Everyone Saw the Earnings. Few Saw This.

Published June 13, 2025

Christopher Gannatti, CFA
Global Head of Research

Key Takeaways

- While Nvidia's record-breaking earnings grabbed headlines, its quiet release of NVLink Fusion reveals a deeper strategy to entrench itself as the indispensable backbone of global AI infrastructure.
- In contrast to AMD's hardware-centric approach, Nvidia is winning where it matters most—in systems integration, developer ecosystems and flexible global partnerships that position it as a platform, not just a chipmaker.
- For investors seeking exposure to enduring AI leadership rooted in profitability and growth fundamentals, the WisdomTree U.S. Quality Growth Fund (QGRW) offers a timely and differentiated approach beyond pure market-cap plays.

It is a strange thing to say, but in 2025, Nvidia is somehow still misunderstood. Everyone knows the earnings numbers. Every news outlet on earth has already covered the $44 billion in quarterly revenue, the 70% year-on-year growth and the $3.3 trillion market cap.1 But the story isn't in the earnings anymore. The story is in the architecture Jensen Huang is building—not just of hardware, but of influence, platform control and geopolitical optionality.

We are no longer watching a chip company. We are watching the formation of something closer to a new operating system for AI infrastructure. And to see it clearly, we have to look where others aren't: at NVLink Fusion, at the strange ballet with China, the delayed MI325X and the rental rates of graphics processing units (GPUs) in a fractured, competitive neocloud2 economy.

Nvidia's Most Important Move No One's Talking About

Back in March, Nvidia announced NVLink Fusion, a development that didn't generate the headlines it should have. On the surface, it looked like a technical detail—a new high-bandwidth interconnect for heterogeneous systems. But in practice, it was a profound statement of strategic repositioning. NVLink Fusion means you can pair Nvidia's GPUs with someone else's central processing units (CPUs) or accelerators. In effect, it dissolves the binary between "buy all-Nvidia" and "build custom." It makes Nvidia the connective tissue of other people's silicon.3 And that has implications far beyond engineering.

One of those implications lies in China. Recently, Nvidia took a $5 billion write-down because it couldn't sell certain chips into the Chinese market.4 On cue, pundits began writing eulogies for Nvidia's China

business. But then Jensen Huang was spotted visiting Beijing. Then came the idea of a research center. There is, it seems, a strange dance happening between what the U.S. government will legally allow and what Nvidia is trying to keep alive. If NVLink Fusion can allow Blackwell GPUs to be paired with domestic Chinese CPUs or accelerators—without violating U.S. export laws—then it might just be the back door through which Nvidia maintains relevance in one of the world's most important AI markets.

This isn't a gamble. It's a hedge. And Nvidia has always hedged well. That is, in part, what makes it different from Advanced Micro Devices (AMD).

The Difference between a Chip and a Platform

Much has been said lately about AMD's potential to compete with Nvidia in inference compute.5 The performance per dollar varies wildly depending on the type of workload—chat applications, document retrieval, long-context reasoning. And for companies that purchase and own their GPUs outright, the story is mixed. Sometimes AMD wins. Sometimes Nvidia does.

But the moment you enter the rental market, Nvidia wins every time. The reason is both mundane and revealing: market depth. There are over a hundred neocloud providers offering Nvidia GPUs—H100s, H200s and, soon, B200s.

To pause, many may not have heard the term "neocloud" before, but those same people may have seen the recent CoreWeave initial public offering. CoreWeave is a major neocloud provider.

AMD GPUs, by contrast, are barely available, and when they are, the rental rates are too high to justify the switch. It doesn't matter how good the MI325X is in isolation when compared to the Nvidia options. What matters is what it costs, when it ships, how easy it is to deploy and whether your engineers can make it work without weeks of configuration headaches.6

And that's where the real gulf between AMD and Nvidia lies: not in floating-point operations per second, but in systems. Nvidia has invested relentlessly in the entire stack. Software like TensorRT-LLM and CUDA7 remains far ahead of AMD's ROCm.8

Simply put, these represent the developer interaction layers such that useful systems can be designed to run on Nvidia and AMD systems. Many of us do not write software, but we have selected either the iOS (Apple) or Android (Google) operating systems for our smartphones. Is it possible for us to switch, one to the other? Yes. Do they make it easy to carry over all of the features you might like? No.

Nvidia provides containers, developer tooling and, increasingly, orchestration frameworks like Dynamo that include features for disaggregated inference, smart routing, KV cache offloading—tools that make the messy reality of serving large models tractable.9

By contrast, AMD is still cleaning up the basics. ROCm has seen improvement, but continuous integration (CI) coverage remains sparse. There are accuracy regressions. Inferencing with AMD means dealing with broken DeepSeek implementations and a stack of flags that require close collaboration with AMD engineers just to get basic performance parity. Worse, AMD spends more quarterly on stock buybacks than

it does on internal GPU cluster resources for ROCm development.10 That may thrill some shareholders, but it doesn't inspire confidence from developers.

The Contest Was Never Justabout Chips

Even when AMD hardware wins a benchmark, Nvidia wins the deployment. The reason is simple: Jensen Huang understands that you cannot build moats out of chips alone. Moats are built from systems, developer mindshare, ecosystem depth and, above all, speed. It is telling that in many inference scenarios, the H200 with TensorRT-LLM still beats MI325X and MI300X, not because the silicon is always better, but because the integration is tighter, the software more mature and the delivery more complete.11

It is also why Nvidia is gaining traction not only with hyperscalers, but with sovereign clients. AI is becoming a strategic national capability, and countries like Saudi Arabia, UAE and Taiwan aren't just looking to buy chips—they're looking to build infrastructure. Nvidia's platform-level approach, combined with its willingness to partner flexibly via NVLink Fusion, offers those clients a tantalizing proposition: sovereignty without fragmentation. They can build with domestic CPUs, or integrate national silicon, while still accessing the best accelerators and orchestration layers in the world.

That is a very different playbook from AMD's, and it's one that makes Nvidia harder to displace not just technically, but politically. In a world where compute is becoming policy, Nvidia has positioned itself as an acceptable partner to multiple sides. It plays nice with Trump-era export restrictions. It opens research centers in China. It talks up Middle East partnerships and ships to European supercomputers. It is hardware diplomacy as much as it is hardware design.

2025: Critical in Cementing Strategic Leadership

That is what makes 2025 such a critical year. Jensen Huang isn't just outpacing competitors; he's redefining the rules of competition. AMD could still win slices of this market—especially in high-latency, dense model inference where MI325X can beat H200s on total cost of ownership. But those wins are becoming pyrrhic. If no one can deploy your chip easily, if few can rent it competitively and if your own software isn't fully tested—then how much does the raw performance really matter?12

The chip war is evolving into a systems war. And Nvidia, for all its flaws and delays, is still the only company building systems that win not just in the lab, but in the world. It is misunderstood because people keep trying to measure it like a semiconductor stock. But it is something else now. Nvidia is becoming the infrastructure layer of the AI century. Not a chip company. Not even a systems company. A new kind of operating system, hiding in plain sight.

Conclusion: AI Trades and Tariff Announcements Have Driven Performance

Enter the WisdomTree U.S. Quality Growth Fund (QGRW), a fundamentally driven alternative, designed to track the total return performance of the WisdomTree U.S. Quality Growth Index, which focuses toward the more profitable, higher-growth names in large-cap U.S. equities. There is a direct emphasis on return on equity, return on assets and earnings growth. This is in contrast with Nasdaq 100 Index constituents, which do not have any such direct requirements. For investors who believe the AI boom still has legs—but

want a basket built on enduring fundamentals, not just market cap—QGRW may offer the right lens for this next leg of the tech trade.

Ever since the launch of ChatGPT in November 2022, we have seen large, U.S. tech companies have led markets higher. That narrative was challenged a bit in 2025, but it seems that on any sort of tariff pause or tariff lowering, the importance of these big tech companies comes back to the fore.

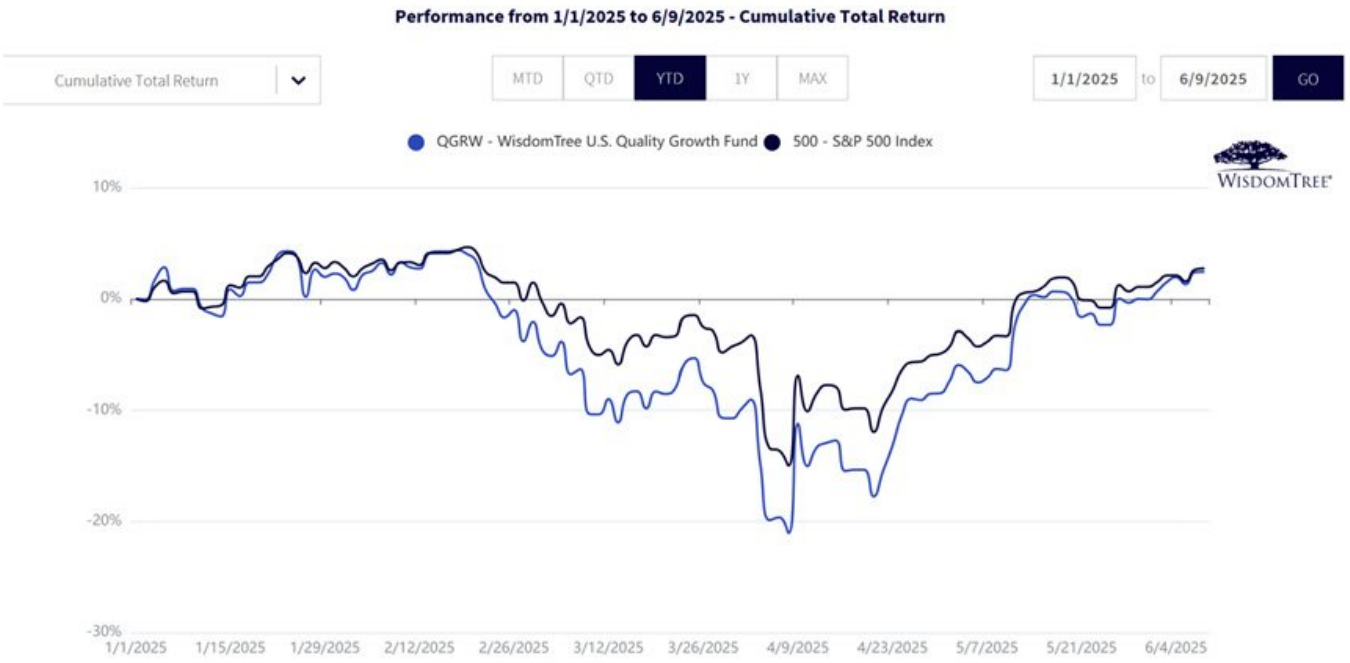
Figure 1: Standardized Returns

Fund Name	Fund Ticker Symbol	Fund Inception Date	Fund Expense Ratio	Year-to-Date	1-Year	3-Year	5-Year	10-Year	Since Fund Inception
WisdomTree U.S. Quality Dividend Growth (NAV	QGRW	12/15/22	0.28%	-10.73%	7.06%	N/A	N/A	N/A	27.77%
WisdomTree U.S. Quality Dividend Growth (MP)	QGRW	12/15/22	0.28%	-10.76%	7.11%	N/A	N/A	N/A	27.78%
S&P 500 Index				-4.27%	8.25%	9.06%	18.59%	12.50%	N/A

Sources: WisdomTree, FactSet, specifically data from the Fund Comparison Tool in the PATH suite of tools, accessed 5/29/25, with returns as of 3/31/25. NAV denotes total return performance at net asset value. MP denotes market price performance. **The performance data quoted represents past performance. Past performance is not indicative of future results. Investment return and principal value of an investment will fluctuate so that an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than the performance data quoted. For the most recent month-end and standardized performance, click here.**

As we speak to investors, it is remarkable that with everything that has happened so far in 2025, we are talking about roughly flat year-to-date performance, as we see in figure 2. People looking for that responsiveness and that exposure to these big AI companies like Nvidia have a notable strategy, QGRW.

Figure 2: A Long Journey to a Roughly Flat Performance Experience



Sources: WisdomTree, FactSet, specifically data from the Fund Comparison Tool in the PATH suite of tools, accessed 6/10/25, with returns as of 6/9/25. NAV denotes total return performance at net asset value. MP denotes market price performance. **Past performance is not indicative of future results. Investment return and principal value of an investment will fluctuate so that an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than the performance data quoted. For the most recent month-end and standardized performance, click here.**

1 Source: "NVIDIA Corp. (NVDA.O): Top Pick – Semiconductors | United States of America," *Morgan Stanley Research*, 5/29/25.

2 In this context, a **neocloud** refers to a new generation of cloud service providers that specialize in renting out high-performance GPUs—like Nvidia's H100 or H200—for AI and machine learning workloads.

3 Source: "NVLink Fusion: Nvidia's Strategy to Dominate AI Infrastructure Through Interoperability," *Investor's Business Daily*, May 2025.

4 Source: Gopalan, N. Nvidia to record $5.5B charge as US cracks down on chip exports to China. *Investopedia*, 4/16/25.

5 Inference compute refers to the running of trained AI models for users such that they can generate outputs.

6 Source: "Inference Benchmarking and Total Cost of Ownership Analysis: H100, H200, B200, MI300X, MI325X, and MI355X," SemiAnalysis, May 2025.

7 TensorRT-LLM and CUDA represents avenues through which developers can interact with Nvidia processors.

8 ROCm represents the avenue through which developers can interact with AMD processors.

9 Source: SemiAnalysis, 2025.

10 Source: SemiAnalysis, 2025.

11 Source: SemiAnalysis, 2025.

12 Source: SemiAnalysis, 2025.

Important Risks Related to this Article

There are risks associated with investing, including the possible loss of principal. Growth stocks, as a group, may be out of favor with the market and underperform value stocks or the overall equity market. Growth stocks are generally more sensitive to market movements than other types of stocks. The Fund is non-diversified; as a result, changes in the market value of a single security could cause greater fluctuations in the value of Fund shares than would occur in a diversified fund. The Fund invests in the securities included in, or representative of, its Index regardless of their investment merit. The Fund does not attempt to outperform its Index or take defensive positions in declining markets, and the Index may not perform as intended. Please read the Fund's prospectus for specific details regarding the Fund's risk profile.