Preview. We believe post-training around supervised fine-tuning (SFT) has driven the key performance enhancements; **4) harness/agentic applications**, where internet mega-caps are leading in workspace applications, Tencent's Workbuddy and Alibaba's product revamp with QwenWork; and **5) continued focus on AI capex**

Ronald Keung, CFA

+852 2978 0856
ronald.keung@gs.com
Goldman Sachs (Asia) L.L.C.

Lincoln Kong, CFA

+852 2978 6603 lincoln.kong@gs.com
Goldman Sachs (Asia) L.L.C.

Timothy Zhao

+852 2978 2673
timothy.zhao@gs.com
Goldman Sachs (Asia) L.L.C.

Steve Qiu

+852 2978 2672 steve.qiu@gs.com
Goldman Sachs (Asia) L.L.C.

Eunice Liu

+852 2978 7472 eunice.liu@gs.com
Goldman Sachs (Asia) L.L.C.

Luqing Zhou

+852 3465 4207 luqing.zhou@gs.com
Goldman Sachs (Asia) L.L.C.

Damian Xie

+852 2978 1398 damian.xie@gs.com
Goldman Sachs (Asia) L.L.C.

Iris Xiao

+852 2978 0202 iris.xiao@gs.com
Goldman Sachs (Asia) L.L.C.

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. Analysts employed by non-US affiliates are not registered/qualified as research analysts with FINRA in the U.S.

spending, on the back of anticipated increasing domestic compute supply over 2H26. Factoring in the latest industry landscape and ARR developments, we raise Z.AI and MiniMax 2026E revenues by 35%/63% respectively, fine-tune Z.AI/increase MiniMax 2027-28E estimates and maintain Neutral on Zhipu, Buy on MiniMax on risk-reward.

Estimate changes for Z.AI and MiniMax: Factoring in latest industry landscape and ARR developments, for Z.AI, we change our 2026-28E revenues by +35%/+4%/-2% and our earnings estimates by -3%/-3%/-1% on strong 2026 YTD ARR ramp up, while factoring in intensified competition in longer term over larger parameter/high-end coding models and sustained R&D investments. For MiniMax, we raise our 2026-28E revenue by 22-64% on favorable mix shift towards fast growing API business from AI-native products, and fine-tune our net income estimates by -3% to +1% to reflect higher R&D and other operating expenses. Our TPs change by -14%/-7% for Z.AI/MiniMax respectively on estimate changes and latest share count (MiniMax). We also refresh our bull/bear scenarios within. 1H26 preview: For Z.AI, we expect revenue to reach Rmb1.8bn (i.e. US\$0.3bn) and Rmb-1.8bn (i.e. US\$-0.3bn) adj. net loss, on the back of training cost of Rmb2.2bn. For MiniMax, we expect revenue to reach US\$120mn and US\$-168mn adj. net loss, on the back of R&D expenses of US\$0.2bn.

Our June app tracker shows the top 400 mobile apps have grown total time spent by +9% yoy (vs. +5% in June 2025), with **Weixin** time spent growth (+2% yoy), continued strong **Douyin** app engagement performance (main app +27% yoy/Lite +30% yoy) and domestic AIGC application total engagement up by +2% mom. We also note ByteDance's continued share gain (33% within top apps, up +6% from 2 years ago) via short-form video apps (Douyin/Red Fruit) and AIGC applications (Doubao). See specific growth by verticals in the below Top apps section.

Our key ideas/stock picks: We continue to rank Cloud & Data Centers as our preferred sub-sector (#1) with key ideas: GDS, VNET, Alibaba and Kingsoft Cloud on proliferation of AI token demand from agentic AI with improving cloud pricing/margin upside and hyperscalers' further capex ramp up into 2H26 as key catalysts. Within AI models, we highlight MiniMax on upside skewed risk-reward. We view the company's articulated strategy in addressing M3 model competitiveness and H3 multi-modal launch as positive drivers for ARR ramp up that could drive significant room for valuation repair over 2H 2026 (vs. ARR of US\$400mn+ in early June 2026 before M3 launch, at a run-rate likely at half of Zhipu at the mid-year mark 2026).

- **H3 multi-modal model launch from MiniMax; Open weight approach and 30%-50% pricing levels of incumbents:** MiniMax's H3 model was released on July 31, more than a video-generation model but a new architecture as an Omni-model, that understands all modals (text, image, video, audio, music) with stable output across video, image, audio based on user needs/prompts. We see implications to the multi-modal industry landscape (offering alternatives to existing video generation products) given its open-weight approach at 30%-50% pricing levels on architectural efficiencies, yet still healthy industry pricing (vs. foundation text models) given the less fragmented market landscape and tight computing resources where demand significantly outpaces capacity.
- **The contest for Harness/workspace agentic applications:** We expect China AI model companies to increasingly focus on positioning their harness/agentic

applications as key entry points, especially in coding/workspace (e.g. Zhipu's ZCode, Tencent's Workbuddy, and Alibaba's Qoder which are aggregate platforms that support the full suite of AI models) as model companies attempt to close the loop in capturing more real-life coding and agentic data for scaling. Co-work and industry expert agentic products are the next priorities. Tencent's WorkBuddy has ranked top in June monthly visits among China desktop agent products (see [Exhibit 10](#)), on the back of rich ecosystem support, rapid feature iteration, and safety governance. We also note difference between mega-caps/individual lab offerings, where WorkBuddy provides in-house (Hy) + third-party models and monetizes on both 1P token & 3P API resale/revenue-sharing, while independents host their in-house models with focus more on promoting their own coding plan/API.

- **Multiple large parameter high-end coding model launches and potentially stricter access to the most advanced Chinese AI models outside of China:** After Kimi K3 launch, we anticipate for further new Chinese AI model updates/launches over 2H26 (Zhipu GLM, MiniMax M3 Pro, Doubao Seed and more) with significantly larger total parameter sizes of 1-5 trillion across Chinese AI model players. Coding/programming segment competition will intensify over 2H26, in our view. The addition of multi-modal/visual understanding will also be the next upgrades amongst Chinese AI foundation models. With Chinese players' focused on post-training data, supervised fine-tuning (SFT) and reinforcement learning that is bringing coding performance uplift across the board (Tencent's Hy3 performance uplift vs. Hy3 preview as an example, and we anticipate Minimax M3.1 to see strong uplift vs. M3), user switching between models can be fast where pricing/cost efficiencies/speed will become other key drivers to user adoption of Chinese AI models (see our [takeaways from MiniMax Founder/CEO meeting](#)). We also note [news reports](#) of Chinese regulators considering restrictions on foreign downloads of model weights and overseas transfer of training data.
- **Continued shift to license terms for Chinese open-weight models; competitive API pricing for the lower end agentic-focused segment:** We note Chinese open-weight models have gradually started to impose usage restrictions, whereby large-scale commercial adoption mandates revenue-sharing payments via formal licenses. We believe open source models' disclosed ARR's are likely understating total deployment and revenue potentials, and we expect more shifts to open weight (with Community License, i.e., commercial terms if for commercial use) among Chinese AI models down the road. We expect API pricing and therefore gross margins to remain under pressure at the lower-end pricing segment (around US\$0.1-0.2 per 1M tokens) into the second half, as China's AI model players have significant cash buffers post-fund raising to subsidize competitive pricing. As a result, we see financial strength as one of the three most important metrics (alongside pricing power and cost efficiencies) in assessing a Chinese AI model company. We note value-for-money agentic models such as Tencent's Hy3 and Xiaomi's MiMo 2.5 continued to gain share over the latest week following DeepSeek's announcement of rush-hour price hike.
- **Increasing domestic compute supply:** We are seeing positive signals on domestic compute adoption for AI inference (and prospectively training). On July 23, Alibaba

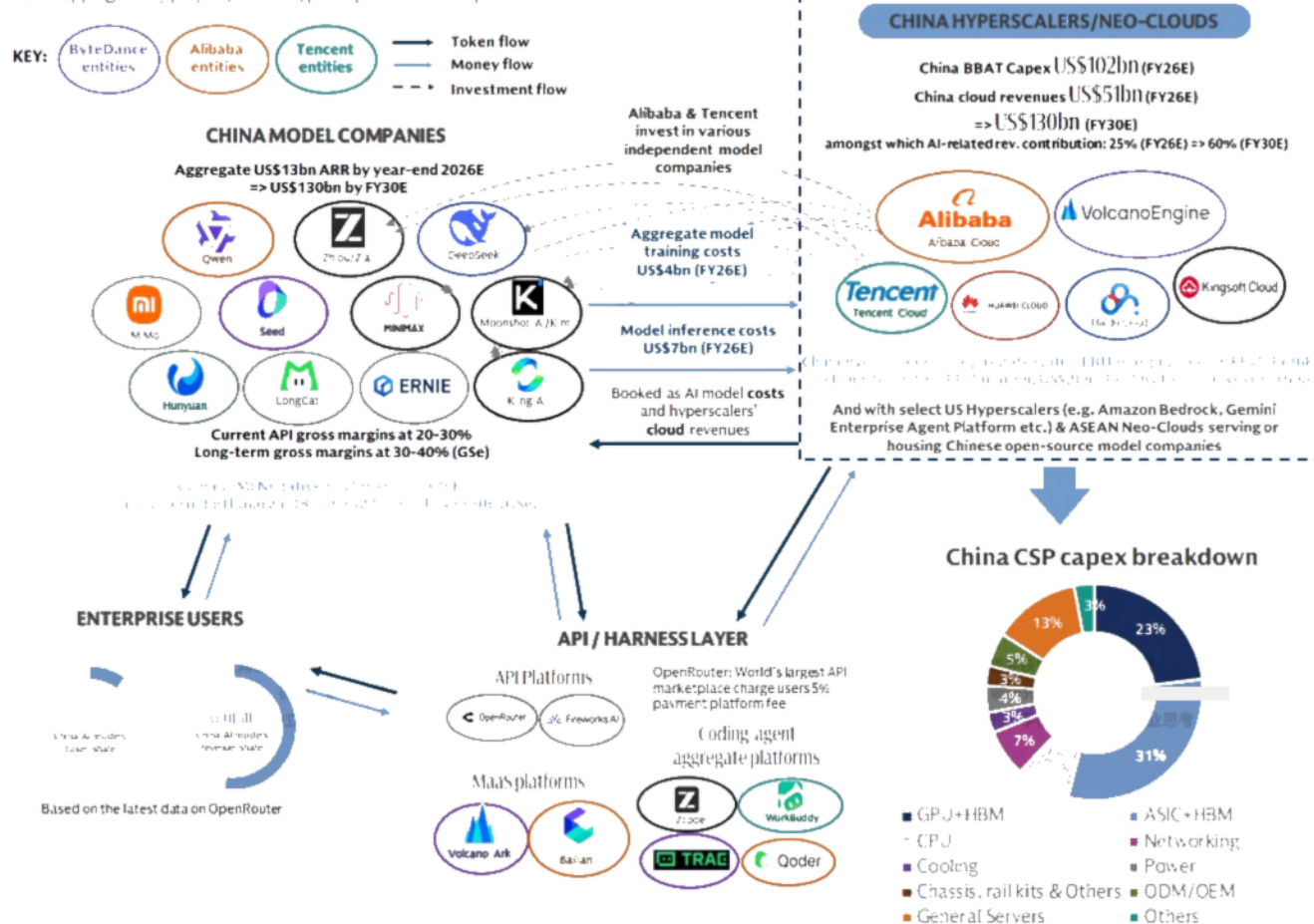
disclosed that its T-head Lingjun Zhenwu M890 Super Node has gone live on the Bailian MaaS platform for inference and has been adapted to Qwen3.8 Max, making it the first super-node in China capable of running AI models with over 2T parameters. Meanwhile, spending commitments from both hyperscalers (Tencent Cloud's announcement at WAIC to aggressively expand domestic AI chip deployment to lower inference costs) and pure plays (Zhipu's reported completion of a 1GW data center powered entirely by domestic chips, alongside its acquisition of compute software team XCore Sigma to improve chip utilization) suggest that model players continue to strengthen their computing infrastructure post-recent fund raising, as competition shifts towards larger-parameter models.

See also from our recent reports

- Navigating China AI models: LLM primer: Critical point of intelligence for global proliferation, July 10, 2026
- Navigating China AI models: New frontier from Kimi K3; from cost efficiency, intelligence to pricing power of Chinese models, July 18, 2026

CHINA AI MODEL & HYPERSCALERS ECOSYSTEM

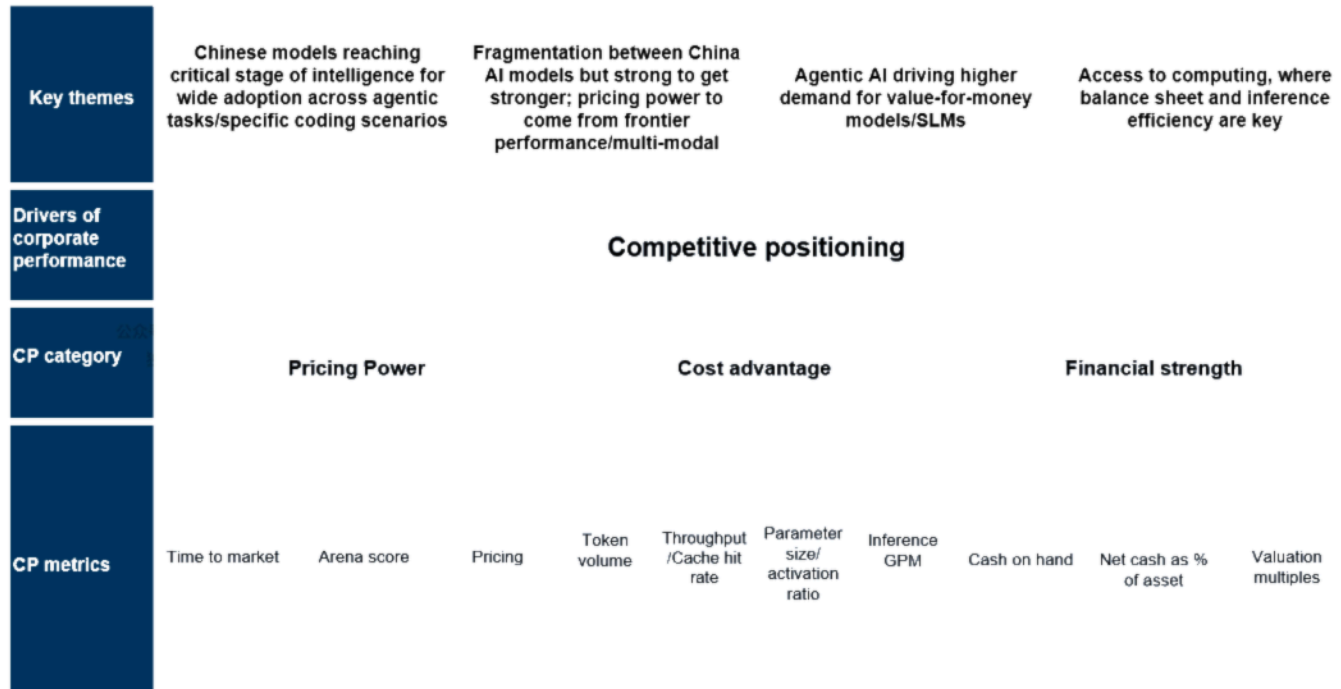
Our mapping of key players, revenue/profit pools & interdependencies



1) The calculation of China cloud revenues includes Alibaba, Tencent, ByteDance, Baidu, and Kingsoft Cloud. 2) China CSP capex breakdown is based on GS Greater China Tech team estimates; 3) Companies included are for illustrative purposes and may not be an exhaustive list.

Source: Company data, Goldman Sachs Global Investment Research

Exhibit 1: Competitive positioning framework for AI model players

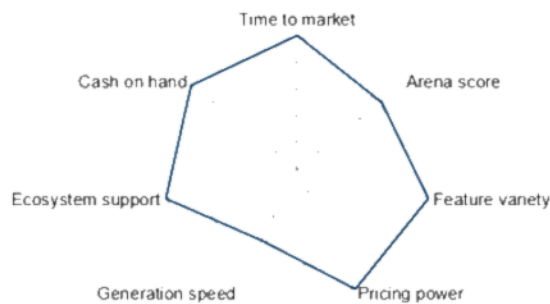


Source: Goldman Sachs Global Investment Research

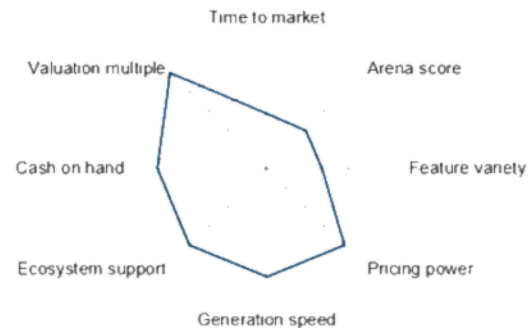
Exhibit 2: Overview of competitive analysis for key multi-modal labs' positioning

ByteDance: leading in model performance with strong financial/ecosystem back-up for quick adoption & commercialization

Kuaishou: prioritizing enterprise & professional users with strong pricing power and clearer monetization path, supporting a more manageable ROI



Alibaba: abundant resources to support rapid iteration for catching up in performance / feature offerings



MiniMax: H3 delivers full-modal generation with strong performance improvements and cost efficiency



Bytedance's (Not Covered) relevant metrics are calculated by analyzing the industry and companies that we cover, and then extrapolating them to Bytedance. N/A for Bytedance's valuation multiple

Source: Goldman Sachs Global Investment Research

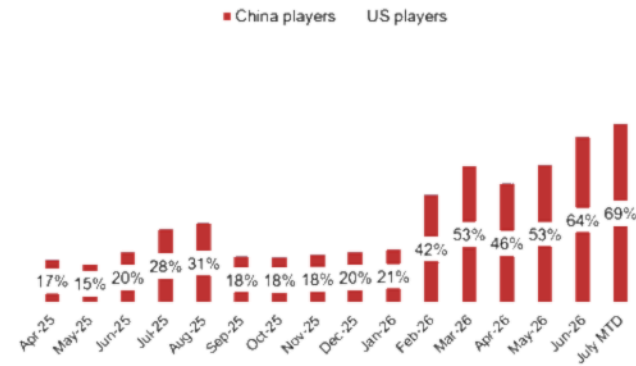
Exhibit 3: China harness/agent product landscape



***Note:** Alibaba is set to consolidate QoderWork, Wukong and MuleRun (prior focus mainly on overseas market) into **QwenWork**, now under beta testing

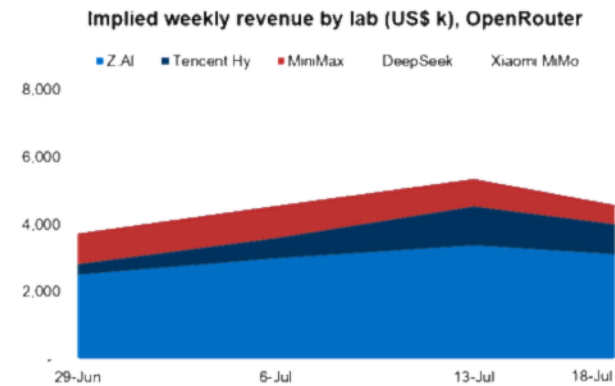
Source: Company data, Data compiled by Goldman Sachs Global Investment Research

Exhibit 4: China model players continue to capture market share by token consumption on OpenRouter



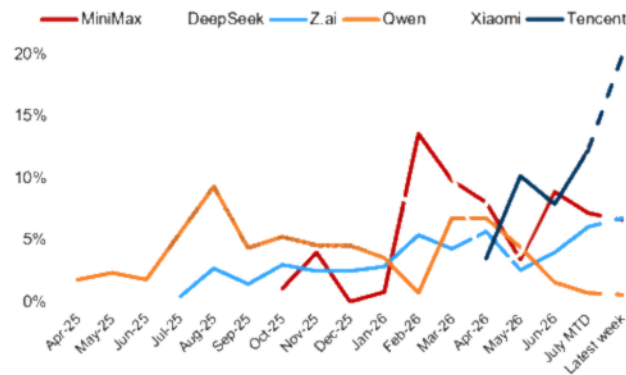
Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

Exhibit 5: Implied weekly revenue rise according with token growth, especially for higher pricing coding models



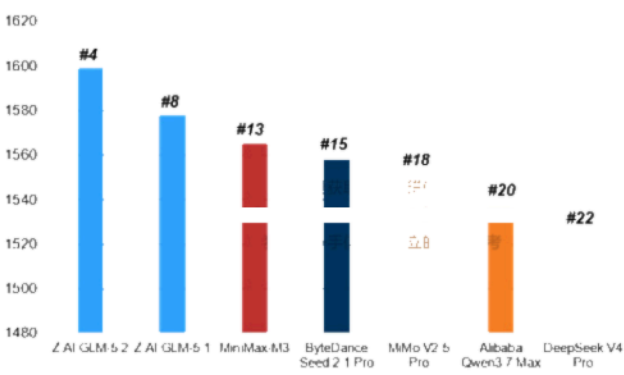
Source: OpenRouter, Company data, Data compiled by Goldman Sachs Global Investment Research

Exhibit 6: We see leading share gain from value-for-money agentic models, esp. Xiaomi/Tencent over the latest week post DeepSeek rush hour price hike
Market share of Chinese AI model players, of token consumption on OpenRouter



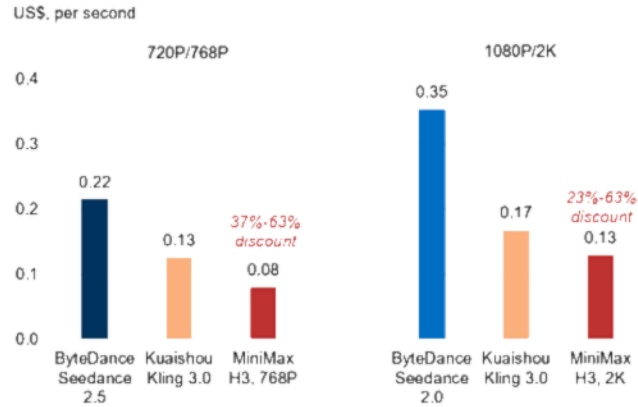
Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

Exhibit 7: China models' ranks on Code Arena leaderboard for full-stack web development tasks
As of late July, 2026



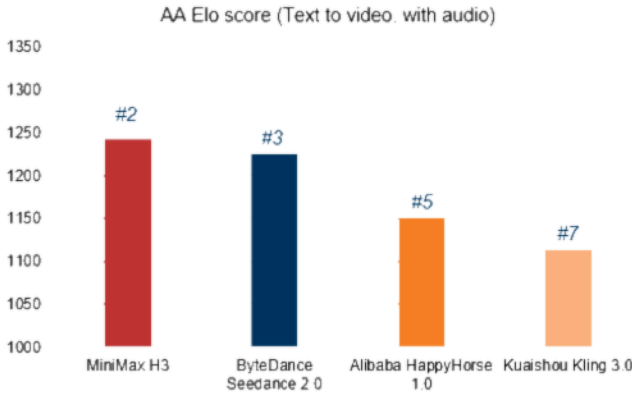
Source: LMArena, Data compiled by Goldman Sachs Global Investment Research

Exhibit 8: Attractive H3 API pricing at 1/3~2/3 of peer level
US\$ per second



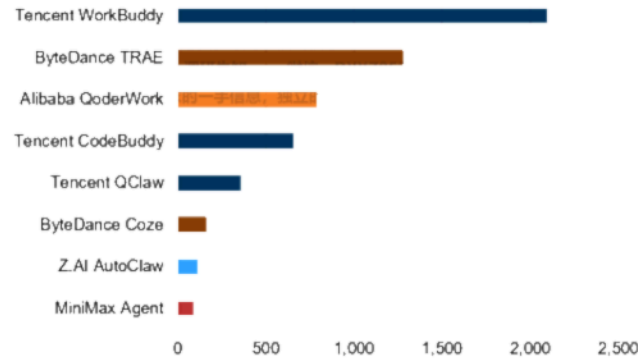
Source: Artificial Analysis, Data compiled by Goldman Sachs Global Investment Research

Exhibit 9: We note substantial improvement in MiniMax H3 ranking on AA Text to video leaderboard (with audio)



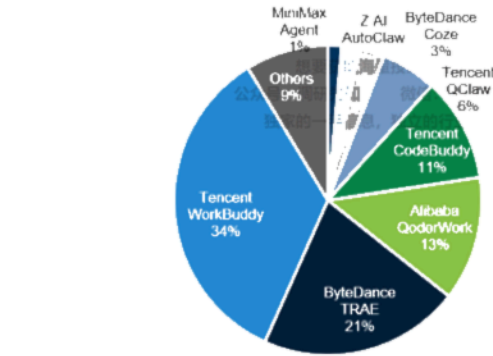
Source: Artificial Analysis, Data compiled by Goldman Sachs Global Investment Research

Exhibit 10: WorkBuddy ranks top on June monthly visits among China desktop AI agent products



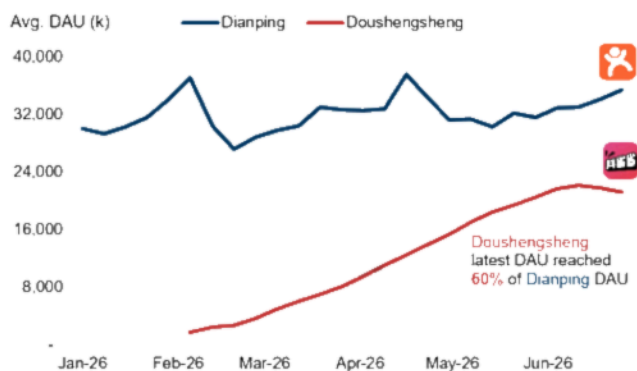
Source: Analysys, Data compiled by Goldman Sachs Global Investment Research

Exhibit 11: ...where WorkBuddy takes 34% market share, followed by ByteDance TRAE and Alibaba QoderWork



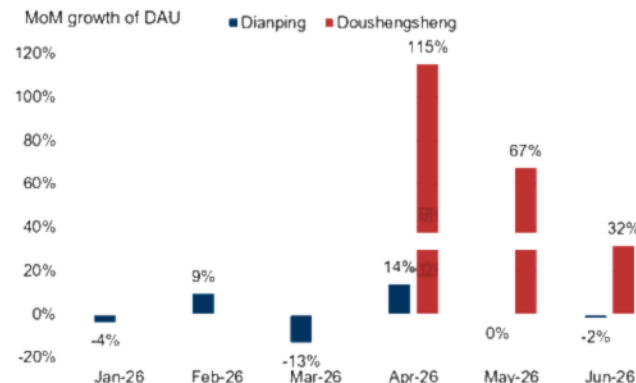
Source: Analysys, Data compiled by Goldman Sachs Global Investment Research

Exhibit 12: Doushengsheng (ByteDance's local service app) surged in traffic since Feb launch with July MTD DAU exceeding 20mn (~60% of Dianping)...



Source: QuestMobile, Data compiled by Goldman Sachs Global Investment Research

Exhibit 13: ...yet MoM DAU growth has begun to slow, suggesting potential normalization in competition into 2H as aggressive subsidies of Doushengsheng gradually scale back



Source: QuestMobile, Data compiled by Goldman Sachs Global Investment Research

AI foundation models/latest API tokens tracker: Proliferation of tokens at Chinese players driven by coding, AI agents

Ongoing token share gain by China open-source/open-weight models

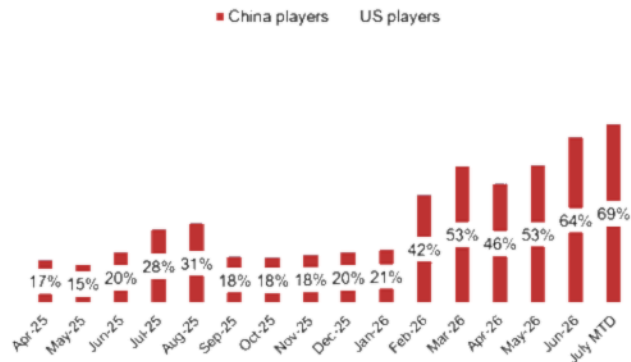
Exhibit 14: China models have dominated top ranks of OpenRouter token usage leaderboard for several weeks

As of Jun 21th, 2026 (Rolling 7 days)			As of Jul 5th, 2026 (Rolling 7 days)			As of July 24th, 2026 (Rolling 7 days)		
Ranking	Model	Tokens usage	Ranking	Model	Tokens usage	Ranking	Model	Tokens usage
1	DeepSeek V4 Flash	4.93T	1	DeepSeek V4 Flash	5.34T	▲ 1	MiMo-V2.5	9.81T
2	MiMo-V2.5	3.94T	2	MiMo-V2.5	4.38T	▲ 2	Hy3*	8.85T
3	MiniMax M3	3.77T	3	MiniMax M3	4.11T	3	DeepSeek V4 Flash	5.59T
4	Hy3 preview	3.63T	4	Hy3 preview	3.13T	▲ 4	GLM-5.2	3.47T
5	LongCat-2.0	2.56T	5	GLM-5.2	2.58T	▲ 5	DeepSeek V4 Pro	3T
6	DeepSeek V4 Pro	2.53T	6	DeepSeek V4 Pro	2.36T	6	MiniMax M3	2.72T
						▲ 13	Kimi K3	1.1T

*Combined usage of Hy3 (free) at 6.79T and Hy3 at 2.06T

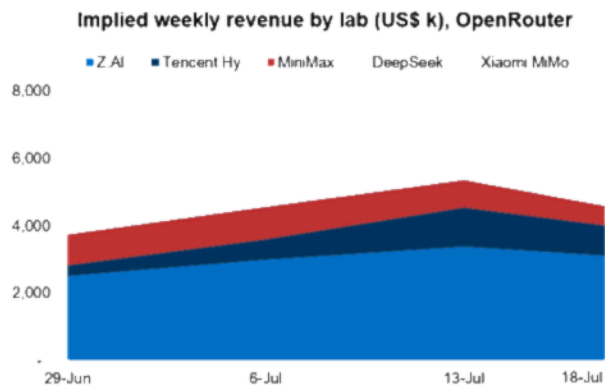
Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

Exhibit 15: China model players continue to capture market share by token consumption on OpenRouter



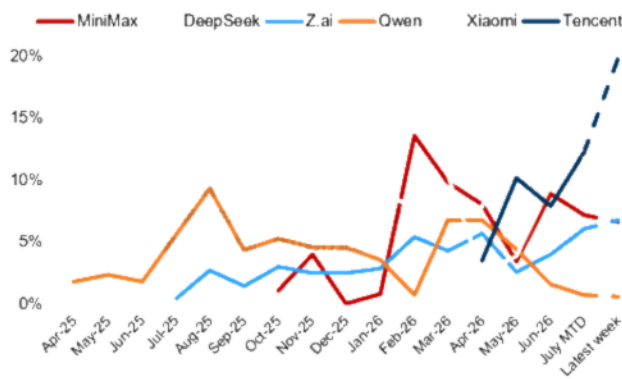
Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

Exhibit 16: Implied weekly revenue rise according with token growth, especially for higher pricing leading coding models



Source: OpenRouter, Company data, Data compiled by Goldman Sachs Global Investment Research

Exhibit 17: We see leading share gain from value-for-money agentic models, esp. Xiaomi/Tencent over the latest week post DeepSeek rush hour price hike
Market share of Chinese AI model players, of token consumption on OpenRouter



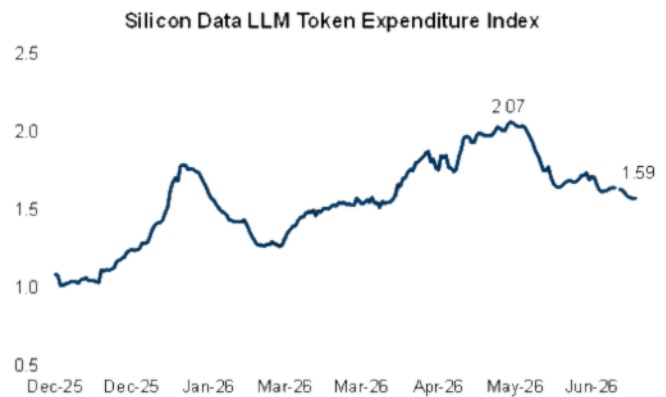
Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

Exhibit 19: China AI models now only account for limited share of spend across different tasks, per OpenRouter
As of week starting Jul 1, 2026



Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

Exhibit 18: Blended LLM token price (SDLLMTK) has been declining since early June, also potentially indicating rising adoption of cost effective China models
US\$ per 1M tokens, date as of July 19



The Silicon Data LLM Token Expenditure Index (SDLLMTK) is a daily blended measure of what the market actually pays to run LLM inference across frontier and open-weight models.

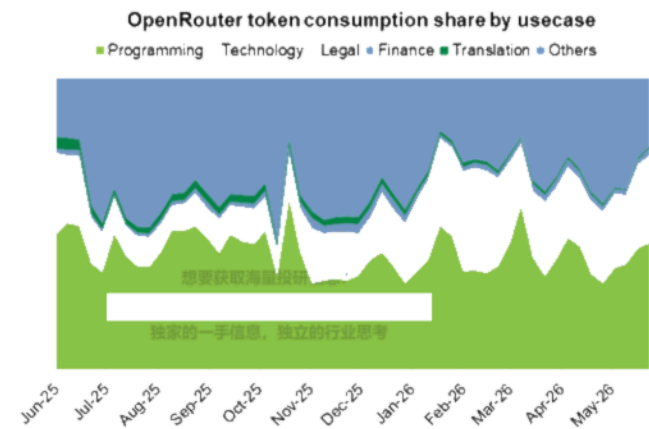
Source: Silicon Data, Bloomberg, Data compiled by Goldman Sachs Global Investment Research

Exhibit 20: While China AI models already capture majority of token usage on price competitiveness, esp. in long duration coding/agentic tasks
As of week starting Jul 1, 2026



Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

Exhibit 21: Token consumption is concentrating in complex, high value scenarios like programming/technology



Source: OpenRouter, Data compiled by Goldman Sachs Global Investment Research

AI multi-modal generation/video models

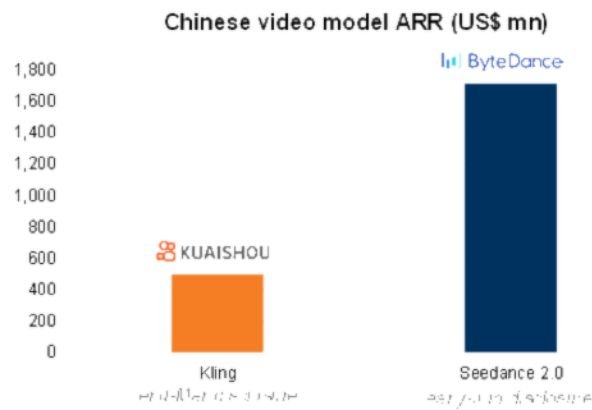
Highlights and surprises from MiniMax H3 launch: Open-weight, performance step-up and pricing

We view MiniMax H3's launch as a meaningful comeback to the frontline of AI video generation (see performance improvements with AA ranking [Exhibit 27](#)), with its planned open-weight release, unified full-modal generation and materially lower pricing, which potentially shift the competition from single-prompt visual quality toward commercially ready production tools. Specifically, we highlight:

- **Unified full-modal understanding and generation:** H3 jointly understands text, images, video and audio as a single creative context, while generating video with native stereo audio rather than treating generation, reference handling and editing as separate specialist tasks. For example, users can specify a camera movement from one video reference, a character from an image, and singing style from an audio reference within one instruction, allowing H3 to infer the relationships among inputs and produce a coherent audiovisual output.
- **Production-ready cost efficiency enabled by architectural innovation:** Per second pricing is set at ~1/3 of peers at 2K and ~1/2 at 768P, with extreme value-for-money on the back of redesigned H3-VAE, which delivers 4X sequence-length efficiency, heterogeneous training optimization in H3-Omni Transformer, and in-context regeneration, through which the base model regenerates higher-resolution results using the original multimodal context rather than relying solely on a standalone super-resolution module.
- **Expanded output specifications and reference capacity:** H3 supports up to 15s video generation at 24 FPS, native stereo audio and major aspect ratios, while accepting up to 9 images, 3 video clips and 3 audio clips per request (12 files in aggregate, substantial improvement from Hailuo 2.3), alongside prompts of up to 7,000 characters. We note that multi-modal references beyond text prompt help anchor character/motion/voice consistency and visual style, which could better support complete advertising beats, product demonstrations, game or UI sequences and short narrative content.
- **Granular control lowers “lottery rate” and project cost:** Stronger instruction following, brand/text rendering and V2V motion transfer enable creators to retain approved elements while making targeted changes, reducing full-video regeneration and improving iteration efficiency across commercial workflows. On LibTV (AI video creation platform hosting multiple APIs), limited-time H3 pricing starts at Rmb0.1/s for 768P and Rmb0.2/s for 2K, and the estimated cost per advertising video/short drama episode is low at Rmb10/Rmb20 each, with high usable output rate effectively bringing down per-clip cost.

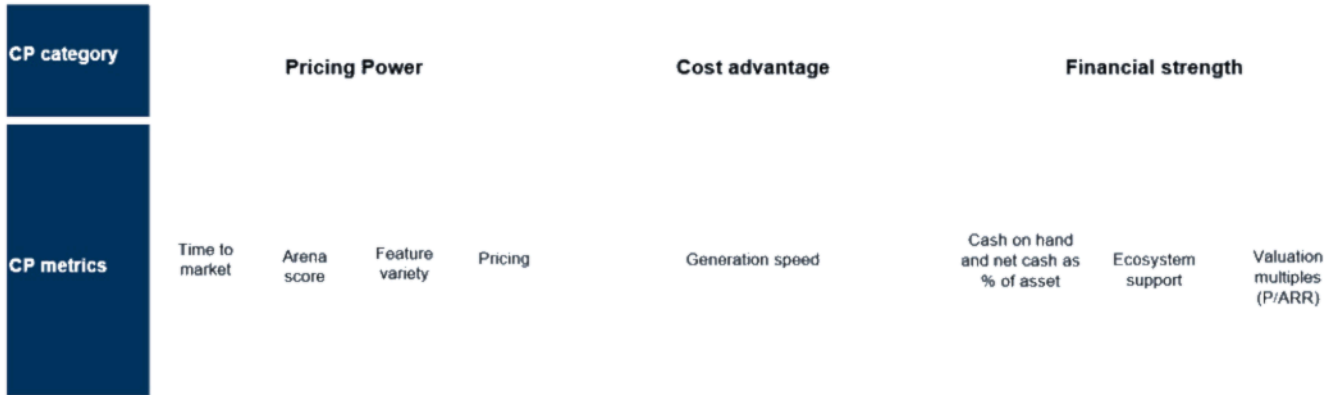


Exhibit 22: Latest disclosure of Chinese AI video model ARR (US\$ mn)



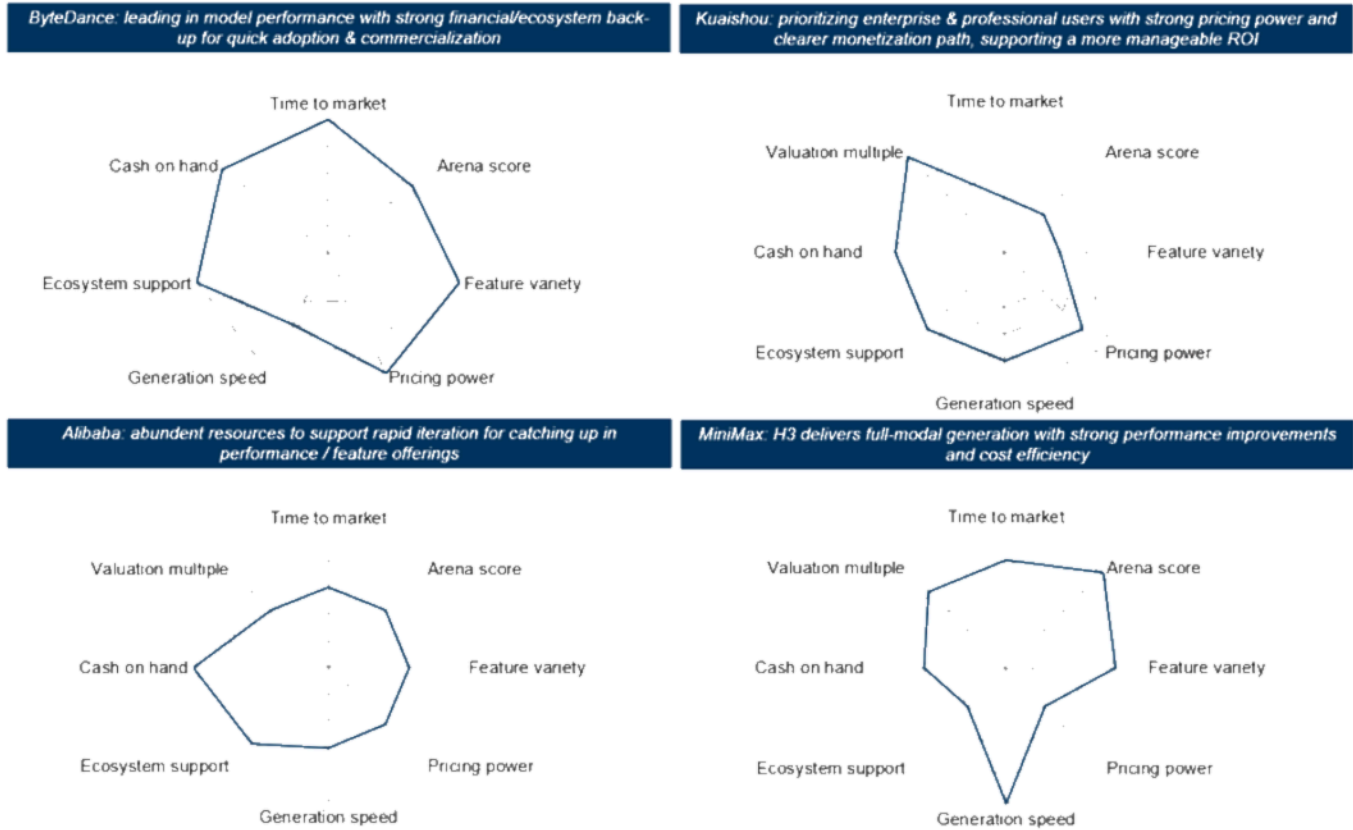
Source: Company data, 36kr, Data compiled by Goldman Sachs Global Investment Research

Exhibit 23: Competitive positioning framework for multi-modal generation players



Source: Goldman Sachs Global Investment Research

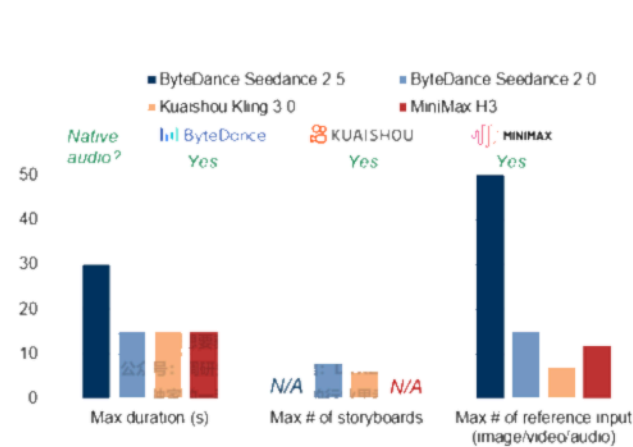
Exhibit 24: Overview of competitive analysis for key multi-modal labs' positioning



ByteDance's (Not Covered) relevant metrics are calculated by analyzing the industry and companies that we cover, and then extrapolating them to ByteDance. N/A for ByteDance's valuation multiple

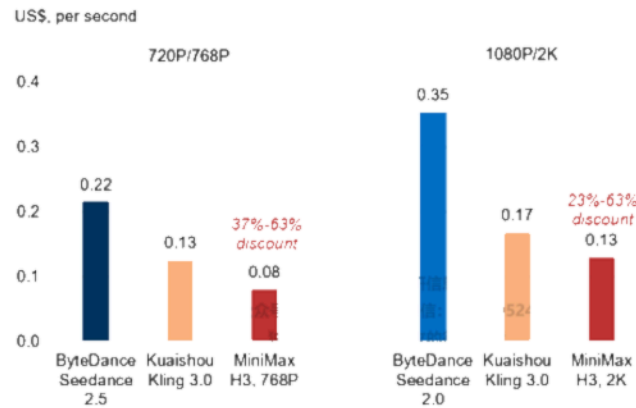
Source: Goldman Sachs Global Investment Research

Exhibit 25: Feature variety of key AI video models



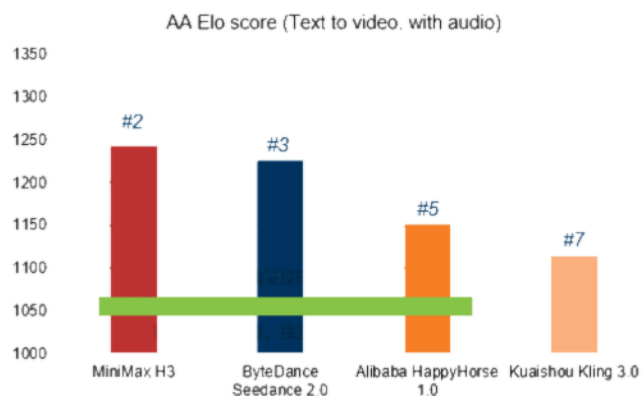
Source: Company data, Data compiled by Goldman Sachs Global Investment Research

Exhibit 26: Attractive H3 API pricing at 1/3~2/3 of peer level
US\$ per second



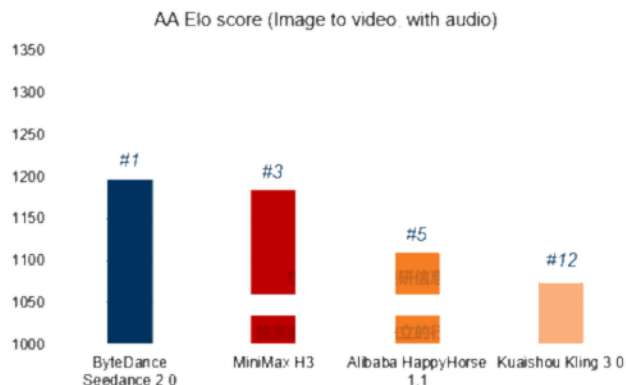
Source: Artificial Analysis, Data compiled by Goldman Sachs Global Investment Research

Exhibit 27: We note substantial improvement in MiniMax H3 ranking on AA Text to video leaderboard (with audio)



Source: Artificial Analysis, Data compiled by Goldman Sachs Global Investment Research

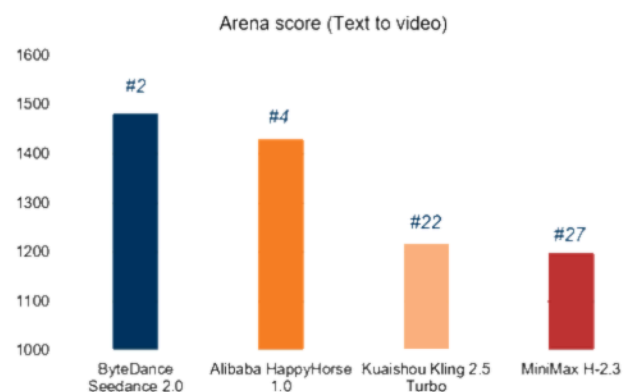
Exhibit 28: ...Same in Image to video leaderboard, despite lagging Seedance 2.0 at second place



Source: Artificial Analysis, Data compiled by Goldman Sachs Global Investment Research

Exhibit 29: Arena.ai score and global ranking of key video models (Text to video generation)

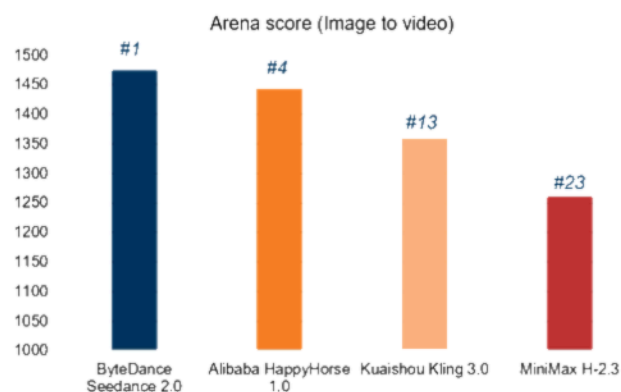
Date as of July, 2026



Source: LMArena, Data compiled by Goldman Sachs Global Investment Research

Exhibit 30: Arena.ai score and global ranking of key video models (Image to video generation)

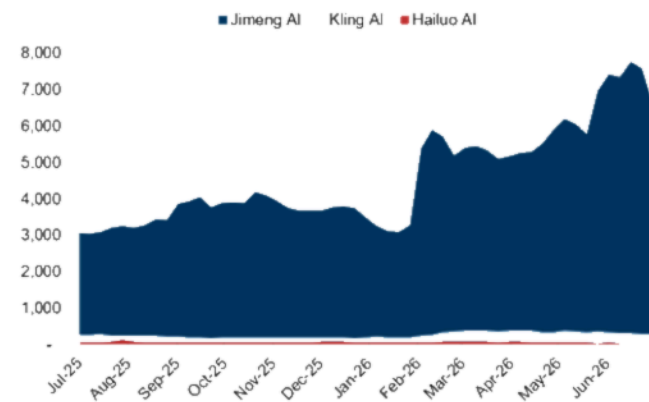
Date as of June, 2026



Source: LMArena, Data compiled by Goldman Sachs Global Investment Research

Exhibit 31: ByteDance's Jimeng AI continued to lead in WAU among consumer AI video apps

Weekly active users (k)



Source: QuestMobile, Data compiled by Goldman Sachs Global Investment Research

AI cloud/infrastructure & capex

Evolving dynamics of domestic & foreign chip supply

Exhibit 32: H100 spot rental pricing continues to rise YTD
US\$ per GPU-hour



Silicon Data H100 Rental Price Index (SDH100RT) from neo-cloud is a daily benchmark for on-demand NVIDIA H100 rental pricing

Source: Silicon Data, Bloomberg

Exhibit 33: ...same for A100 spot rental pricing
US\$ per GPU-hour



Silicon Data A100 Rental Price Index (SDA100RT) from neo-cloud is a daily benchmark for on-demand NVIDIA A100 rental pricing

Source: Silicon Data, Bloomberg

Exhibit 34: While B200 spot rental pricing shows higher volatility
US\$ per GPU-hour

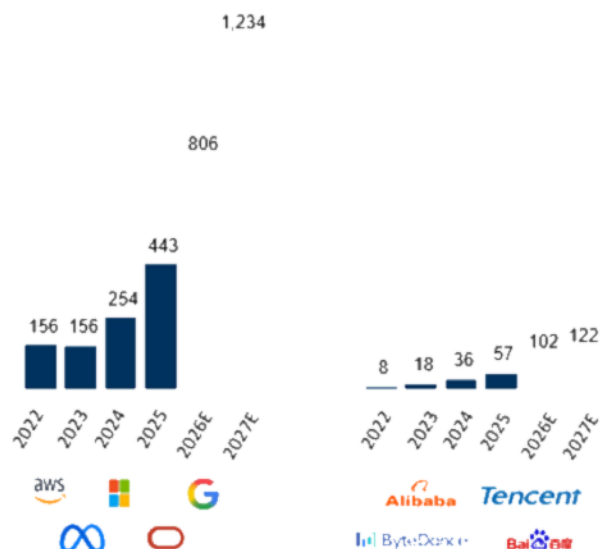


Silicon Data B200 Rental Price Index (SDB200RT) is a daily benchmark for on-demand NVIDIA B200 rental pricing

Source: Silicon Data, Bloomberg

AI capex/infrastructure: Our latest China hyperscaler capex forecasts and capex-to-cloud revenue framework

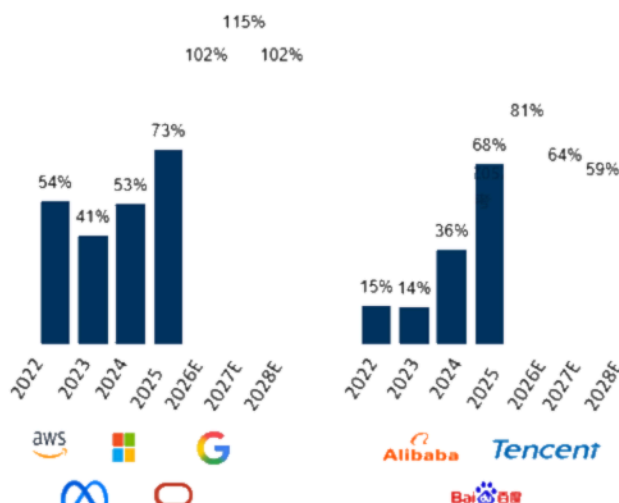
Exhibit 35: Historical and projected capex for major US & China cloud service providers
US\$bn



We calculate ByteDance's (Not Covered) relevant operating metrics by analyzing the industry and companies that we cover, and then extrapolating them to ByteDance.

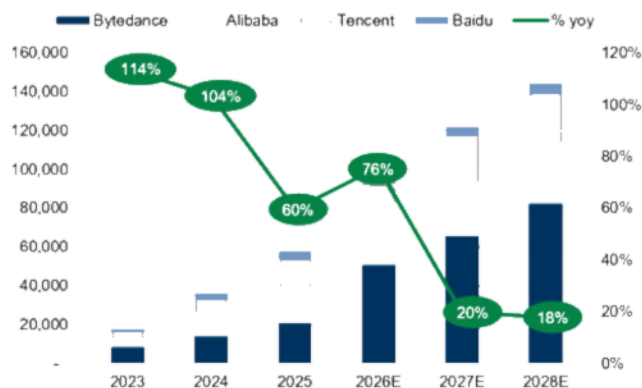
Source: Company data, Goldman Sachs Global Investment Research

Exhibit 36: Capex to operating cash flow ratio is still healthy for China hyperscalers



Source: Company data, Goldman Sachs Global Investment Research

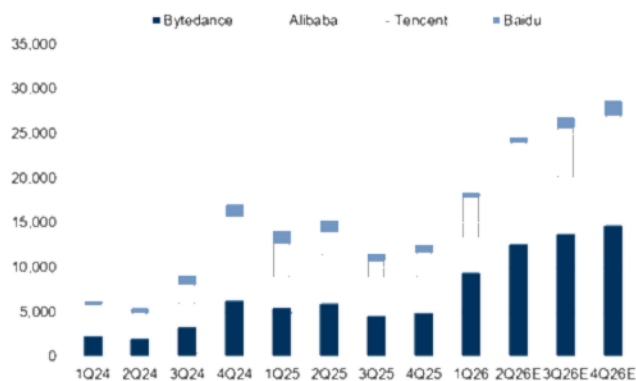
Exhibit 37: China Internet yearly capex forecast
US\$mn



For ByteDance (Not Covered), we impute capex based on industry trends.

Source: Company data, Goldman Sachs Global Investment Research, Bloomberg

Exhibit 38: China CSP's capex quarterly forecasts
US\$mn

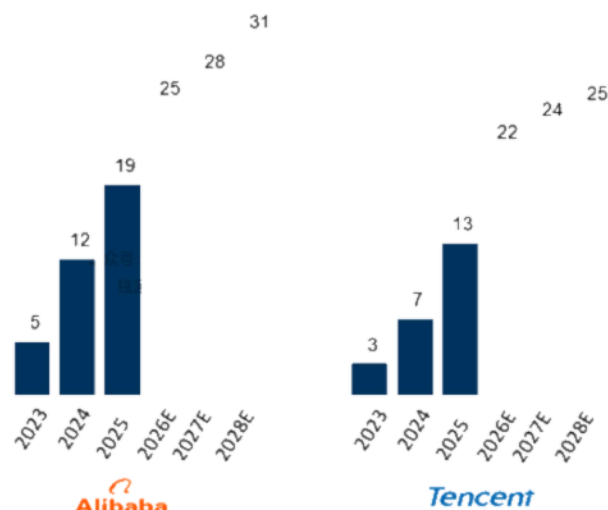


For ByteDance (Not Covered), we impute capex based on industry trends.

Source: Company data, Goldman Sachs Global Investment Research, Bloomberg

Exhibit 39: Historical and projected capex for Alibaba and Tencent

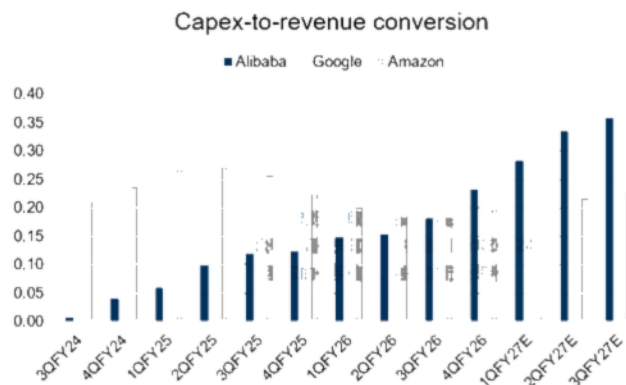
US\$bn



Source: Company data, Goldman Sachs Global Investment Research

Exhibit 40: Alibaba capex-to-revenue conversion ratio vs. Amazon and Google

Rolling 12m incremental cloud revenue/rolling 12m capex (lagged by 1 quarter)

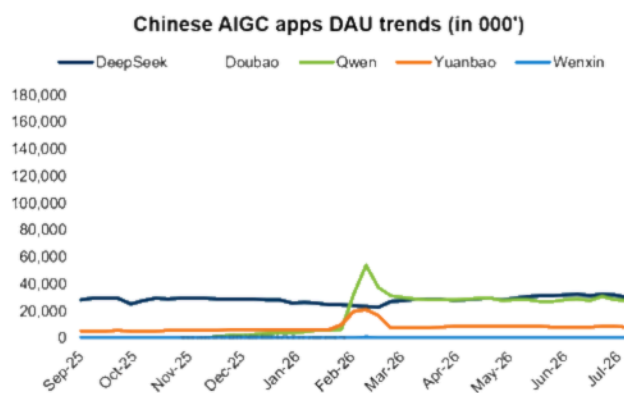


We benchmark Alibaba external cloud revenue growth to Google and Amazon Cloud (lagged c.2 years)

Source: Company data, Goldman Sachs Global Investment Research

AI applications: Harness/agent products and To-C apps**To-C application / chatbot engagement trends****Exhibit 41: Chinese AIGC apps DAU trends**

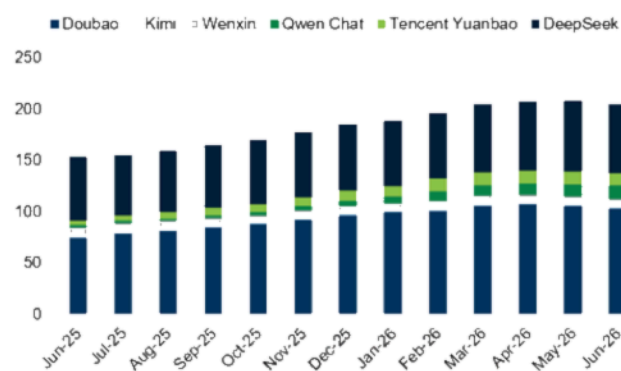
DAU K in China



Source: QuestMobile

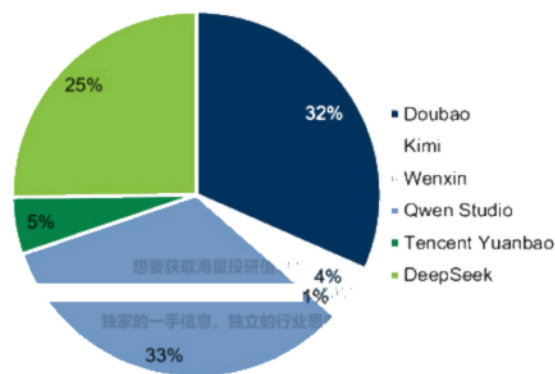
Exhibit 42: China AIGC to-C / chatbots MAU trends

MAU mn



Source: SensorTower

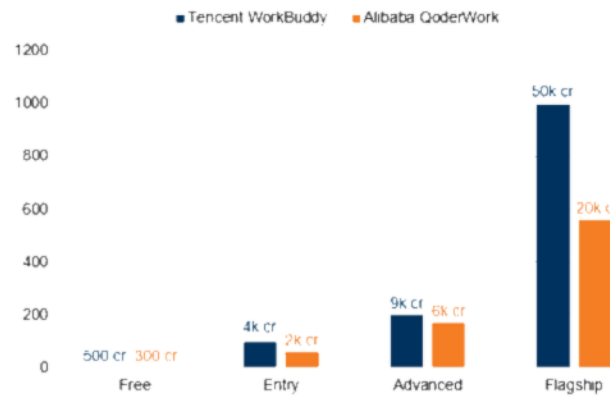
Exhibit 43: China AIGC to-C / chatbots monthly download share: Qwen surpassing Doubao in June



Source: SensorTower

Harness/agent application

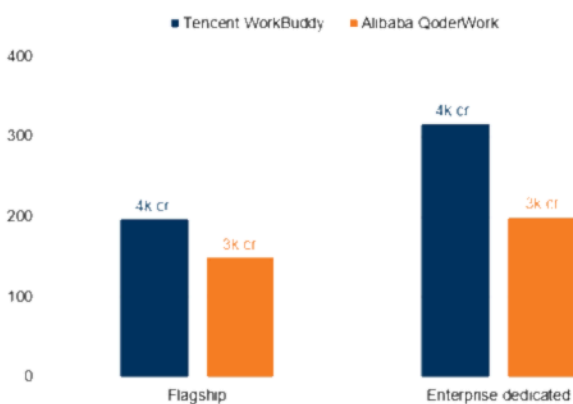
Exhibit 44: Personal subscription price for Tencent WorkBuddy and Alibaba QoderWork, under tiers of different credit offerings



We note that the monthly subscription tiers differ mainly in the amount of included credits, while credits across products are not directly comparable, since each vendor defines its own credit unit and consumption rate

Source: Company data, Data compiled by Goldman Sachs Global Investment Research

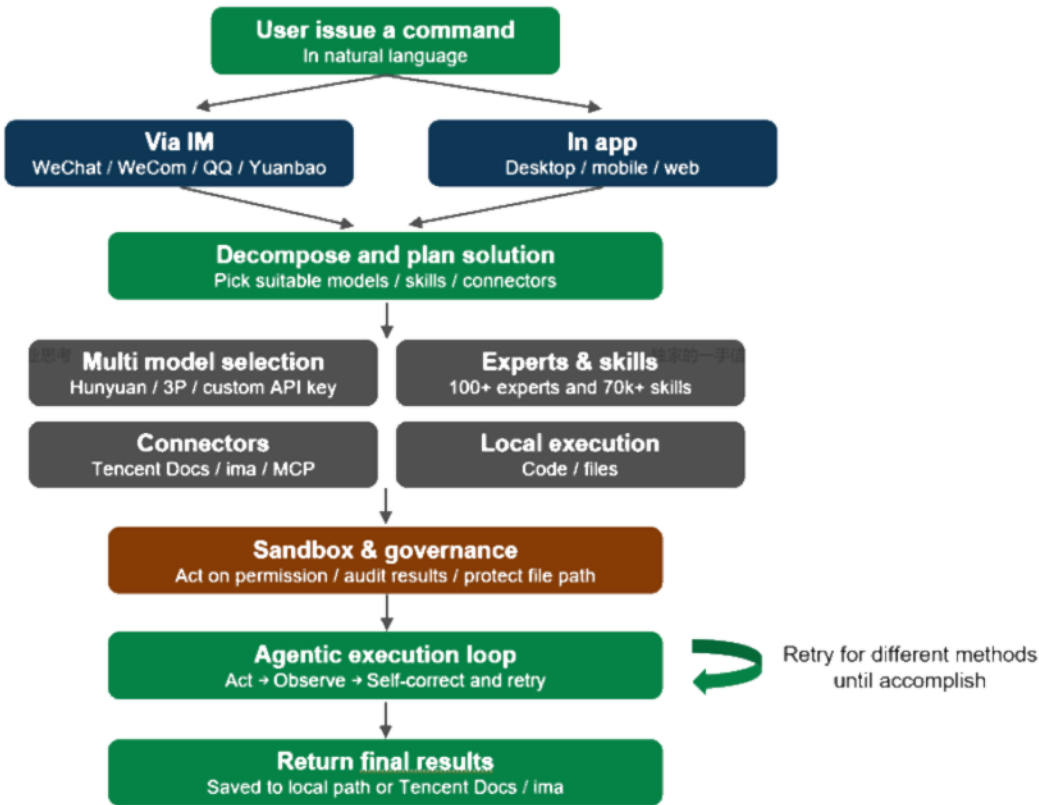
Exhibit 45: Enterprise subscription price for Tencent WorkBuddy and Alibaba QoderWork



Enterprise dedicated version gives customers its own isolated instance and data storage

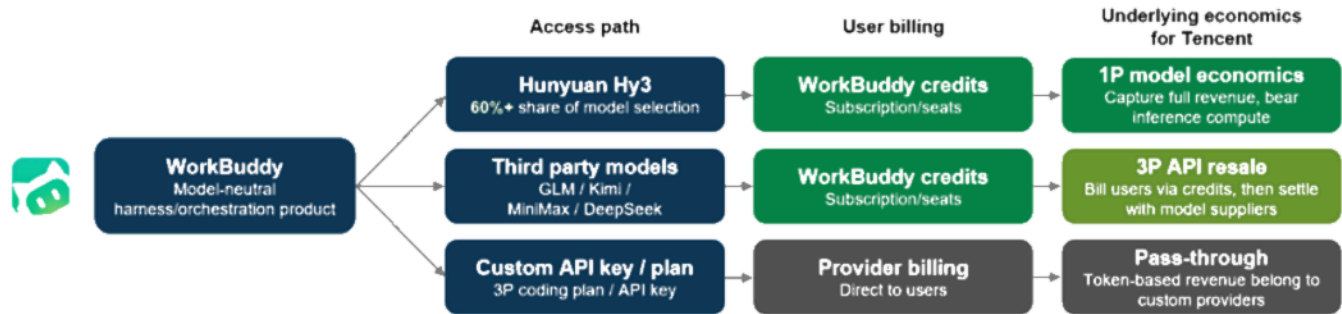
Source: Company data, Data compiled by Goldman Sachs Global Investment Research

Exhibit 46: How WorkBuddy works from receiving commands to final deliverable



Source: Company data, Data compiled by Goldman Sachs Global Investment Research

Exhibit 47: WorkBuddy business model and underlying economics for Tencent



Note: WorkBuddy monetizes through consumer subscription + enterprise seats + credit add-on packs; Built-in model usage is measured by credits (similar to token-based pricing), with more capable models carrying higher credit multipliers

Source: Company data, Data compiled by Goldman Sachs Global Investment Research

Estimate changes for Z.AI and MiniMax

Z.AI/Zhipu

We change our 2026-28E revenues by +35%/+4%/-2% and our earnings estimates by -3%/-3%/-1% on strong 2026 YTD ARR ramp up, while factoring in intensified competition in longer term over larger parameter/high-end coding models and sustained R&D investments.

Our 12-m TP moves to HK\$1,610 (prior HK\$1,880, implied US\$95bn valuation) is based on a DCF valuation, with 12% WACC and 2% terminal growth rate (consistent with our coverage) that assumes continued market share gain to 21% market share by 2030E (prior: 24%), and a long-term adj. EBIT margin of 26% by 2035E. Accordingly, we also refresh our bull case implied valuation to HK\$3,266 (vs. prior HK\$3,350) and bear case implied valuation to HK\$835 (vs. prior HK\$900) with key assumptions in [Exhibit 50](#).

Key risks: (+) Stronger-than-expected model intelligence, **(+)** Faster-than-expected path to profit visibility, **(+)** Additional non-inference take rate revenue streams, **(+)** Stronger-than-expected commercialization capability, **(-)** Competition in global foundation model industry, **(-)** Limited near-term profit visibility due to high R&D expenses, **(-)** Cash burn/self-funding capability, **(-)** Risk around geopolitics under an intensified tech race between the US and China.

Exhibit 48: Z.AI: New vs. old estimates

December year-end (Rmb mn)	New 2026E	Old 2026E	Diff (%)	New 2027E	Old 2027E	Diff (%)	New 2028E	Old 2028E	Diff (%)
Open platform and API yoy %	7,473	5,248	42%	22,790	21,931	4%	55,188	56,144	-2%
On-premise deployment yoy %	1,103	1,091	1%	1,836	1,794	2%	2,712	2,666	2%
Revenue yoy %	8,575	6,339	35%	24,626	23,725	4%	57,900	58,809	-2%
Cost of sales	(6,246)	(4,528)	38%	(17,641)	(17,264)	2%	(41,218)	(42,438)	-3%
Gross (loss)/profit yoy %	2,329	1,812	29%	6,985	6,460	8%	16,682	16,371	2%
	685%	511%		200%	257%		139%	153%	
Selling and distribution expenses	(586)	(508)	-15%	(774)	(737)	-5%	(1,107)	(1,105)	0%
Administrative expenses	(733)	(657)	-12%	(975)	(953)	-2%	(1,433)	(1,429)	0%
Research and development expenses	(6,234)	(5,725)	-9%	(11,532)	(10,877)	-6%	(16,630)	(16,305)	-2%
Operating profit	(5,190)	(5,045)	-3%	(5,855)	(5,673)	-3%	(2,283)	(2,269)	-1%
Adj. operating profit - non-IFRS yoy %	(4,492)	(4,375)	-3%	(5,017)	(4,869)	-3%	(1,320)	(1,304)	-1%
	-39%	-36%		-12%	-11%		74%	73%	
(Loss)/Profit before tax	(5,229)	(5,081)	-3%	(5,893)	(5,709)	-3%	(2,322)	(2,304)	-1%
Income tax expense	-	-		-	-		-	-	
Net profit/(loss)	(5,229)	(5,081)	-3%	(5,893)	(5,709)	-3%	(2,322)	(2,304)	-1%
Adj. net profit/(loss) - non-IFRS yoy %	(4,531)	(4,411)	-3%	(5,056)	(4,905)	-3%	(1,359)	(1,340)	-1%
	-42%	-39%		-12%	-11%		73%	73%	
GPM	27%	29%	-141bps	28%	27%	113bps	29%	28%	97bps
Adj. OPM	-52%	-69%	1663bps	-20%	-21%	15bps	-2%	-2%	-6bps
Adj. NPM	-53%	-70%	1675bps	-21%	-21%	14bps	-2%	-2%	-7bps

Source: Goldman Sachs Global Investment Research

Exhibit 49: Z.AI - Income statement

December year-end (Rmb mn)	2023	2024	2025	2026E	2027E	2028E
On-premise deployment	113	264	534	1,103	1,836	2,712
yoy %		134%	102%	107%	66%	48%
Cloud-based deployment	12	48	190	7,473	22,790	55,188
yoy %		307%	293%	3825%	205%	142%
Revenue	125	312	724	8,575	24,626	57,900
yoy %		151%	132%	1084%	187%	135%
Cost of sales	(44)	(137)	(428)	(6,246)	(17,641)	(41,218)
Gross (loss)/profit	80	176	297	2,329	6,985	16,682
yoy %		-119%	69%	685%	200%	139%
GP Margin %	65%	56%	41%	27%	28%	29%
Other income			2	10	19	15
Selling and distribution expenses	(101)	(387)	(391)	(586)	(774)	(1,107)
Administrative expenses	(66)	(134)	(505)	(733)	(975)	(1,433)
Research and development expenses	(529)	(2,195)	(3,180)	(6,234)	(11,532)	(16,630)
Operating profit	(626)	(2,538)	(3,787)	(5,190)	(5,855)	(2,283)
Adj. operating profit - non-IFRS	(620)	(2,515)	(3,228)	(4,492)	(5,017)	(1,320)
yoy %	-535%	-305%	-28%	-39%	-12%	74%
OP Margin %	-498%	-805%	-446%	-52%	-20%	-49%
(Loss)/Profit before tax	(788)	(2,958)	(4,718)	(5,229)	(5,893)	(2,322)
Income tax expense	-	-	-	-	-	-
Net profit/(loss)	(788)	(2,958)	(4,718)	(5,229)	(5,893)	(2,322)
Adj. net profit/(loss) - non-IFRS	(621)	(2,466)	(3,182)	(4,531)	(5,056)	(1,359)
yoy %	-537%	-297%	-29%	-42%	-12%	73%
NP Margin %	-499%	-789%	-439%	-53%	-21%	-2%

Source: Company data, Goldman Sachs Global Investment Research

Exhibit 50: Z.AI: Our valuation bull/bear/base scenarios

Scenarios for valuation	Bear case 2027E EV/Sales-based	Base case DCF-based	Bull case DCF-based, mid-2028E ARR-implied
Z.AI	Applying 6X EV/Sales multiple to 27E open API platform revenue. Bear case valuation implies Zhipu's revenue pool share declines from 18% to 15% from 2027 to 2030E on higher competition.	Assuming Zhipu's revenue pool share to expand from 18% in 2027 to 21% by 2030E, and 26% EBIT margin (by 2035E). Applying 12% WACC and 2% terminal growth rate (consistent with our coverage).	Bull case valuation assumes Zhipu's revenue pool share to expand from 18% in 2027 to 32% by 2030E. Implies 28X mid-2028E ARR, discounting back using 12% p.a.
Valuation per share (HK\$)	835	1610	3266
Valuation (US\$bn)	50	95	195

Source: Company data, Goldman Sachs Global Investment Research

MiniMax

We raise our 2026-28E revenue by 22-64% on favorable mix shift towards fast growing API business from AI-native products, and fine-tune our net income estimates by -3% to +1% to reflect higher R&D and other operating expenses.

Our 12-m TP moves to HK\$800 (prior HK\$860, implied US\$38bn valuation) largely due to more share counts following the completion of recent financing round, partly offset by higher API revenue forecast. We also refresh our bull case implied valuation to HK\$1392 (vs. prior HK\$1350) and bear case implied valuation to HK\$220 (vs. prior HK\$250) with key assumptions in [Exhibit 53](#).

Key risks: Weaker-than-expected model performance amid the competition in the global foundation model industry; Slower-than-expected path to profit visibility;

Weaker-than-expected commercialization capability; risks around IP/content generation; cash burn/self-funding capability; risks around geopolitics under an intensified tech race between the US and China.

Exhibit 51: MiniMax: New vs. old estimates

	New	Old	Diff (%)	New	Old	Diff (%)	New	Old	Diff (%)
December year-end (US\$ mn)	2026E	2026E		2027E	2027E		2028E	2028E	
AI-native products	171	149	14%	429	558	-23%	869	1,622	-46%
yoy %	222%	182%		151%	273%		102%	191%	
Open API platform and other AI-based enterprise services	322	151	114%	842	322	162%	2,136	848	152%
yoy %	1140%	480%		161%	114%		154%	163%	
Revenue	493	300	64%	1,271	880	44%	3,005	2,470	22%
yoy %	524%	280%		158%	193%		136%	181%	
Cost of sales	(344)	(222)	55%	(928)	(666)	39%	(1,964)	(1,625)	21%
Gross (loss)/profit	149	78	91%	343	214	60%	1,041	845	23%
yoy %	644%	289%		129%	174%		204%	295%	
Selling and distribution expenses	(88)	(75)	-17%	(203)	(180)	-13%	(391)	(400)	2%
Administrative expenses	(66)	(60)	-11%	(96)	(80)	-21%	(180)	(130)	-38%
Research and development expenses	(531)	(468)	-14%	(674)	(549)	-23%	(947)	(792)	-20%
Operating profit	(536)	(524)	-2%	(630)	(595)	-6%	(477)	(477)	0%
Adj. operating profit - non-IFRS	(467)	(465)	-3%	(558)	(523)	-7%	(401)	(401)	0%
yoy %	-57%	-53%		-19%	-15%		28%	23%	
(Loss)/Profit before tax	(505)	(493)	-2%	(559)	(546)	-2%	(423)	(428)	1%
Income tax expense	-	-		-	-		-	-	
Net profit/(loss)	(505)	(493)	-2%	(559)	(546)	-2%	(423)	(428)	1%
Adj. net profit/(loss) - non-IFRS	(436)	(425)	-3%	(487)	(473)	-3%	(347)	(352)	1%
yoy %	-74%	-69%		-12%	-12%		29%	26%	
GPM	30%	26%	424bps	27%	24%	267bps	35%	34%	44bps
Adj. OPM	-95%	-152%	5707bps	-44%	-59%	1550bps	-13%	-16%	291bps
Adj. NPM	-88%	-141%	5308bps	-38%	-54%	1548bps	-12%	-14%	270bps

Source: Goldman Sachs Global Investment Research

Exhibit 52: MiniMax: P&L summary

December year-end (US\$ mn)	2023	2024	2025	2026E	2027E	2028E
AI-native products	1	22	53	171	429	869
yoy %		2777%	143%	222%	151%	102%
Open API platform	3	9	26	322	842	2,136
yoy %		223%	198%	1140%	161%	154%
Revenue	3	31	79	493	1,271	3,005
yoy %		782%	159%	524%	158%	136%
Cost of sales	(4)	(27)	(59)	(344)	(928)	(1,964)
Amongst which: Inference cost	(4)	(26)	(55)	(264)	(596)	(1,105)
Gross (loss)/profit	(1)	4	20	149	343	1,041
yoy % Buy		538%	437%	644%	129%	204%
GP Margin %	-25%	12%	25%	30%	27%	35%
Selling and distribution expenses	(23)	(87)	(52)	(88)	(203)	(391)
Administrative expenses	(8)	(14)	(37)	(66)	(96)	(180)
Research and development expenses	(70)	(189)	(253)	(531)	(674)	(947)
Amongst which: Training cost	(47)	(141)	(180)	(360)	(400)	(545)
Operating profit	(101)	(287)	(321)	(536)	(630)	(477)
Adj. operating profit - non-IFRS	(98)	(280)	(297)	(467)	(558)	(401)
yoy %	-637%	-186%	-6%	-57%	-19%	28%
OP Margin %	-2831%	-917%	-376%	-95%	-44%	-46%
(Loss)/Profit before tax	(269)	(465)	(1,872)	(505)	(559)	(423)
Income tax expense	-	-	-	-	-	-
Net profit/(loss)	(269)	(465)	(1,872)	(505)	(559)	(423)
Adj. net profit/(loss) - non-IFRS	(89)	(244)	(251)	(436)	(487)	(347)
yoy %	-633%	-174%	-3%	-74%	-12%	29%
NP Margin %	-2574%	-800%	-317%	-88%	-38%	-12%
GPM	-25%	12%	25%	30%	27%	35%
Adj. OPM	-2831%	-917%	-376%	-95%	-44%	-46%
Adj. NPM	-2574%	-800%	-317%	-88%	-38%	-12%

Source: Company data, Goldman Sachs Global Investment Research

Exhibit 53: MiniMax: Our bear, base and bull case valuations and assumptions

Scenarios for valuation	Bear case 2027E EV/Sales-based	Base case DCF-based	Bull case DCF-based, mid-2028E ARR-implied
MiniMax	Applying 6X EV/Sales multiple to 27E revenue of US\$1.3bn. Bear case valuation implies MiniMax's revenue pool share to gradually decline over 2027E-2030E	Assuming MiniMax's revenue pool share to expand 0.4-0.9 ppts p.a. from 2027 to 2030E, reach market share of 8% in 2030E, and 18% EBIT margin (by 2035E). Applying 12% WACC and 2% terminal growth rate (consistent with our coverage).	Bull case valuation assumes MiniMax's revenue pool share to expand 6ppts p.a. from 2027 to 2030E, reaching a total market share of 24% by 2030E. Implies 25X mid-2028E ARR, discounting back using 12% p.a.
Valuation per share (HK\$)	220	800	1392
Valuation (US\$bn)	10	38	66

Source: Goldman Sachs Global Investment Research

Xunce Technology (NC) NDR highlights

We hosted Shenzhen Xunce Technology (3317.HK, Not Covered)'s Investor Relations team on July 23. Key takeaways include:

1) Strategic positioning as the data layer within AI value chain: The company focuses on transforming enterprise side private data, workflows and industry know-how into standardized data that can be directly fed into foundation models for efficient processing. Unlike foundation model providers (whose training data is largely derived from public internet), Xunce produces vertical tokens designed to solve enterprise-specific use cases. Relative to traditional ERP vendors, the company aims to unify fragmented data across departments and systems. Meanwhile, the company has recently partnered with domestic GPU vendors to help customers build and optimize internal AI infrastructure, with the shared goal of improving enterprises' ability to organize proprietary data while enhancing the efficiency of AI-driven analysis and decision-making.

2) Growing momentum of by-token business model: Management highlights growing adoption of token-based pricing since 1Q, with token revenue accounting for >10% of 1H26 revenue and expected to reach 20-30% for FY26. The company attributes the premium pricing (at US\$10-100 per 1M tokens) and ARPU of Rmb8mn (vs. FY25 company-level ARPU of Rmb5.5mn) to the higher effectiveness of vertical tokens, which compared to raw data inputs into foundation models, help reduce token consumption and hallucination. Management also notes the plug-and-play nature of its product, which allows enterprises to directly evaluate performance improvements through A/B testing, making ROI more measurable and easier to justify.

3) Continued expansion into new verticals: The company has now entered 11 industries and continues to expand its industry coverage (auto and low-altitude economy in FY26), leveraging proprietary data processing capabilities and reusable business modules. Management notes that entering a new industry typically requires 1-3 years of upfront investment in data connectivity, industry knowledge accumulation and workflow construction, with gross margins initially at 40-50% before scaling toward 70-80% as industry-specific modules mature and become reusable across customers.

4) Capital allocation plans: The company plans to accelerate investment across three priorities: further development of Token OS infrastructure and technologies, accelerating vertical industry expansion and overseas market penetration. Management views its current cash position as sufficient to support near-term operations before turning profitable.

5) Competitive landscape and margin outlook: Management views enterprise data governance as a differentiated layer within the AI stack which requires long-term capital, talent and engineering investments, and believes foundation model players are unlikely to penetrate with model performance improvements and commercialization remaining key priorities. As higher-margin token revenue continues to scale (with 90-95% gross margin expected at mature stage), the company expects a structurally improving gross margin trajectory.

Top apps: June time spent trends at a glance

Overall, the top 400 mobile apps have grown total time spent by +9% yoy in June 2026 (vs. +5% in June 2025), with **Weixin** time spent growth (+2% yoy) and continued strong **Douyin** app engagement performance (main app +27% yoy/Lite +30% yoy). We also note ByteDance's continued share gain by time spent (33% within top apps, up +6% from 2 years ago) via short-form video apps (Douyin/Red Fruit) and AIGC applications (Doubao). Specifically by vertical:

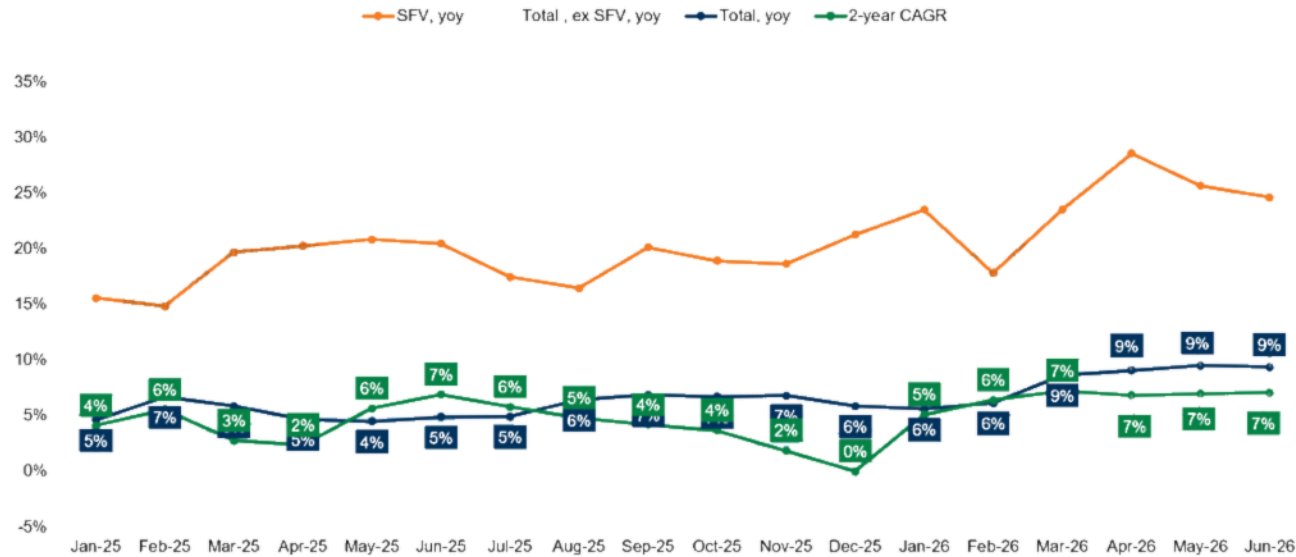
- **(+) AI engagement**, where domestic AIGC application engagement increased by +2% mom driven by Qwen (+16% mom) and Zhipu Qingyan (+11%). To-C AI chatbots continue to be mostly free-to-use, where the path to monetization is still evolving for China. Meanwhile, harness/agent applications have sustained strong traffic growth momentum, notably Tencent's WorkBuddy (WAU achieved 20%+ WoW growth for 6 consecutive weeks).
- **(+/-) eCommerce engagement**: Amongst top platforms, JD timespent decreased by -7% yoy on high base, Taobao timespent remained flat, while Pinduoduo grew its timespent by +6%.
 - **On Temu**, global MAU decreased by -3% mom at 474mn in Jun as per SensorTower, likely reflecting marketing budget reduction in regions including 1) ASEAN (-16% mom in June) and Europe (-4~-5% mom in June) due to official removal of de minimis exemption; and 2) US (decreased by -6% mom, accounting for 16% of global MAU in Jun) before indefinite suspension of de minimis exemption and stricter clearance requirements newly announced in July. That said, we view commencement of PDD's first-party brand ("Xin Pin Mu") investment cycle for the next three years (targeting to rebuild Pinduoduo) as a multi-year growth driver for Temu + Pinduoduo.
- **(+/-) Social engagement**: Weixin time spent increasing by +2% yoy and Weibo/QQ's timespent decreased by -3%/-4% yoy. **(+/-) Video engagement**: Douyin engagement sustained strong momentum (main app +27% yoy/Lite +30% yoy) yet offset by a continuous softening in user engagement at Kuaishou (main app -14% yoy/Express -6% yoy). Bilibili grew healthy at +7% yoy, while long-form video apps continued to decelerate with iQiyi at -11% yoy.
- **(+) Games engagement**: HoK (Honour of Kings) time spent growth accelerated to +19% yoy in Jun; Delta Force engagement maintained robust momentum at +94% yoy; Game for Peace timespent decreased by -5% yoy in Jun (vs. +1% yoy in May).
- **(+) Local services engagement**: DiDi timespent growth reaccelerated to +24% yoy (vs. +10% in May), while Meituan engagement decreased -8% yoy (vs. -6% yoy in May) and Amap's engagement decreased -2% yoy (vs. -3% in May). We also note continuous disruption from Doushengsheng (抖省省), ByteDance's local services discount/coupon app, to Dianping (timespent -5% yoy in June, negative growth since March), where Doushengsheng DAU has exceeded 20mn in July MTD (~60% of Dianping), see [Exhibit 12](#). However, deceleration of Doushengsheng's MoM DAU growth (see [Exhibit 13](#)) may suggest potential normalization into 2H, especially given its aggressive subsidy level.

■ **Other internet verticals:**

- **(+) Education:** Despite a seasonally weak month, DAUs of **Doubao Study** sustained strong growth of +161% yoy to 3.4mn in Jun 2026. In overseas markets, Gauth maintained its global leadership with monthly billings +8% yoy to c.US\$0.7mn. **TAL Peiyou** sustained positive MAU growth of +13% yoy in Jun vs. -7% yoy for **Xueersi.com** and -6% yoy for **EDU**.
- **(+) Recruitment:** Boss Zhipin showed **+9% yoy** MAU growth and **+7% yoy** growth in time spent in Jun 2026 (vs. +10%/+14% yoy growth in MAU/timespent in May 2026), faster than +7%/flat yoy growth of MAU/time spent for major online recruitment platforms in Jun 2026. Boss Zhipin maintained its leadership in time spent share at **67%** in Jun 2026 (vs. 66% in May 2026) in Mainland China. In Hong Kong market, **OfferToday** (BossZhipin's HK App) solidified its **No.1** HK recruitment platform position in terms of app DAU, MAU and download volume, reaching c.17k DAUs with +71% yoy growth in Jun 2026 (vs. +188% yoy growth in May 2026) and c.287k MAUs with +145% yoy growth in Jun 2026 (vs. +202% yoy growth in May 2026).
- **(-) Real estate: Time spent** on Beike's three major apps dropped -7% yoy/-6% mom in Jun 2026 (vs. +1% yoy/-1% mom in May 2026). **For DAU**, Beike Zhaofang DAU saw -11% yoy/-2% mom decline in Jun 2026 (vs. -6% yoy/-2% mom in May 2026).
- **(+) Auto aftermarket service: Tuhu app MAU** increased +11% yoy/+2% mom to 16.9mn in Jun 2026 (vs. +8% yoy/+1% mom in May 2026). On the merchant side, **Tuhu merchant app MAU** increased +3% yoy or dropped -2% mom to 85.1k in Jun 2026 (vs. +3% yoy/-1% mom in May 2026), with a DAU/MAU ratio of 77% and +1% yoy growth in daily time spent per DAU.

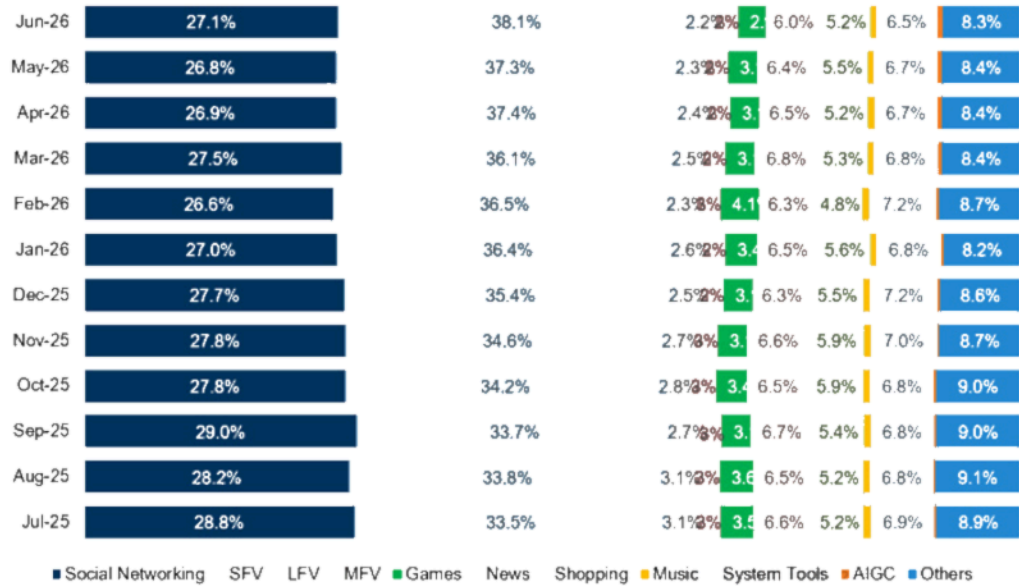
China Internet industry

Exhibit 54: Aggregate time spent growth continued at +9% yoy in Jun



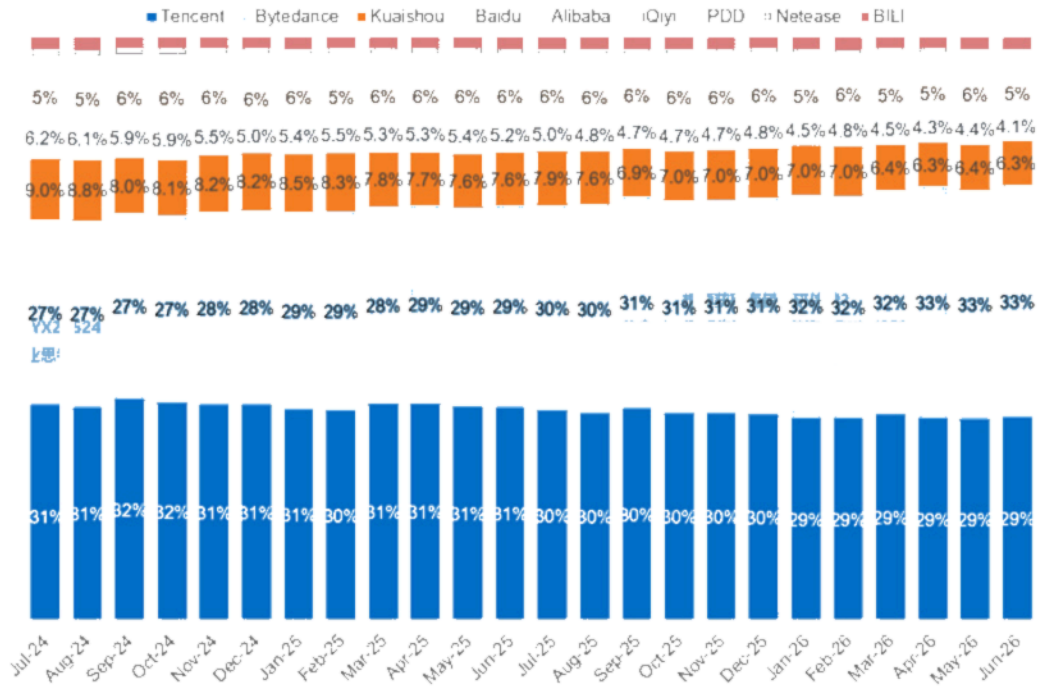
Source: QuestMobile

Exhibit 55: China Internet: time spent share % within top 400 apps by vertical



Source: QuestMobile

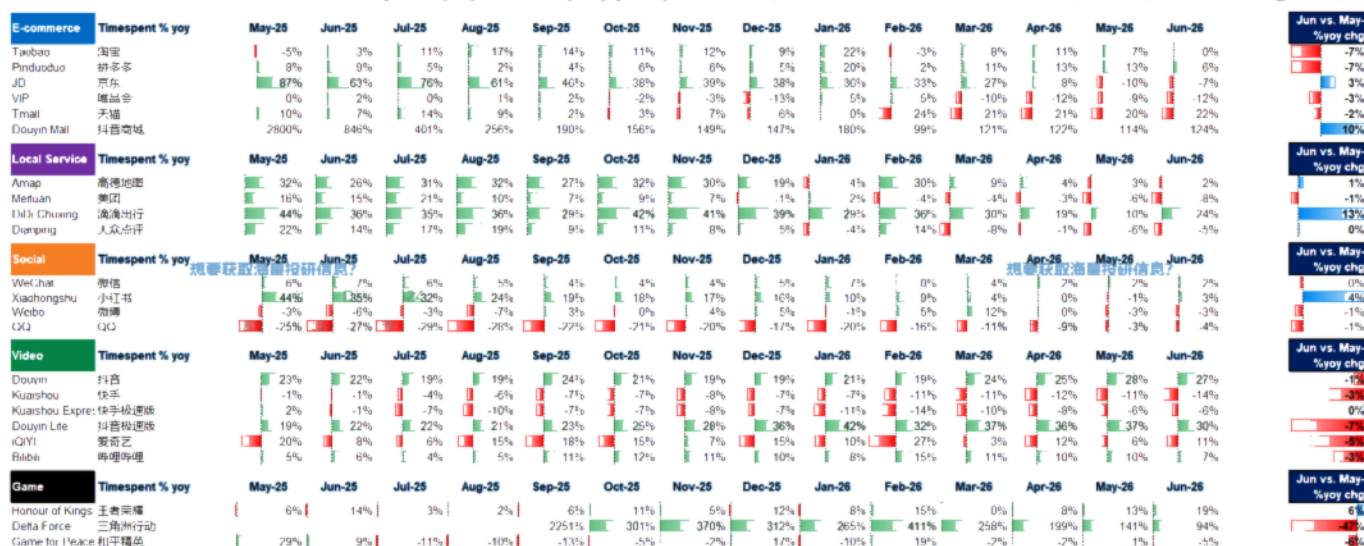
Exhibit 56: China Internet: time spent share % within top 400 apps by company



Source: QuestMobile

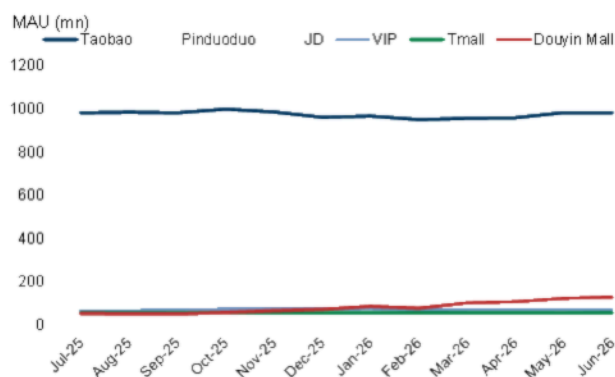
China online transaction platforms and entertainment

Exhibit 57: China Internet: time-spent yoy % of top apps by vertical (E-commerce, local service, social, video and game)



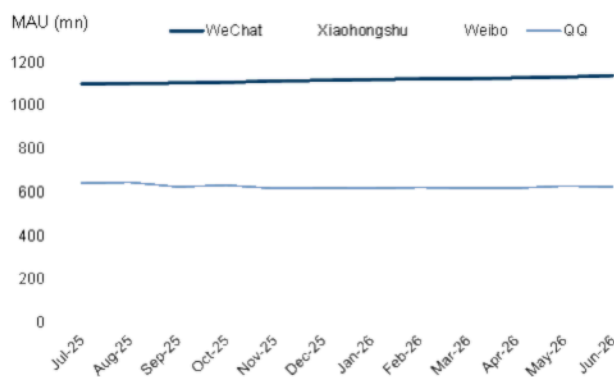
Source: QuestMobile

Exhibit 58: China e-commerce: major app MAU trend (mn)



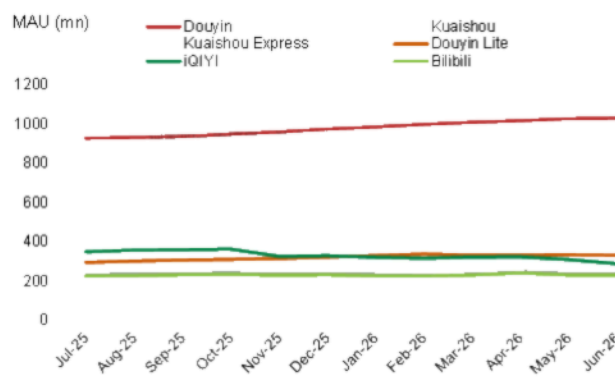
Source: QuestMobile

Exhibit 59: China online social: major app MAU trend (mn)



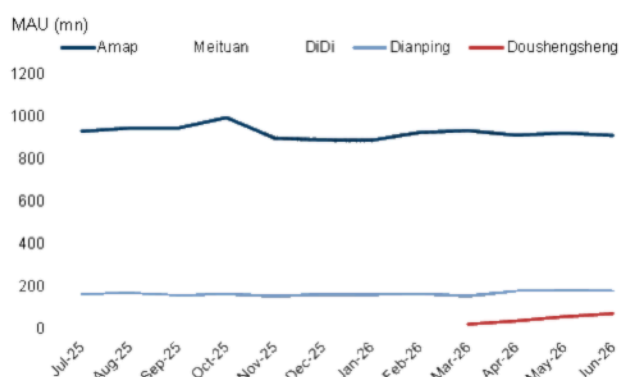
Source: QuestMobile

Exhibit 60: China online video: major app MAU trend (mn), where Douyin + Douyin Lite sustain strong growth momentum



Source: QuestMobile

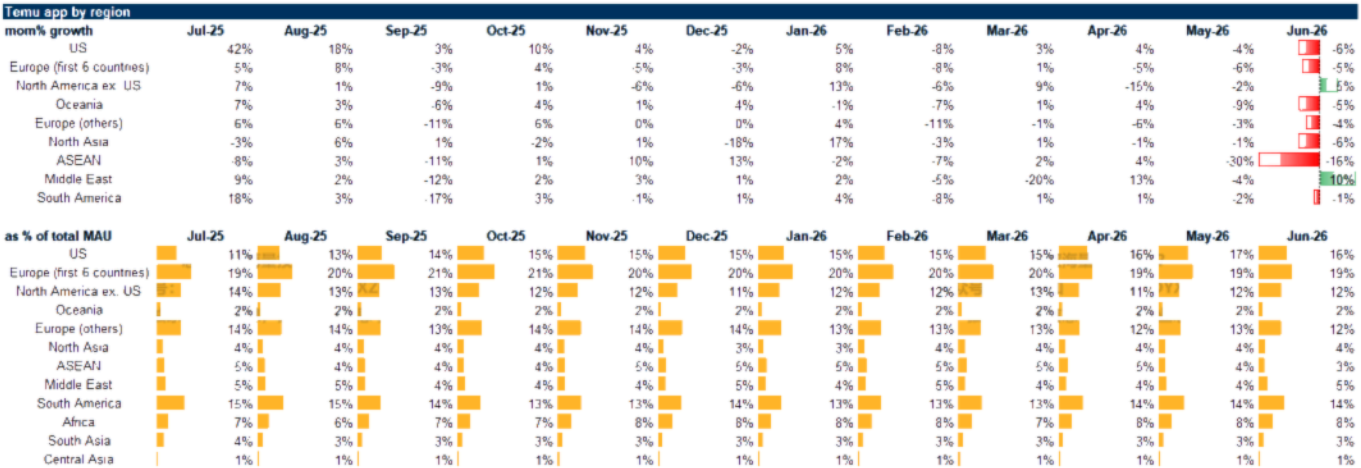
Exhibit 61: China online local services: major app MAU trend (mn) with Doushengsheng entering in March



Source: QuestMobile

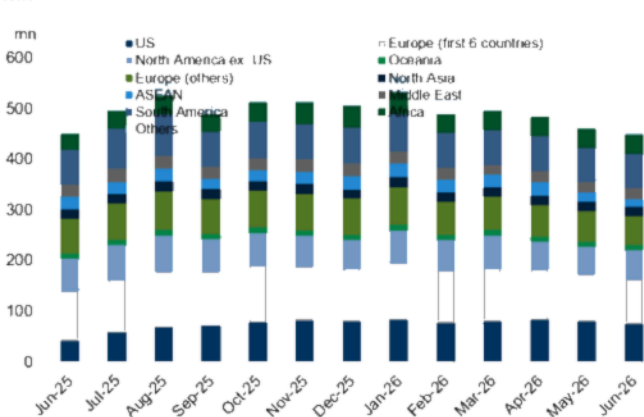
Temu operating data

Exhibit 62: Temu app MAU by region



Source: SensorTower

Exhibit 63: Temu global MAU reached 474mn in Jun 2026 (-3% mom)

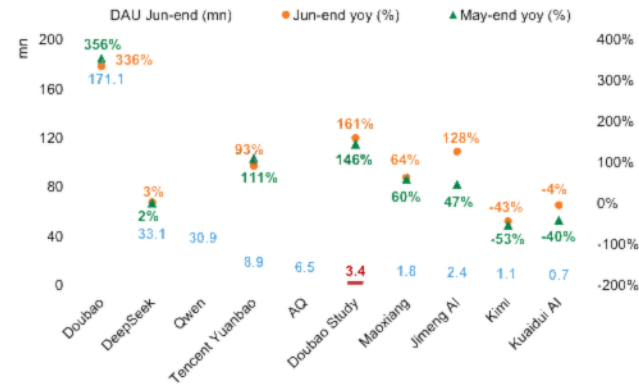


The first 6 European countries include France, Germany, Italy, the Netherlands, Spain, and the UK (launched in April 2023)

Source: SensorTower

Other internet verticals: education, recruitment, real estate, auto
Education

Exhibit 64: Doubao Study DAU growth accelerated to +161% yoy as of Jun-end and sustained the position of the 6th largest AI native app in China in terms of DAUs
DAUs of major AI apps as of Jun-end 2026

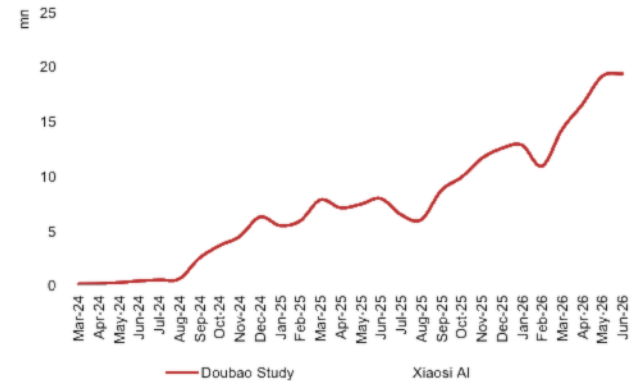


Qwen: 78x/94x for May-end/Jun-end yoy growth

Source: Questmobile, Data compiled by Goldman Sachs Global Investment Research

Exhibit 66: MAU of TAL's Xiaosi AI increased +50% yoy to c.0.6mn vs. MAU of Doubao Study increased +142% yoy to c.19.4mn in Jun 2026

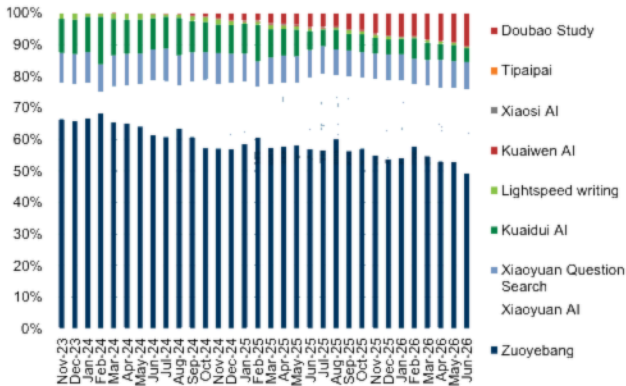
MAUs of Doubao Study vs. Xiaosi AI



Source: Questmobile, Data compiled by Goldman Sachs Global Investment Research

Exhibit 65: Doubao Study continued to expand time spent market share among major K-12 learning tool/study companion apps in Jun

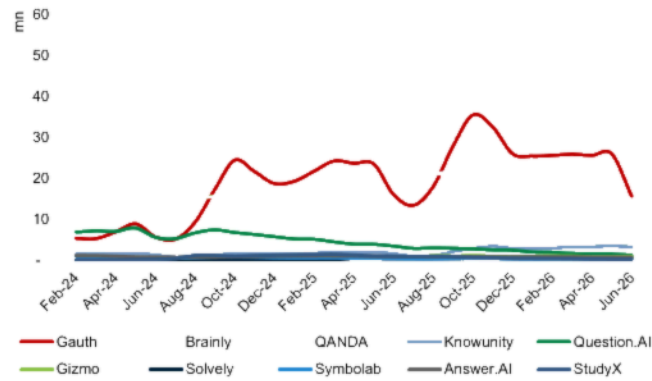
Time spent share among major K-12 learning tool/study companion apps



Source: Questmobile, Data compiled by Goldman Sachs Global Investment Research

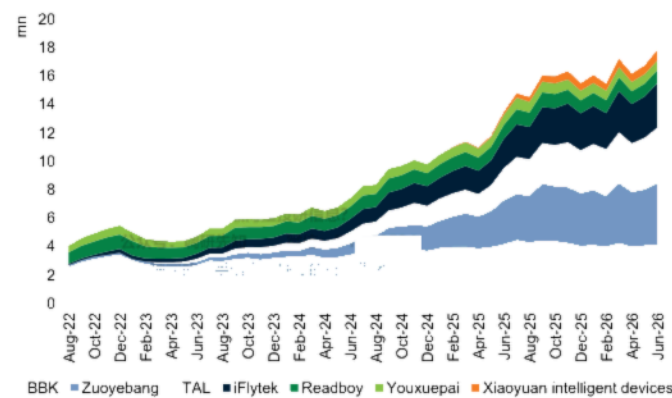
Exhibit 67: Gauth has outperformed global AI-native education peers in terms of MAUs

MAUs by AI education apps (mn)



Source: Sensor Tower, Data compiled by Goldman Sachs Global Investment Research

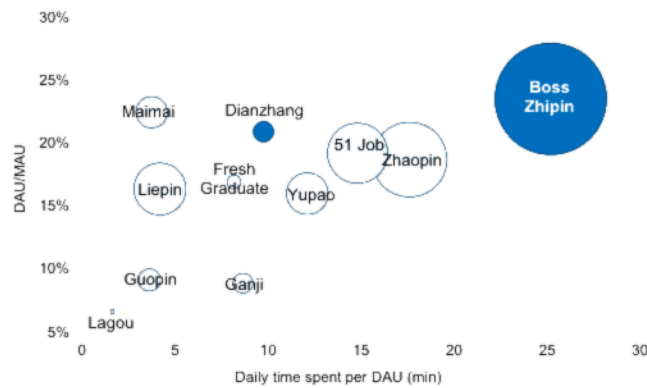
Exhibit 68: Active users of TAL and Zuoyebang learning tablets grew +73% yoy/+41% yoy in Jun (vs. +100% yoy/+56% yoy in May)
MAUs of smart education device apps (mn)



Source: Questmobile, Data compiled by Goldman Sachs Global Investment Research

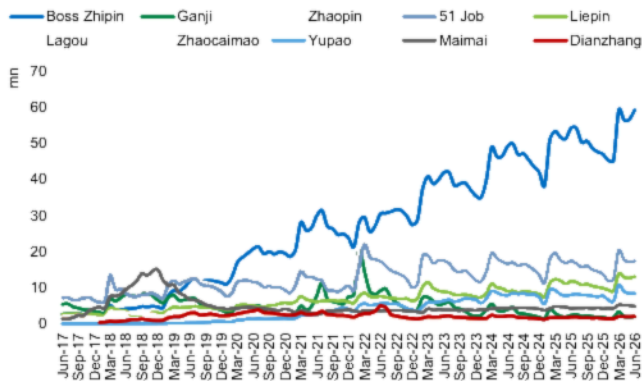
Online recruitment

Exhibit 69: As China’s largest online recruitment platform, Boss Zhipin continues to lead the industry in user engagement, including DAU/MAU and time spent per DAU (Jun 2026)



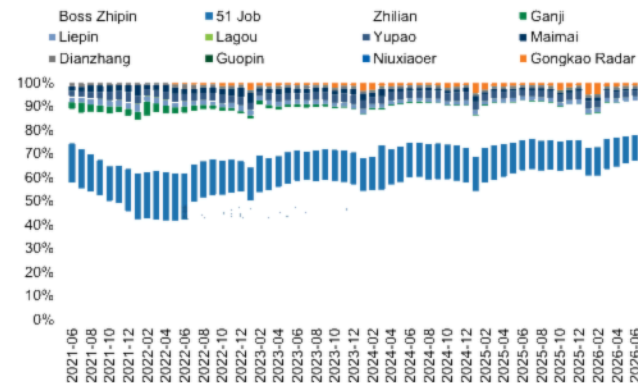
Source: QuestMobile

Exhibit 70: Boss Zhipin remains one of the largest online recruitment platforms in China by MAU, registering 59.2mn MAUs with +9% yoy growth in Jun 2026
MAUs of major recruitment platforms (mn)



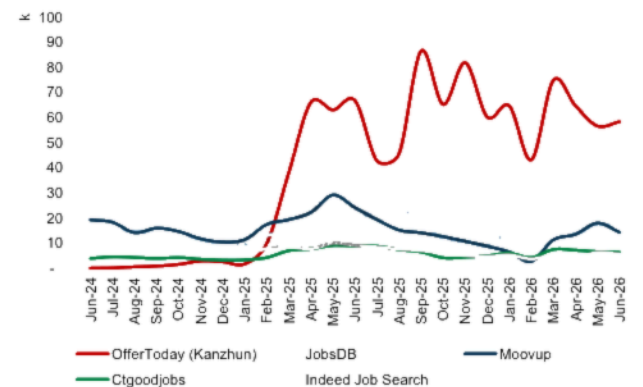
Source: QuestMobile

Exhibit 71: ...and claimed 67% time spent share amongst major recruitment platforms in China (Jun 2026)



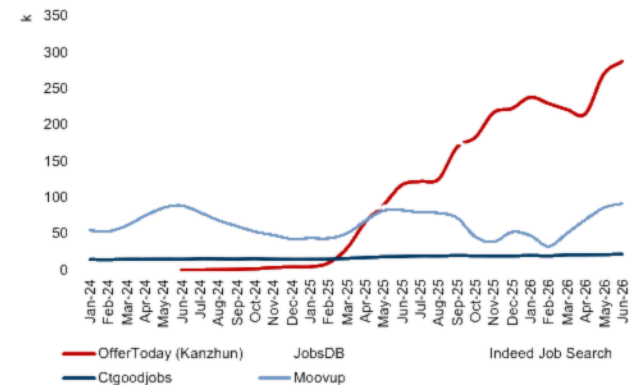
Source: QuestMobile

Exhibit 72: OfferToday, as the HK platform of Boss Zhipin, significantly outperformed its peers in HK for app downloads (3.3 times vs. No.2 player JobsDB in Jun 2026)
App downloads of major HK recruitment platforms



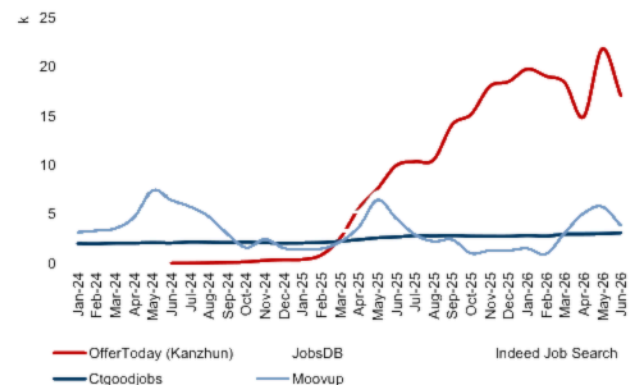
Source: Sensor Tower

Exhibit 73: OfferToday continues to lead in MAU, registering c.287k MAUs with +145% yoy growth in Jun 2026
MAUs of major HK recruitment platforms



Source: Sensor Tower

Exhibit 74: OfferToday remains its leading position for DAU, registering c.17k DAUs with +71% yoy growth in Jun 2026
DAUs of major HK recruitment platforms



Source: Sensor Tower

Real estate brokerage/software

Exhibit 75: Time spent on Beike's three major apps dropped -7% yoy/-6% mom in Jun 26 (vs. +1% yoy/-1% mom in May 2026)

Time spent yoy change



Source: Questmobile

Exhibit 76: Beike Zhaofang DAU saw -11% yoy/-2% mom decline in Jun 2026, while A+ DAU saw +3% yoy growth or flat mom in Jun 2026

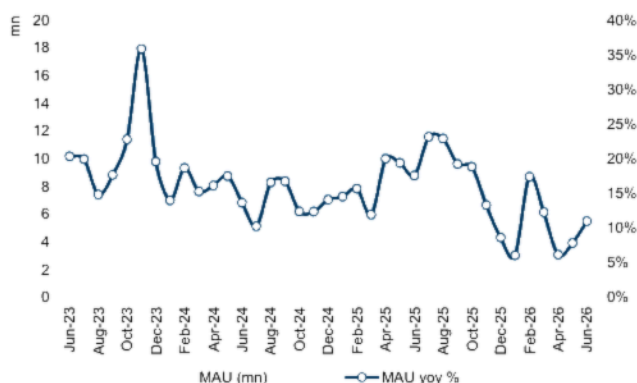
Average DAUs (10k)



Source: Questmobile

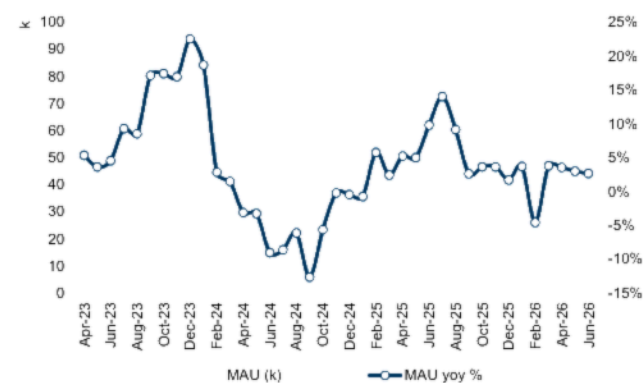
Auto aftermarket service

Exhibit 77: Tuhu app's MAU increased +11% yoy/+2% mom to 16.9mn in Jun 2026 (vs. +8% yoy/+1% mom in May 2026)



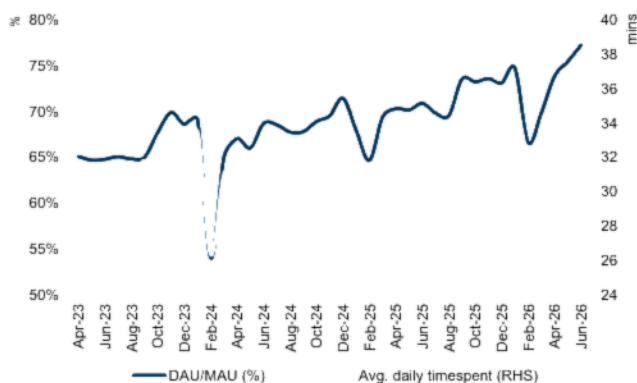
Source: Questmobile

Exhibit 78: Tuhu merchant app MAUs increased +3% yoy or dropped -2% mom to 85.1k in Jun 2026 (vs. +3% yoy/-1% mom in May 2026)



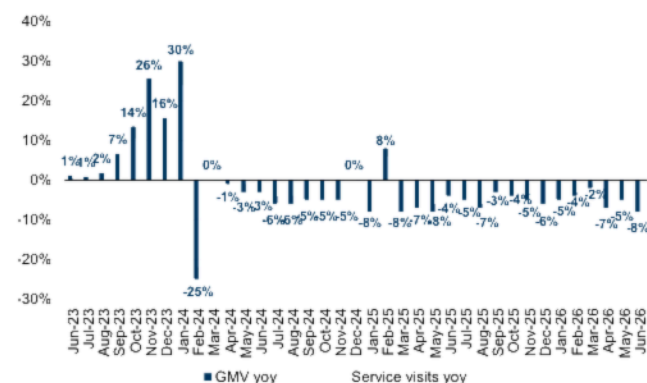
Source: Questmobile

Exhibit 79: Tuhu merchant DAU/MAU ratio increased +6pp yoy to 77% in Jun 2026, while daily time spent per DAU saw +1% yoy growth to 32.5 mins in Jun 2026



Source: Questmobile

Exhibit 80: For auto aftermarket stores served by F6, GMV saw -8% yoy/-3% mom decline in Jun 2026, while service visits saw -8% yoy/-3% mom decline in Jun 2026



Source: F6 data

Price Target Risks and Methodology – MiniMax Group

We are Buy rated on MiniMax with a 12-month target price of HK\$800. Our valuation is based on DCF valuation, with 12% WACC and 2% terminal growth rate.

Key risks: Weaker-than-expected model performance amid the competition in the global foundation model industry; Slower-than-expected path to profit visibility; Weaker-than-expected commercialization capability; risks around IP/content generation; cash burn/self-funding capability; risks around geopolitics under an intensified tech race between the US and China.

Price Target Risks and Methodology – Z.AI

We are Neutral rated on Z.AI with a 12-month target price of HK\$1,610. Our valuation is based on DCF valuation, with 12% WACC and 2% terminal growth rate (consistent with our coverage).

Key risks: **(+)** Stronger-than-expected model intelligence, **(+)** Faster-than-expected path to profit visibility, **(+)** Additional non-inference take rate revenue streams, **(+)** Stronger-than-expected commercialization capability, **(-)** Competition in global foundation model industry, **(-)** Limited near-term profit visibility due to high R&D expenses, **(-)** Cash burn/self-funding capability, **(-)** Risk around geopolitics under an intensified tech race between the US and China.