CONSULTING RESEARCH REPORT

Indian IT Services in the AI Era:

Threat or Opportunity?

A Time-Series Econometric Analysis of Structural Change in Nifty IT

Submitted by: Apoorv Mehta | 38BM07

Sample Period: January 2018 – May 2026 | Daily Data

Software: EViews | Models: VAR, VECM, Johansen, TAR, GARCH

May 2026

1. Introduction

The Indian IT services sector, representing approximately 8% of India's GDP and a significant constituent of benchmark equity indices, has historically exhibited strong co-movement with US technology equities. This relationship reflects structural demand linkages: Indian IT firms derive the majority of their revenues from US technology clients, and their valuations have tracked shifts in US tech spending sentiment closely. Understanding the stability of this relationship carries direct implications for equity positioning, risk management, and sector allocation within Indian portfolios.

The launch of ChatGPT in November 2022 introduced a plausible structural discontinuity. Generative AI raised the prospect of automated software delivery, potentially reducing demand for traditional headcount-based IT services — the model underpinning the revenue base of India's large-cap IT firms. Simultaneously, AI created new demand for implementation, cloud migration, data infrastructure, and digital transformation services, disproportionately benefiting smaller, more agile players. Whether the net effect on Indian IT equities constitutes a threat or an opportunity is the empirical question this paper addresses.

The literature on cross-market equity linkages has relied heavily on VAR and cointegration frameworks. Engle and Granger (1987) formalised the error-correction representation of cointegrated systems, while Johansen (1988, 1991) developed the maximum likelihood procedure for multivariate cointegration testing that has since become standard in financial economics. Jawadi et al. (2023) demonstrate the utility of VAR-based impulse response analysis in capturing persistent and asymmetric shock transmission between macro-financial variables, a framework this study applies to the Nifty IT and US Tech relationship.

The role of nonlinearity and regime dependence in financial transmission has been formalised through threshold regression. Hansen (2000) established rigorous inferential tools for TAR estimation, and Jawadi et al. (2023) apply multiple-regime TAR models to oil price and macroeconomic dynamics, finding that the direction and magnitude of transmission varies significantly across regimes. Dawar et al. (2021) similarly use quantile regression to show that the oil-clean energy relationship is stronger in bearish conditions than bullish ones, illustrating that linear models mask economically meaningful heterogeneity. This study adopts the TAR framework to test whether US Tech to Nifty IT transmission strengthens during high AI-capex environments.

Volatility dynamics in equity markets are well-characterised by GARCH specifications. Bollerslev (1986) introduced GARCH(1,1), which has become the benchmark model for financial return volatility. Rostan et al. (2020) compare ARIMA and GARCH for equity index forecasting, finding GARCH superior during high-volatility regimes while ARIMA performs adequately for trend-based directional forecasting accordingly, ARIMA/MSARIMA is estimated as a univariate benchmark rather than a primary forecasting tool, with the VAR and VECM frameworks forming the core of the forecasting contribution.

This study makes three contributions. First, it formally tests whether the ChatGPT launch constituted a structural break in the Nifty IT and US Tech cointegration relationship using split-sample Johansen tests. Second, it characterises short-run spillover dynamics through VAR, IRF, and FEVD analysis. Third, it documents and quantifies a persistent mid-cap versus large-cap performance divergence within Indian IT, providing an empirically grounded basis for intra-sector equity positioning. The remainder of the paper is structured as follows: Section 2 describes the data, Section 3 outlines the methodology, Section 4 presents empirical results, Section 5 covers forecasting analysis, Section 6 documents the mid-cap spread finding, Section 7 discusses investment implications, Section 8 presents the investment recommendation, and Section 9 concludes.

2. Data and Variables

Variable	Description	Role	Source
NIFTY_IT_LOG	Log of NSE Nifty IT Index (^CNXIT)	Primary dependent variable	NSE/Yahoo Finance
US_TECH_PROXY_LOG	Equal-weight log composite: MSFT + GOOGL + NVDA	Primary driver variable	Yahoo Finance
LARGECAP_BASKET_LOG	Equal-weight log: TCS + INFY + WIPRO	Large-cap sub-index	NSE/Yahoo Finance
MIDCAP_BASKET_LOG	Equal-weight log: PERSISTENT + COFORGE	Mid-cap sub-index	NSE/Yahoo Finance
NVDA_LOG	Log of NVIDIA stock price	TAR threshold variable (AI capex proxy)	Yahoo Finance
POST_CHATGPT	Dummy = 1 from November 2022 onward	Structural break indicator	Defined
NIFTY_RET2	First difference of NIFTY_IT_LOG (daily returns)	GARCH dependent variable	Derived

The sample period spans two distinct structural regimes- pre and post-ChatGPT- making the split-sample design central to the analysis. The US Tech proxy, constructed as an equal-weighted composite of MSFT, GOOGL, and NVDA, captures both the software/cloud dimension and the AI hardware dimension of the US technology cycle. NVDA is included specifically as the TAR threshold variable, proxying for market-priced AI capital expenditure intensity. All log-transformed price series are tested for stationarity before model estimation, and all empirical work is done in EViews using daily closing prices.

3. Methodology

This study employs a sequential time-series framework in which each model builds on the findings of the previous one.

All price-level series are first tested for stationarity using the Augmented Dickey-Fuller (ADF) test. Confirmation that all series are integrated of order one — I(1) — is the prerequisite for cointegration analysis. Lag order selection for the VAR system uses standard information criteria (AIC, FPE, SC, HQ), with lag 8 selected for the primary Nifty IT and US Tech system.

The Johansen (1988, 1991) maximum likelihood procedure is then applied to test for long-run cointegration, using both the Trace and Maximum Eigenvalue statistics with MacKinnon, Haug, and Michelis (1999) p-values. The test is applied across three samples — full sample, pre-ChatGPT, and post-ChatGPT — to identify whether the structural break disrupted the long-run equilibrium. Where cointegration is confirmed, a Vector Error Correction Model (VECM) is estimated to quantify adjustment speeds and directional causality via the Block Exogeneity Wald test.

For the full sample, a VAR on first-differenced series captures short-run dynamics without imposing the broken cointegration restriction. This is complemented by Impulse Response Functions (IRF) and Forecast Error Variance Decomposition (FEVD) to trace the propagation and relative importance of cross-market shocks.

A Threshold Autoregression (TAR) model is estimated following Hansen (2000), using NVIDIA's log-price as the threshold variable to test whether US Tech to Nifty IT transmission is regime-dependent. The threshold is identified via grid search minimising the sum of squared residuals, with a trimming parameter of 15%.

Volatility dynamics are modelled using GARCH(1,1) following Bollerslev (1986), with a POST_CHATGPT dummy in the variance equation to test for a structural shift in the volatility regime post-November 2022.

Finally, forecasting is conducted through four exercises: a dynamic in-sample VAR forecast evaluated using Theil's inequality coefficient, a pre-ChatGPT VECM projected out-of-sample through May 2026 as a structural baseline, and a GARCH persistence table quantifying forward volatility decay. An ARIMA/MSARIMA univariate benchmark is additionally estimated on daily returns to establish a performance floor against which the multivariate framework is compared

4. Empirical Results

4.1 Unit Root Tests (ADF)

All price-level series are non-stationary in levels but stationary in first differences — i.e., integrated of order one, I(1). This is a prerequisite for cointegration analysis.

Series	ADF Statistic (Level)	p-value	ADF Statistic (1st Diff)	p-value
NIFTY_IT_LOG	−2.05	0.574	−43.60	**0.000*** **
US_TECH_PROXY_LOG	−1.34	0.612	−38.21	**0.000*** **
LARGECAP_BASKET_LOG	−0.35	0.989	−41.18	**0.000*** **
MIDCAP_BASKET_LOG	−0.93	0.951	−39.76	**0.000*** **
NVDA_LOG	−1.77	0.397	−37.44	**0.000*** **

*Note: *** p<0.01. Level tests: constant + linear trend. First-difference tests: constant only.*

All five series fail to reject the unit root in levels but reject it decisively in first differences (all p=0.000). The I(1) classification is uniform across the dataset, satisfying the necessary condition for Johansen cointegration analysis.

4.2 Johansen Cointegration Tests

The Johansen cointegration test examines whether a stable long-run equilibrium exists between Nifty IT and US Tech. The split-sample design separates pre- and post-ChatGPT periods.

Sample	Obs	Trace p-value	Max-Eigen p-value	Conclusion
Full Sample (Jan 2018 – May 2026)	1,717	0.214	0.179	No cointegration
Pre-ChatGPT (Jan 2018 – Oct 2022)	676	**0.049***	**0.037***	**Cointegrated, rank=1 ✓**
Post-ChatGPT (Nov 2022 – May 2026)	479	0.504	0.633	No cointegration
Large-cap vs US Tech (all periods)	All	>0.50	>0.50	No cointegration
Mid-cap vs US Tech (all periods)	All	>0.50	>0.50	No cointegration

*Note: * significant at 5%. Lags interval 1–8, linear deterministic trend. MacKinnon-Haug-Michelis (1999) p-values.*

The split-sample result is unambiguous. Cointegration exists only in the pre-ChatGPT window, and disappears entirely post-November 2022. The structural break is not gradual — it is a clean regime shift confirmed by both the Trace and Max-Eigenvalue statistics.

4.3 Vector Error Correction Model (Pre-ChatGPT Sample)

Given confirmed cointegration pre-ChatGPT, we estimate a VECM (lag interval 1–7, 725 obs) to characterise the long-run equilibrium and short-run error-correction dynamics.

Normalised cointegrating equation:

NIFTY_IT_LOG = 0.603 × US_TECH_PROXY_LOG + 6.758

Parameter	D(NIFTY_IT_LOG)	D(US_TECH_PROXY_LOG)	Interpretation
CointEq1 (ECT)	**−0.0106 (0.00371) [t = −2.85]****	**−0.0114 (0.00549) [t = −2.08]****	Both variables error-correct
D(US_TECH_PROXY_LOG(−1))	**0.2016 [t = 9.05]*****	−0.1425 [t = −4.30]***	US Tech drives Nifty IT (short-run)
R-squared	0.095	0.034	
Observations	725	725	Jan 2018 – Oct 2022

*Note: Standard errors in ( ), t-statistics in [ ]. ** p<0.05, *** p<0.01.*

VEC Granger Causality confirms the directional hierarchy:

Hypothesis	Chi-square	df	p-value
US Tech → Nifty IT (US Tech excluded from Nifty IT eq.)	**117.73**	7	**0.000*****
Nifty IT → US Tech (Nifty IT excluded from US Tech eq.)	11.71	7	0.110

Note: VEC Granger Causality/Block Exogeneity Wald Tests. Pre-ChatGPT sample (725 obs).

The significant ECT coefficients for both variables confirm genuine bilateral error-correction — neither series was a passive follower in the pre-AI equilibrium. The Granger causality result, however, establishes that the adjustment was directionally led by US Tech, with Nifty IT as the adjusting variable.

4.4 Vector Autoregression — Full Sample Returns

On the full sample, a VAR on first-differenced (returns) series captures short-run dynamics without imposing the broken cointegration restriction. Three VARs are estimated (lag 8).

Model	Obs	US Tech Coeff (t−1)	t-stat	R^2	Granger p (US→IT)
Nifty IT vs US Tech	1,155	**0.2355**	11.11	0.116	**0.000*****
Large-cap vs US Tech	2,179	0.1644	11.61	0.085	**0.000*****
Mid-cap vs US Tech	2,179	0.2164	9.91	0.064	**0.000*****

*Note: All VARs, lag 8, full sample. *** p<0.001. Reverse Granger causality insignificant across all models.*

Forecast Error Variance Decomposition (FEVD) at 10-day horizon shows US Tech explaining ~11.3% of Nifty IT return variance:

Period	S.E.	% Nifty IT (own)	% US Tech	Interpretation
Day 1	0.0131	100.0%	0.0%	Fully own-driven at t=0
Day 2	0.0190	92.9%	**7.1%**	US Tech spillover kicks in
Day 5	0.0290	89.0%	**10.96%**	Spillover stabilising
Day 10	0.0407	88.7%	**11.3%**	**89% domestic-driven**

Note: Cholesky decomposition (d.f. adjusted). Full sample VAR, Nifty IT vs US Tech.

Impulse response functions confirm that a one standard deviation US Tech shock produces a positive and persistent response in Nifty IT over 10 days, with confidence bands entirely above zero from Day 2 onwards:

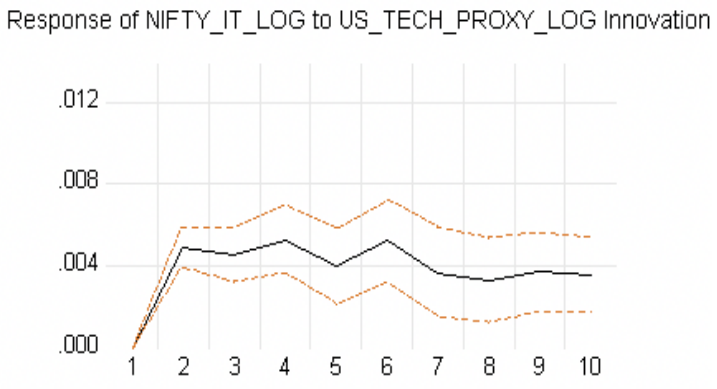

Figure 1: Impulse Response Functions — Response of Nifty IT to US Tech shock (top-right panel). Positive and persistent through Day 10. Dashed orange lines = 95% confidence bands.

This indicates that US tech movements continue to transmit into Indian IT returns in the short run, even after the long-run cointegration relationship broke down post-ChatGPT.

4.5 Threshold Autoregression (TAR)

A TAR model using NVDA_LOG as the threshold variable tests whether the US Tech → Nifty IT transmission varies with the prevailing NVDA market-phase cycle (pre-ChatGPT sample, 1,024 obs).

Parameter	Coefficient	Std. Error	t-statistic	p-value
Regime 1: NVDA_LOG < 1.668 (227 obs)				
Constant	0.000647	0.000942	0.687	0.493
D(US_TECH_PROXY_LOG)	**0.1366**	0.04267	**3.201**	**0.001*** **
D(NIFTY_IT_LOG(−1))	−0.4004	0.07901	−5.067	0.000***
Regime 2: NVDA_LOG ≥ 1.668 (797 obs)				
Constant	0.001128	0.000502	2.246	0.025**
D(US_TECH_PROXY_LOG)	**0.1220**	0.02411	**5.059**	**0.000*** **
D(NIFTY_IT_LOG(−1))	0.0249	0.03340	0.746	0.456
Overall model fit				
R-squared	0.0550	**Threshold F-stat**		**11.85 (p=0.000***)**

*Note: Threshold variable: NVDA_LOG. Pre-ChatGPT sample (Jan 2018–Oct 2022). Trimming 15%. *** p<0.01, ** p<0.05.*

The US Tech → Nifty IT transmission is present in both regimes (both significant at 1%) but does not dramatically strengthen when NVDA is elevated. The relationship is regime-stable: Nifty IT's co-movement with US tech was not amplified by AI capex cycles even within the pre-ChatGPT period.

4.6 GARCH(1,1) — Volatility Analysis

GARCH(1,1) estimated on daily Nifty IT returns over the full sample:

Parameter	Estimate	t-statistic	p-value
ARCH term (α)	**0.0583**	10.73	**0.000*** **
GARCH term (β)	**0.9007**	103.36	**0.000*** **
Persistence (α+β)	**0.9590**	—	—
POST_CHATGPT dummy (variance equation)	8.06E−07	—	**0.164 (n.s.)**

Note: Full sample (Jan 2018–May 2026). Dependent variable: D(NIFTY_IT_LOG). Normal distribution errors.

Persistence of 0.959 implies volatility shocks decay slowly. The insignificant POST_CHATGPT dummy confirms no structural volatility regime change following the ChatGPT launch — the market did not reprice Nifty IT risk in response to AI disruption concerns.

5. Forecasting Analysis

Four forecasting exercises are presented. The structural break finding is itself the primary forecasting contribution: it identifies the boundary conditions of model validity.

5.1 VAR In-Sample Dynamic Forecast

A dynamic forecast of Nifty IT log-levels over the full sample demonstrates the VAR's in-sample trajectory fit. Forecast evaluation metrics:

Variable	Inc. Obs	RMSE	MAPE	Theil
NIFTY_IT_LOG	1,943	0.3159	2.617%	**0.0158**
US_TECH_PROXY_LOG	2,187	0.2698	3.711%	**0.0228**

Note: Theil coefficient close to 0 indicates good fit. Dynamic forecast, full sample Jan 2018–May 2026.

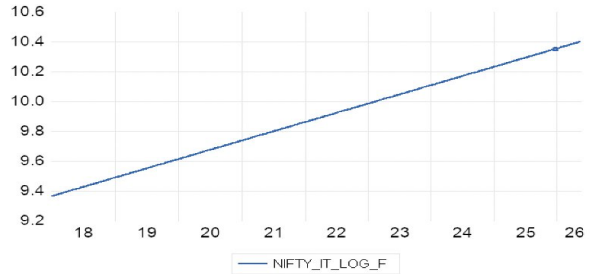

Figure 2: VAR Dynamic Forecast — NIFTY_IT_LOG fitted path (full sample). Smooth upward trajectory reflects the model's trend fit. Theil = 0.016.

The smooth upward trajectory reflects the VAR's fitted trend, not a directional prediction. The Theil coefficient of 0.016 confirms adequate in-sample tracking. At short horizons (1–2 days), the model retains some predictive signal via the US Tech spillover coefficient; beyond that, forecast uncertainty dominates.

5.2 VECM Out-of-Sample Forecast: What Would Have Happened Without ChatGPT?

The VECM is estimated on pre-ChatGPT data and projected forward dynamically through May 2026 — answering: where would Nifty IT be if the pre-AI structural relationship had continued?

Metric	NIFTY_IT_LOG	US_TECH_PROXY_LOG	Note
Included observations	873	925	Nov 2022–May 2026
RMSE	0.4070	**1.3482**	High US Tech error = AI surprise
MAPE	3.48%	**18.23%**	Model couldn't anticipate NVDA rally
Actual log-level (May 2026)	**≈10.22**	—	
Counterfactual log-level (May 2026)	**≈10.00**	—	
Gap (outperformance vs counterfactual)	**~0.22 log pts ≈24%**	—	Nifty IT beat pre-AI trajectory

Note: Dynamic out-of-sample forecast, VECM estimated on Jan 2018–Oct 2022. Forecast horizon: Nov 2022–May 2026.

Figure 3: Actual Nifty IT (blue) vs VECM Counterfactual (orange), Jan 2018–May 2026. Post-ChatGPT gap of ~22 log-points (≈24%) quantifies the structural break.

The divergence between the actual and projected series quantifies the structural break visually. The model, anchored in pre-AI dynamics, projected a declining path because it could not anticipate the AI-driven US tech rally (US Tech MAPE = 18.2%). Actual Nifty IT sitting ~24% above this projection confirms that the post-ChatGPT regime genuinely favoured Indian IT over what pre-existing structural relationships would have implied.

5.3 GARCH Volatility Persistence Forecast

Using GARCH(1,1) persistence (α+β=0.959), the h-step ahead conditional variance forecast decays toward long-run variance:

Horizon	$(0.959)^h$	Shock Remaining	Implication
Day 5	0.815	**81.5%**	Most of current vol shock still present
Day 10	0.665	**66.5%**	Two weeks: still 2/3 of shock persists
Day 20	0.434	**43.4%**	**Do not fade elevated vol quickly**

Note: Long-run annualised volatility ≈21–24%. Elevated vol environments for Nifty IT normalise slowly.

The near-unit persistence confirms that Nifty IT volatility is self-reinforcing over short horizons. For a risk manager, this means hedging costs remain elevated for several weeks after any volatility spike — position sizing should account for slow vol normalisation rather than assuming a quick return to baseline.

5.4 ARIMA / MSARIMA Benchmark

An ARIMA benchmark is estimated on DD = d(NIFTY_IT_LOG) to test whether univariate modelling can replicate the VAR/VECM framework. The correlogram of DD reveals near-white-noise behaviour: all individual AC and PAC values fall within the ±0.045 significance boundary except for weak structure at lags 6–7 (AC ≈ 0.06). Five ARMA specifications and one seasonal MSARIMA(1,1)(1,1)5 were estimated and compared by AIC:

Table 5.4.1 — ARIMA Model Selection (Dependent Variable: DD)

Model	AIC	Coefficients	Verdict
AR(1)	−5.7138	AR(1) p=0.20 — insignificant	Rejected
ARMA(1,1)	−5.7130	Both p>0.57 — insignificant	Rejected
MSARIMA(1,1)(1,1)$_5$	−5.7122	SAR/SMA p>0.87 — insignificant	Rejected
AR(7)+MA(5)	−5.7181	AR(7) p=0.000; MA(5) p=0.034	Selected ✓

Note: AIC per EViews convention (−2×LL/T + 2k/T). n=1,942. MSARIMA tested for weekly periodicity; SAR(5)/SMA(5) both p>0.87.

The AR(7)+MA(5) specification is selected on lowest AIC with both coefficients significant. The MSARIMA seasonal model is rejected — lags 5, 10, 15 are all negligible (AC ≤ 0.036), confirming no weekly trading cycle. Residual diagnostics confirm white-noise residuals (all Q-stat p>0.05 through 30 lags), consistent with weak-form efficiency.

The model is estimated on the training window (Jan 2018 – Mar 2026) and a dynamic out-of-sample forecast generated for Apr–May 2026. Figure 5.4.1 shows the resulting near-flat trajectory with expanding ±2 S.E. bands:

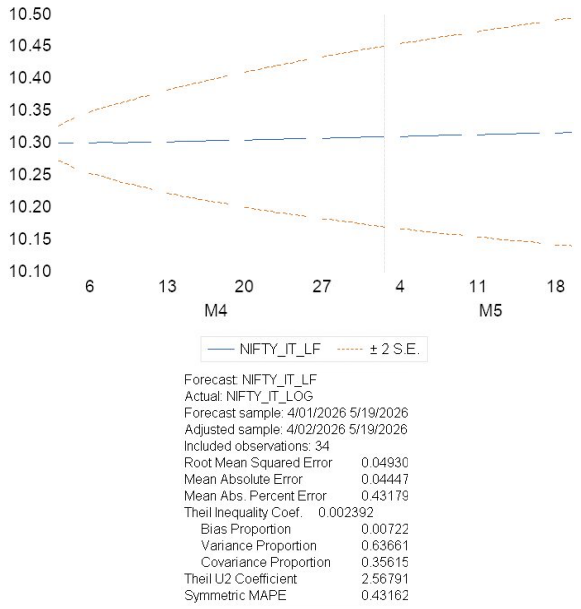

Figure 5.4.1 — AR(7)+MA(5) Dynamic Forecast: NIFTY_IT_LOG (Apr–May 2026). Near-flat trajectory with expanding ±2 S.E. bands.

A direct MAPE comparison between ARIMA and VAR is not reported, as the VAR in-sample metrics (Section 5.1) and ARIMA out-of-sample metrics reflect fundamentally different evaluation conditions — a comparison that would systematically favour in-sample fit. Instead, Theil U2 = 2.567 serves as the standalone benchmark, comparing ARIMA against a naïve random walk on the same out-of-sample window. A Theil U2 substantially above unity confirms the AR(7)+MA(5) model offers no improvement over a random walk over the test period, while the VAR's cross-market structure — US Tech Granger-causing Nifty IT at p = 0.000 with FEVD explaining 11.3% of variance at Day 10 — provides directional signal that no univariate model can replicate.

6. Mid-Cap vs Large-Cap Spread: The Investable Finding

The most actionable result is the persistent and accelerating outperformance of mid-cap IT over large-cap IT since the ChatGPT launch.

Metric	Value	Note
Spread at ChatGPT launch (Nov 2022)	~1.00	MIDCAP − LARGECAP log-spread
Spread at May 2026	**~1.75**	Continued widening post-ChatGPT
Post-ChatGPT incremental widening	~0.75	Over 3.5 years
Annualised outperformance (post-ChatGPT)	**~21%/yr**	**Mid-cap over large-cap**
Cumulative outperformance (2018–2026)	~90 log pts	Full sample

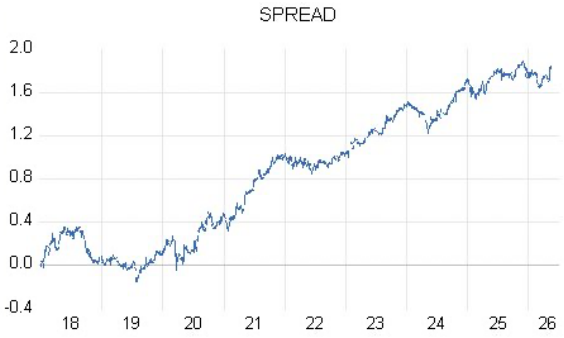

Figure 4: Mid-cap vs Large-cap IT Log-Spread (MIDCAP − LARGECAP), January 2018 – May 2026. Spread ≈0 pre-ChatGPT; ≈1.75 by May 2026.

The spread's consistent upward trend since 2020, accelerating post-ChatGPT, rules out mean-reversion as an explanation. At ~21% annualised outperformance over 3.5 years, the mid-cap advantage has been persistent, not episodic — consistent with a structural demand shift toward agile, AI-native service providers rather than a cyclical rotation.

The business logic is clear: large-cap IT (TCS, Infosys, Wipro) face headwinds from AI-driven automation of headcount-based delivery. Mid-cap names (Persistent Systems, Coforge) have benefited from AI-native demand — digital transformation, GCC buildouts, cloud migration — and pivoted faster to AI-enabled service delivery.

7. Investment Implications and Limitations

7.1 Key Implications

- US tech momentum is no longer a reliable leading indicator for Nifty IT positioning. Factor models anchored in US tech co-movement will generate alpha-eroding lag effects post-ChatGPT.
- Within IT, tilt intra-sector allocation toward mid-caps. The 21% annualised outperformance reflects a durable structural shift, not a mean-reverting anomaly.
- Monitor NVDA as a tactical regime indicator (TAR threshold: NVDA_LOG = 1.668). When above the threshold, short-run US Tech → Nifty IT transmission remains active.
- Do not fade elevated volatility quickly. GARCH persistence of 0.959 implies shocks take weeks to dissipate.

7.2 Limitations

- US Tech Proxy includes NVDA (semiconductor), whose 2,000%+ post-ChatGPT rally may distort what is conceptually an IT services benchmark. A cleaner proxy (MSFT+GOOGL, or Accenture + Cognizant) is recommended for robustness.
- Low R^2 values (VAR: 0.064–0.116). Models characterize structure and direction, not magnitude, with precision.
- Mid-cap spread strategy lacks a formal backtest with transaction costs and liquidity constraints. PERSISTENT and COFORGE have meaningful impact costs at institutional size.
- Post-ChatGPT sample (~479 obs, 3.5 years) is relatively short for robust long-run cointegration inference.

8. Investment Recommendation

Scenario	View	Position	Trigger
Base Case	AI transition creates winners & losers within IT	Long mid-cap, reduce large-cap weight	NVDA above threshold; mid-cap spread expanding
Bull Case	GCC buildout + AI-native demand accelerates	Overweight mid-cap; add on dips	Deal wins accelerating; spread > 2.0
Bear Case	Global recession reduces IT spending sharply	Reduce sector weight; hedge via USDINR	US tech FEVD contribution rises above 20%

The base case favours a long mid-cap / short large-cap spread trade, supported by the ~21% annualised outperformance trend and the NVDA regime indicator. The bull case adds to mid-cap exposure as AI implementation demand — GCC buildouts, cloud migration, enterprise AI adoption — accelerates beyond automation headwinds. The bear case exits the spread entirely and reduces sector weight, as a global growth shock would compress IT budgets indiscriminately. Across all scenarios, NVDA log-level 1.668 serves as the tactical entry signal for sector exposure.

9. Conclusion

This study provides formal econometric evidence that the ChatGPT launch constituted a structural break in the Nifty IT – US Tech relationship. Pre-November 2022, the two series were cointegrated with rank one, error-correcting at approximately 1% per day, with US Tech as the clear structural leader. Post-November 2022, cointegration is absent and the short-run VAR transmission explains only 11% of Nifty IT daily return variance. The AI era did not merely add noise to an existing relationship — it rendered the pre-existing framework non-operative.

This is not a story of Indian IT destruction. The GARCH analysis finds no structural increase in volatility post-ChatGPT, and the out-of-sample forecast shows Nifty IT outperforming its pre-AI structural trajectory by approximately 24%. The TAR analysis confirms that US Tech transmission was regime-stable throughout — not amplified by AI capex cycles. The sector decoupled from its historical anchor without deteriorating.

The clearest investable conclusion is the mid-cap versus large-cap divergence. At ~21% annualised outperformance since ChatGPT, the spread between Persistent/Coforge and TCS/Infosys/Wipro has widened persistently — reflecting a structural shift in where AI-era demand accrues, not a cyclical rotation. Large-caps face headwinds from automation of headcount-based delivery; mid-caps are benefiting from AI-native service demand, GCC buildouts, and digital transformation.

For equity desk positioning: retire US tech as a Nifty IT leading indicator; tilt intra-sector allocation toward mid-caps using the NVDA threshold as a tactical entry signal; and do not fade elevated volatility quickly. The AI era did not destroy Indian IT — it bifurcated it, and the data tell us clearly which side to be on.

References

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–328.

Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Dawar, I., Dutta, A., Bouri, E., & Saeed, T. (2021). Crude oil prices and clean energy stock indices: Lagged and asymmetric effects with quantile regression. Renewable Energy, 163, 288–299.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. Econometrica, 55(2), 251–276.

Hansen, B. E. (2000). Sample splitting and threshold estimation. Econometrica, 68(3), 575–603.

Jawadi, F., Idi Cheffou, A., & Bu, R. (2023). Oil price shocks and macro-financial variables: A TAR analysis. Energy Economics, 127, 107058.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica, 59(6), 1551–1580.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. Oxford Bulletin of Economics and Statistics, 52(2), 169–210.

MacKinnon, J. G., Haug, A. A., & Michelis, L. (1999). Numerical distribution functions of likelihood ratio tests for cointegration. Journal of Applied Econometrics, 14(5), 563–577.

Rostan, P., Rostan, A., & Nurunnabi, M. (2020). Options trading strategy based on ARIMA forecasting. PSU Research Review, 4(1), 56–69.

Sims, C. A. (1980). Macroeconomics and reality. Econometrica, 48(1), 1–48.