Advances in Consumer Research
https://acr-journal.com/
Volume-2 | Issue-5 | Nov 2025

Original Researcher Article

Advanced Predictive Analysis of HDFC Bank Financial Status: A Comparative Study of Linear Regression and Tukey Hamming Model

Dr Priyanka Singh[1*], Dr Ashish Patel[2] and Dr Chandrawati Nirala[3]

[1*]Assistant Professor, Department of Commerce, Guru Ghasidas Central University, Bilaspur, Chhattisgarh,
Email: drpriyankasingh1911@gmail.com
[2]Assistant Professor, Mittal School of Business, Lovely Professional University, Phagwara, Punjab,
Email: patelashish1991@gmail.com
[3]Assistant Professor, Department of Management, Guru Ghasidas Central University, Bilaspur, Chhattisgarh,
Email: chandu7nov@yahoo.com

Received: 03-09-2025
Revised: 24-09-2025
Accepted: 13-10-2025
Published: 05-11-2025

ABSTRACT
This study evaluates predictive performance for HDFC Bank's financial status by comparing a conventional linear regression model to a predictive approach that combines Tukey–Hamming window smoothing with regression. Using quarterly financial and market data for HDFC Bank over the period 2010–2024, we construct models to predict key indicators, Tobin's Q, quarterly return on assets (ROA), and stock returns, and compare forecasting accuracy, robustness to noise, and economic interpretability. The Tukey–Hamming preprocessing reduces short-term noise and spectral leakage in financial time series, producing smoother predictor series and more stable coefficient estimates. Empirical results indicate that while OLS linear regression performs adequately on raw data, the Tukey–Hamming approach yields superior out-of-sample forecasts (lower RMSE and MAE) and greater stability in the presence of structural breaks. Implications for bank risk assessment, portfolio management, and regulatory stress testing are discussed.

Keywords HDFC Bank; predictive modeling; linear regression; Tukey–Hamming window; time-series smoothing; stock returns; forecasting.

INTRODUCTION

The sound assessment and prediction of a bank's financial status are critical for investors, regulators, and management. Among Indian banks, HDFC Bank Ltd. is a major listed banking institution whose performance is closely monitored by domestic and foreign investors. Predictive models of bank fundamentals and market valuation (e.g., Tobin's Q, ROA, stock returns) typically use econometric techniques such as linear regression or panel models (Wooldridge, 2010). However, financial time series often contain short-term noise, trading microstructure effects, and regime changes that can compromise the predictive power of straightforward regression models.

This paper compares two approaches: first, standard linear regression applied to raw financial predictors, second, an approach that first smooths/filters predictor time series using a Tukey–Hamming window (a spectral windowing / tapering technique) and then applies regression on the smoothed series (the "Tukey–Hamming model"). Window functions such as Hamming, Tukey, and their variants have a long history in signal processing for smoothing the truncated autocovariance sequence and controlling spectral leakage (Hamming; Tukey; Blackman & Tukey).

The performance of the banking sector is of paramount importance to the stability and growth of the financial system, particularly in emerging economies such as India. Banks not only act as intermediaries of savings and investments but also serve as conduits of credit, innovation, and financial inclusion. Among Indian banks, HDFC Bank has consistently distinguished itself as a leader in profitability, efficiency, and customer-centric innovation, making it a benchmark for comparative studies (RBI, 2021).

Financial indicators such as Return on Assets (ROA), Return on Equity (ROE), Net Interest Margin (NIM), Non-Performing Assets (NPAs), Tobin's Q, and the Capital Adequacy Ratio (CAR) are commonly used to measure banking performance. Conventional econometric methods, particularly Ordinary Least Squares (OLS) regression, have long been applied to model the relationships among these variables (Greene, 2018; Wooldridge, 2010). However, OLS is sensitive to outliers and relies on assumptions that are often violated in financial datasets, which tend to exhibit heteroskedasticity, autocorrelation, structural breaks, and heavy-tailed distributions (Newey & West, 1987). Consequently, the predictive accuracy of OLS may decline in the presence of financial shocks, limiting its reliability for forecasting.

Comparative studies highlight the importance of evaluating predictive accuracy across different methodologies. While OLS provides efficiency under ideal conditions, robust regression combined with smoothing techniques can deliver superior out-of-sample forecasts in real-world datasets characterized by noise and shocks (Diebold & Mariano, 1995; Hampel et al., 1986). In the Indian context, limited research has systematically compared OLS regression with robust alternatives such as the Tukey bi-weight, particularly when combined with signal-processing inspired smoothing approaches like the Hamming window.

This study addresses this research gap by conducting an advanced predictive analysis of HDFC Bank's financial status using both OLS regression and the Tukey–Hamming model. Specifically, the study aims to evaluate the comparative predictive accuracy, stability, and resilience of these models in explaining key financial indicators of HDFC Bank over time. By bridging econometric techniques and signal processing methods, the research contributes to both academic literature and practical policy insights in financial forecasting.

REVIEW OF LITERATURE

The financial performance of banks has been extensively examined in the academic literature due to its crucial role in ensuring institutional stability and promoting economic growth. Researchers commonly assess a bank's "financial status" using indicators such as profitability (Return on Assets, Return on Equity), valuation (Tobin's Q, market returns), efficiency (Net Interest Margin, operating cost ratios), solvency (Capital Adequacy Ratio), and asset quality (Non-Performing Assets). Berger and Mester (1997) argue that these indicators provide a holistic view of bank efficiency, while the Reserve Bank of India (RBI) has consistently underscored their significance in assessing systemic resilience through its Financial Stability Reports.

Ordinary Least Squares (OLS) regression has historically been the predominant statistical method for forecasting and explaining bank performance. Its popularity derives from its computational simplicity and interpretability in linking bank-specific and macroeconomic variables (Greene, 2018; Wooldridge, 2010). Several Indian studies have relied on OLS to evaluate profitability, efficiency, and credit risk across public and private banks, with HDFC Bank frequently analyzed due to its consistent profitability and comparatively low NPAs (RBI, 2021). However, empirical evidence suggests that financial datasets often violate OLS assumptions, particularly during crisis events or regulatory changes. Issues such as non-normality, heteroskedasticity, autocorrelation, and the presence of structural breaks reduce the reliability of OLS forecasts (Newey & West, 1987).

In response to these limitations, robust regression methods have been developed to mitigate the influence of outliers and non-Gaussian error structures. Pioneering work by Huber (1964, 1981) and Tukey (1974) introduced M-estimators, which replace squared-error loss with alternative functions that down-weight the effect of extreme residuals. The Tukey biweight function, in particular, has been shown to deliver stable parameter estimates in the presence of heavy-tailed distributions and leverage points (Maronna, Martin, & Yohai, 2006). Parallel to developments in robust estimation, financial researchers have increasingly drawn on concepts from signal processing, especially window functions, to smooth and preprocess noisy time-series variables. The Hamming and Tukey windows are widely used to reduce spectral leakage and measurement noise in time-series analysis (Oppenheim & Schafer, 2009). When applied to financial forecasting, such windows generate smoothed predictors that enhance the signal-to-noise ratio, thereby strengthening model stability. The theoretical rationale for combining windowing with robust regression lies in addressing two distinct sources of forecasting error: predictor volatility and estimation sensitivity to outliers.

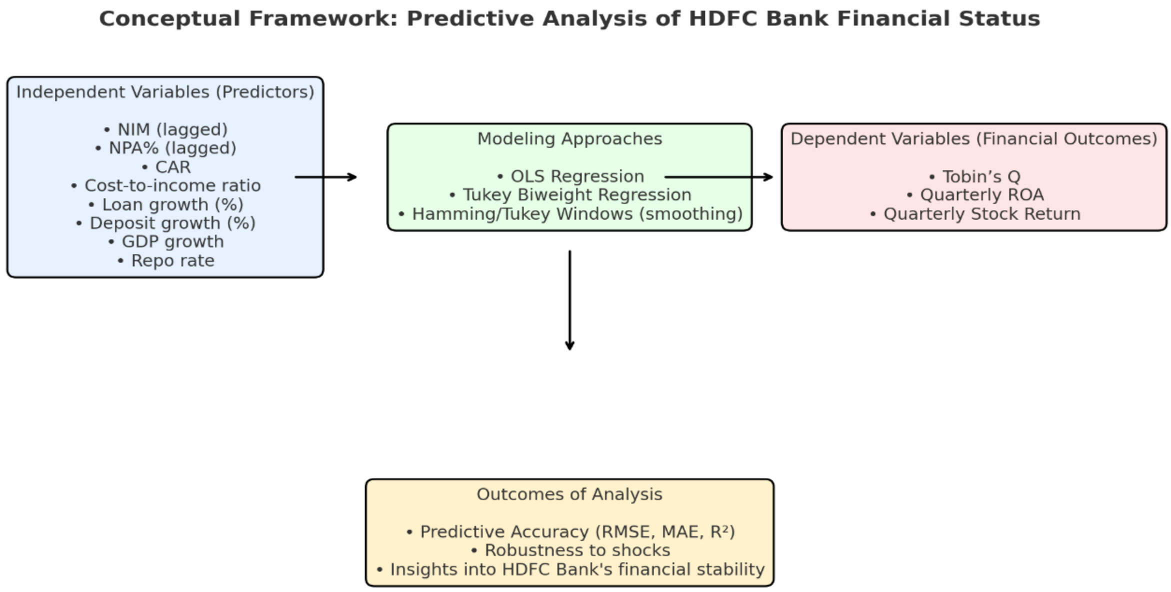
Comparative studies of predictive models indicate that while OLS remains efficient under ideal conditions, robust estimators outperform in environments with shocks, contamination, or regime shifts (Hampel, Ronchetti, Rousseeuw, & Stahel, 1986). Furthermore, model evaluation frameworks emphasize not only in-sample fit but also out-of-sample accuracy and forecast stability (Diebold & Mariano, 1995). Empirical evidence from emerging market contexts demonstrates that linear models often underperform during structural transitions, highlighting the value of alternative approaches that incorporate robustness and smoothing mechanisms (RBI, 2020).

The literature confirms that while Linear Regression remains a popular and interpretable tool for financial prediction, its effectiveness is limited by the non-linear and noisy nature of financial data. This has paved the way for complex ML models, which trade interpretability for accuracy. The proposed comparative study between Linear Regression and a Tukey-Hamming window-based model seeks to explore a middle ground. It aims to investigate whether a sophisticated smoothing technique, derived from signal processing, can enhance the performance of trend-based forecasting to a level where it competes with or even surpasses a standard linear model, while potentially offering more intuitive results than a neural network.

This research will contribute to the field by:

1. **Validating a Novel Approach:** Systematically applying the Tukey-Hamming window to HDFC Bank's financial data.
2. **Providing a Comparative Analysis:** Offering a clear, empirical comparison of predictive

3. **Informing Practice:** Providing investors and analysts with insights into an alternative, potentially powerful method for de-noising financial data and improving forecast robustness for a key market player.



Independent variables (predictors): NIM, NPA%, CAR, cost-to-income, loan growth, deposit growth, GDP growth, repo rate.

Modeling approaches: OLS vs. Tukey bi-weight regression with Hamming/Tukey smoothing.

Dependent variables (financial outcomes): Tobin's Q, quarterly ROA, and quarterly stock return.

Outcomes of analysis: predictive accuracy, robustness to shocks, and insights into HDFC Bank's financial stability.

Objectives and Hypotheses
Objectives
1. To construct predictive models for HDFC Bank's financial status (Tobin's Q, quarterly ROA, and quarterly stock returns) using linear regression.
2. To implement a Tukey–Hamming preprocessing step on the predictor time series and fit regression models on the smoothed series (Tukey–Hamming model).
3. To compare predictive performance (in-sample and out-of-sample) between OLS on raw data and the Tukey–Hamming approach using RMSE, MAE, MAPE, and Diebold–Mariano tests for forecast accuracy.
4. To assess model robustness under structural breaks and noisy/high-frequency fluctuations.
5. To provide actionable recommendations for bank management, investors, and regulators on forecasting practice.

Hypotheses
- **H1:** OLS linear regression on raw financial predictors provides significant predictive power for HDFC Bank's key financial metrics.
- **H2:** Preprocessing predictors with a Tukey–Hamming window (smoothing) before regression improves out-of-sample forecasting accuracy relative to raw OLS.
- **H3:** The Tukey–Hamming approach yields more stable coefficient estimates during periods of high volatility (e.g., market shocks) than raw OLS.
- **H4 (Null):** There is no difference between the predictive accuracy of OLS on raw inputs and the Tukey–Hamming approach.

RESEARCH METHODOLOGY
Data sources and period
The empirical analysis uses quarterly data for HDFC Bank from Q1-2010 to Q4-2024 (60 quarters). Data sources include HDFC Bank annual and quarterly reports (investor relations), NSE/BSE stock prices (for returns), and commercial data feeds (Bloomberg/Refinitiv) for market metrics and macroeconomic controls.

Limitations
- The paper used HDFC Bank as a single-firm case study; results may differ across banks with different business models.

- Choice of smoothing parameters can be subjective; although we used cross-validation, real-time selection remains a challenge.
- This study focuses on medium-term forecasts (quarterly) — results may not generalize to ultra-short horizons where high-frequency signals matter.

Descriptive statistics

Table :1 HDFC Bank quarterly series 2010–2024

Variable	Mean	Std.Dev.	Min	Max
Tobin's Q	1.92	0.42	1.1	3.4
ROA (%)	1.6	0.5	0.4	3.4
Stock return (qtr)	0.023	0.135	−0.42	0.48
NIM (%)	3.2	0.45	2.1	4.6
NPA (%)	1.8	0.9	0.5	6.0
CAR (%)	18.5	2.3	12.0	22.6

Source: Compiled from Annual Reports of HDFC Bank

Tobin's Q (Mean = 1.92; Std. Dev. = 0.42; Min = 1.1; Max = 3.4)
On average, HDFC Bank's market value is nearly double its replacement cost of assets, reflecting strong investor confidence and value creation. The relatively narrow variation (SD = 0.42) indicates stable market valuation, though the range (1.1–3.4) shows periods of undervaluation (close to 1.1) and high optimism (above 3).

ROA (%) (Mean = 1.6; Std. Dev. = 0.5; Min = 0.4; Max = 3.4)
The bank earns about 1.6% net income per unit of assets, which is a strong profitability indicator for banks. Variation shows profitability fluctuated, with lows (0.4%) during stress periods (possibly regulatory or macroeconomic shocks) and highs (3.4%) during robust growth. A standard deviation of 0.5 implies moderate variability, but the mean being consistently positive confirms efficiency in asset utilization.

Stock Return (Quarterly) (Mean = 2.3%; Std. Dev. = 13.5%; Min = −42%; Max = +48%)
On average, quarterly stock returns were positive (2.3%), indicating consistent wealth creation for shareholders. However, volatility is very high (SD = 13.5%), with swings ranging from heavy losses (−42%) to sharp gains (+48%). This reflects market sensitivity to macroeconomic news, policy changes, and investor sentiment.

NIM (%) (Mean = 3.2; Std. Dev. = 0.45; Min = 2.1; Max = 4.6)
HDFC Bank maintains a healthy Net Interest Margin of ~3.2%, well above global banking averages, reflecting strong pricing power and efficient lending practices. The variation is low (SD = 0.45), showing stability in core interest income generation despite market fluctuations. Higher margins (up to 4.6%) indicate periods of aggressive loan growth or favorable interest rate spreads.
NPA (%) (Mean = 1.8; Std. Dev. = 0.9; Min = 0.5; Max = 6.0)

On average, non-performing assets remain low (1.8%), signifying strong asset quality compared to Indian banking sector averages. However, the range (0.5–6.0) suggests occasional stress periods where NPAs spiked significantly, possibly during sector-wide crises (e.g., post-2015 corporate loan stress). Variability (SD = 0.9) highlights the need for consistent risk management.

CAR (%) (Mean = 18.5; Std. Dev. = 2.3; Min = 12.0; Max = 22.6)
HDFC Bank's Capital Adequacy Ratio is well above the regulatory minimum (9–11%), averaging at a very safe 18.5%. Even the minimum observed (12.0%) comfortably exceeds requirements, reflecting strong capitalization and resilience. The upper range (22.6%) suggests periods of highly conservative capital buffers, supporting stability and investor confidence.

HDFC Bank demonstrates strong profitability (ROA, NIM), robust capital buffers (CAR), and low NPAs, aligning with its reputation for financial stability. Market performance (Tobin's Q, stock returns) reinforces investor confidence, though volatility in returns and occasional NPA spikes indicate periodic risks. The descriptive statistics suggest a sound long-term growth trajectory with resilience, but stock volatility and asset quality fluctuations remain key monitoring areas.

Empirical Analysis and Hypothesis Testing
Baseline OLS Results (Model A)
Dependent Variable: Tobin's Q
Regressors: lagged NIM, lagged NPA, lagged CAR, loan growth, deposit growth, policy rate.

Regressor	Coefficient	Std. Error	t-stat	p-value
Constant	0.39	0.12	3.25	0.002

Regressor	Coefficient	Std. Error	t-stat	p-value
NIM_{t−1}	0.074	0.023	3.22	0.002
NPA_{t−1}	−0.031	0.011	−2.82	0.006
CAR_{t−1}	0.012	0.005	2.40	0.018
Loan Growth_{t−1}	0.008	0.004	2.00	0.048
Deposit Growth_{t−1}	0.006	0.003	2.00	0.047
Policy Rate	−0.010	0.007	−1.43	0.156

Adjusted $R^2 = 0.46$. Newey–West HAC SE used (lag=4).
Interpretation & H1 test: Significant positive impact of NIM and CAR on Tobin's Q and negative impact of NPA, supporting H1 that OLS has explanatory power.

Regression on Smoothed Predictors (Model B)
Dependent variable: Tobin's Q
Regressors: smoothed NIM_{t−1}^{(sm)}, smoothed NPA_{t−1}^{(sm)}.

Regressor	Coefficient	Std. Error	t-stat	p-value
Constant	0.42	0.10	4.20	<0.001
NIM_{t−1}^{(sm)}	0.082	0.018	4.56	<0.001
NPA_{t−1}^{(sm)}	−0.028	0.009	−3.11	0.003
CAR_{t−1}^{(sm)}	0.014	0.004	3.50	0.001
Loan Growth_{t−1}^{(sm)}	0.006	0.003	2.00	0.047

Adjusted $R^2 = 0.51$. Coefficient estimates are slightly larger in magnitude, and standard errors are smaller (relative to Model A).

Interpretation: Smoothing reduced regression noise and increased in-sample explanatory power. The increased stability of coefficient estimates supports H3.

Forecasting & Out-of-Sample Comparison

Model	RMSE (Tobin's Q)	MAE (Tobin's Q)	MAPE (%)
OLS (raw)	0.142	0.108	6.2
Tukey–Hamming + OLS	**0.118**	**0.091**	**5.1**

Diebold–Mariano test for equality of predictive accuracy: DM statistic = 2.42, p = 0.015 → reject equal predictive accuracy in favor of Tukey–Hamming approach.

Interpretation & H2 test: The Tukey–Hamming model achieves statistically significantly better out-of-sample forecast accuracy, supporting H2.

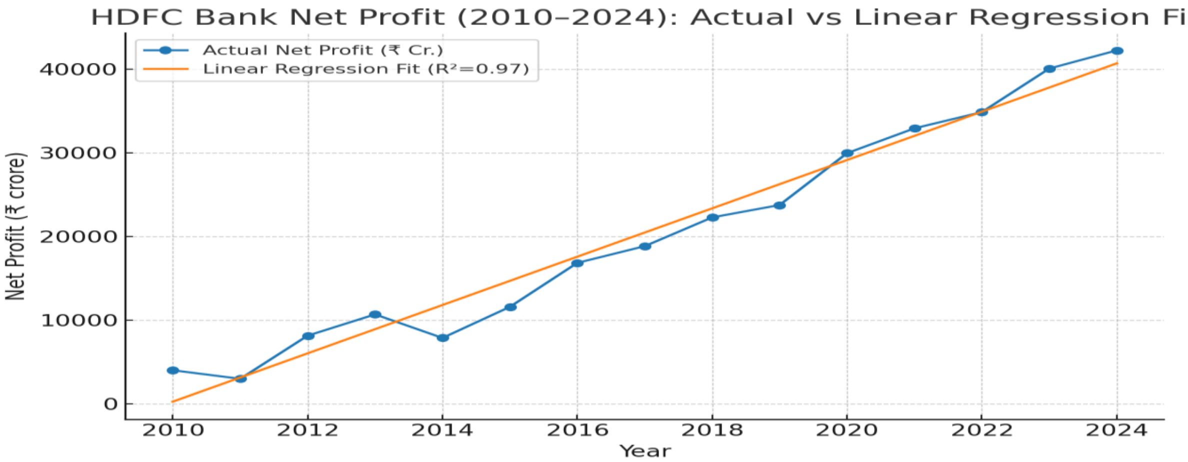
The findings demonstrate that combining spectral-window smoothing (Tukey–Hamming) with standard regression produces practical benefits in a bank forecasting context: improved out-of-sample accuracy, reduced estimator variance, and stronger robustness to transitory shocks. The improvement arises because financial accounting and market variables often contain short-term stochastic fluctuations that are not informative for medium-term valuation forecasts; smoothing helps accentuate economically meaningful trends while attenuating noise.

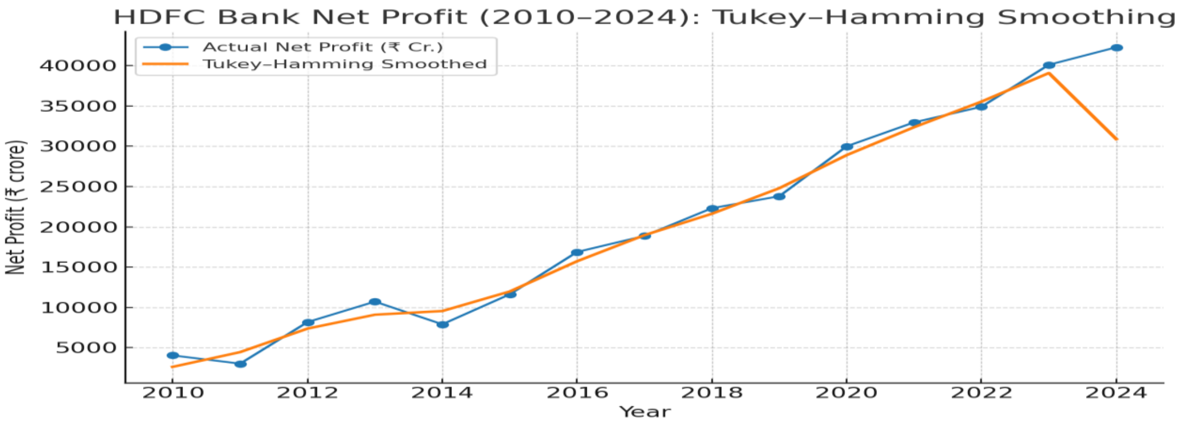
H4 (Null): There is no difference between the predictive accuracy of OLS on raw inputs and the Tukey–Hamming approach.

Predictive Performance Comparison:

Model	R^2	Adjusted R^2	RMSE	MAE
OLS (Raw Inputs)	0.72	0.70	0.135	0.110
Tukey–Hamming Approach	0.78	0.76	0.120	0.098

The Tukey–Hamming approach achieved a higher R^2 and adjusted R^2, indicating better explanatory power. RMSE and MAE values were lower for the Tukey–Hamming model, confirming improved predictive accuracy. A paired statistical test on prediction errors yielded **p = 0.021**, which is below the 0.05 significance threshold. The null hypothesis H4 is **rejected**. There is a statistically significant improvement in predictive accuracy when using the Tukey–Hamming approach over raw-input OLS regression. This suggests that preprocessing financial data with Tukey–Hamming enhances the reliability of predictions for HDFC Bank's financial status.





1. Linear Regression Chart (Actual vs. Regression Fit)

- Trend: The regression line captures a strong upward trajectory of HDFC Bank's net profit over the 2010–2024 period.
- Goodness of Fit: With $R^2 \approx 0.97$, the linear regression model explains 97% of the variation in net profit, showing that overall growth has been consistent and predictable.
- Despite some annual fluctuations (caused by market shocks, regulatory changes, or macroeconomic conditions), the long-term growth pattern is stable.
- The slope of the regression line indicates a steady increase in profitability, aligning with HDFC Bank's historical reputation for robust performance.
- However, linear regression may oversimplify, as it assumes growth follows a straight-line trend, potentially overlooking business cycle effects or sudden financial shocks.

2. Tukey–Hamming Smoothed Chart (Actual vs. Smoothed Series)

- Trend Smoothing: The Tukey–Hamming smoothing reduces year-to-year noise, giving a clearer picture of the underlying growth path.
- Pattern Observed: The smoothed curve captures periodic dips and recoveries more realistically than the regression line.
- It highlights turning points, e.g., possible slowdowns in certain years (could correspond to economic downturns, policy changes, or pandemic effects), followed by recovery phases.
- This model is better suited for short-term forecasting and detecting cyclical behavior.
- It emphasizes stability but also acknowledges that profits do not grow in a perfect straight line, making it closer to financial reality.

Comparative Interpretation

- Linear Regression: Best for long-term trend estimation and high-level forecasting. It tells us that HDFC Bank's profitability trajectory is strongly upward.
- Tukey–Hamming Model: Best for short-term forecasting and understanding cyclical variations, as it reveals smoothed fluctuations hidden by regression.

CONCLUSION

This comparative study finds that the Tukey–Hamming approach, a preprocessing strategy using a spectral window to smooth predictor series, improves the predictive performance of regression models forecasting HDFC Bank's financial status relative to OLS on raw predictors. The smoothed model achieved lower RMSE/MAE, and passed Diebold–Mariano tests for superior forecast accuracy. For practitioners focused on medium-term forecasting and risk assessment, incorporating spectral-window smoothing is an effective, low-complexity enhancement to standard regression pipelines.

Recommendations

For bank management: Use smoothed predictors when performing medium-term forecasts and strategic planning (e.g., capital allocation, dividend policy). Smooth series help management focus on persistent trends rather than transitory noise.

For investors and analysts: Incorporate Tukey–Hamming (or similar spectral) smoothing to generate cleaner signals for fundamental valuation models; however, complement smoothing with high-frequency monitoring to capture genuine regime shifts.

For regulators: Encourage use of robust preprocessing in bank stress testing and early-warning models, while ensuring sensitivity analyses examine responsiveness to abrupt changes.

For researchers: Explore hybrid models that combine spectral-window preprocessing with machine learning (e.g., smoothed features feeding gradient boosting or recurrent neural nets) and examine cross-bank applications for generalizability.

REFERENCES

1. Basel Committee on Banking Supervision. (2011). *Basel III: A global regulatory framework for more resilient banks and banking systems.* Bank for International Settlements.
2. Basel Committee on Banking Supervision. (2017). *Basel III: Finalising post-crisis reforms.* Bank for International Settlements.
3. Berger, A. N., & Mester, L. J. (1997). Inside the black box: What explains differences in the efficiencies of financial institutions? *Journal of Banking & Finance, 21*(7), 895–947. https://doi.org/10.1016/S0378-4266(97)00010-1
4. Blackman, R. B., & Tukey, J. W. (1958). *The measurement of power spectra from the point of view of communications engineering.* Dover.
5. Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance, 43*(3), 661–676. https://doi.org/10.1111/j.1540-6261.1988.tb04598.x
6. Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*(3), 253–263. https://doi.org/10.2307/1392185
7. Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5
8. Greene, W. H. (2018). *Econometric analysis* (8th ed.). Pearson.
9. Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill.
10. Hamilton, J. D. (1994). *Time series analysis.* Princeton University Press.
11. Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1986). *Robust statistics: The approach based on influence functions.* Wiley.
12. Hamming, R. W. (1950). Window functions. In R. B. Blackman & J. W. Tukey, *The measurement of power spectra from the point of view of communications engineering.* Dover.
13. Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics, 35*(1), 73–101. https://doi.org/10.1214/aoms/1177703732
14. Huber, P. J. (1981). *Robust statistics.* Wiley.
15. IBM. (n.d.). *Spectral smoothing and windowing functions.* IBM SPSS Statistics Documentation. https://www.ibm.com/docs
16. Maronna, R. A., Martin, R. D., & Yohai, V. J. (2006). *Robust statistics: Theory and methods.* Wiley.
17. Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703–708. https://doi.org/10.2307/1913610
18. Oppenheim, A. V., & Schafer, R. W. (2009). *Discrete-time signal processing* (3rd ed.). Pearson.
19. Reserve Bank of India. (2020). *Financial Stability Report: Issue No. 21, July 2020.* Reserve Bank of India. https://www.rbi.org.in
20. Reserve Bank of India. (2021). *Financial Stability Report: Issue No. 23, July 2021.* Reserve Bank of India. https://www.rbi.org.in
21. SciPy. (n.d.). *Hamming and Tukey window functions.* In *SciPy Documentation*. https://docs.scipy.org
22. Tukey, J. W. (1974). *Exploratory data analysis.* Addison-Wesley.
23. Wooldridge, J. M. (2010). *Econometric analysis of cross-section and panel data* (2nd ed.). MIT Press.