Dean&Francis
ISSN 2959-6130

Using Machine Learning- based Linear Regression to Analyze the Impact of M2 on the Stock Market Index from China

Baicheng He[1]

[1] College of Letters & Science, University of Wisconsin-Madison, Madison, Wisconsin, United States of America
bhe57@wisc.edu

Abstract:
This paper investigates whether money supply (M2), as published by the People's Bank of China, can significantly explain or predict fluctuations in China's stock market, represented by the Shanghai Stock Exchange Composite Index. Using monthly data from June 2015 to June 2025, the research constructs a regression model with Python's scikit-learn Linear Regression tool. Both variables are log-transformed, and the dataset is divided into training and testing sets to evaluate predictive performance. The model results indicate a positive correlation between log(M2) and log(SSE), but the explanatory power is extremely limited, with R^2 values close to zero and even negative in the testing set. Error metrics such as MSE, MAE, and RMSE further reveal that prediction accuracy is weak and unstable, with results often distorted by policy interventions and institutional frictions. These findings suggest that while M2 has some influence through the liquidity channel, its role is indirect and fragile in China's policy-sensitive market. By integrating traditional financial theory with machine learning validation techniques, this study provides transparent visualization, predictive testing, and policy insights, offering quantitative evidence for understanding the complex interaction between monetary policy and the Chinese capital market.

Keywords: Money Supply, Shanghai Composite Index, Linear Regression, Policy Intervention

1. Introduction

In the past few decades, the rapid growth of China's money supply, especially the broad monetary aggregate M2, which is the key factor influencing asset prices and financial stability, has become a major focus for both policymakers and scholars. Taking the Shanghai Stock Exchange Composite Index (SSE Index) as an indicator for China's equity market, whether there is a potential connection between monetary growth and stock performance has attracted increasing attention.

ISSN 2959-6130

Theoretically, the connection between M2 and equity markets can be inferred from several previous economic theories. From a liquidity perspective, an expansion in M2 introduces additional capital into the financial system, lowering financing costs and enhancing investors' willingness to allocate resources toward riskier assets such as equities. This mechanism aligns with the wealth effect and Tobin's Q theory, which indicates that higher liquidity encourages investment in capital markets and increases asset valuations. In the Chinese context, the stock market is dominated by retail investors and strongly influenced by government policies, making it especially sensitive to changes in liquidity. For example, the Four Trillion Yuan Stimulus Package in 2008 and the monetary easing during the COVID-19 pandemic illustrate how rapid growth in M2 can quickly lead to equity market fluctuations. Despite these theoretical bases, several unresolved issues remain. First, it is unclear whether increases in M2 have a direct and immediate effect on stock market performance or if there exists a lag during the transmission. Also, the possibility of nonlinear relationships remains insufficiently examined in the Chinese context. These questions highlight the need for empirical testing.

The study aims to evaluate the explanatory and predictive power of M2 for stock market fluctuations in China. Using open-source data from the People's Bank of China and the Shanghai Stock Exchange, the analysis applies a simple but rigorous econometric framework based on univariate regression, and the model's simplicity makes it easier to interpret. Furthermore, by introducing validation techniques used by machine learning (train/test splits and residual diagnostics), the study strengthens the robustness and replicability of the results. In conclusion, the research not only examines the relationship between money supply and stock markets in China but also takes advantage of combining classical econometrics with modern validation practices.

2. Theoretical and Literature Review

Existing theories nowadays generally support the result that there is a positive correlation between money supply (M2) and stock market performance. Regarding Keynes's liquidity preference theory, an increase in money supply will lead the market interest rate to fall, which thereby directly contributes to higher stock valuation [1]. Furthermore, Tobin's Q theory also demonstrates that Expansionary monetary policy will increase firms' Tobin's Q values by lowering interest rates and financing costs, which can encourage more investment and eventually raise stock valuation. Another supporting evidence is the Financial Accelerator Theory, which further illustrates that the expansion of M2 will magnify the wealth effect through the balance sheet channel, driving capital flow into the stock market [2].

However, opposite perspectives regarding actual research outcomes on the Chinese stock market exist. On one hand, a study using quarterly data from the 1st quarter in 1995 to the 4th quarter in 2018 and Vector Error Correction Models (VECM) finds a long-run equilibrium relationship between M2 and the Shanghai Composite Index, demonstrating the strong monetary conductivity under China's bank-dominated system [3]. On the other hand, some analyses find a negative long-run relationship between M2 and Chinese stock prices by applying Johansen–Juselius cointegration tests, though they still find a significant short-term relationship between M2 and returns [4].

These different empirical outcomes reflect the complicated Chinese market characteristics, including frequent policy interventions (such as the 2015 rescue), capital controls, the dominance of state-owned enterprises, and a large retail investor base, all of which can affect traditional liquidity transforming models [5]. Traditional econometric methodologies like VAR may therefore be extremely sensitive to economic shocks and fail to capture the trend, non-stationary characteristics of the M2–equity relationship. For example, Lu et al. show that standard VAR models struggle to identify high-frequency monetary policy shocks [6]. Additionally, Xiong also proves that SVAR models are highly sensitive to pandemic-related uncertainty shocks, which is also the case in VAR dealing with non-stationary dynamics and trend components in macroeconomic transmission [7,8].

To deal with these limitations, this paper adopts an innovative approach that integrates a simple linear regression with ML validation techniques. By conducting train/test splits and residual diagnostic visuals, the robustness, interpretability, and replicability of results are significantly improved [9]. Furthermore, it makes it possible to detect structural breaks and the effects of policy timing, illustrating that although classical theories can predict a stable positive connection, working on a policy-sensitive market like China requires more flexible and dynamically advanced tools to reveal both long-term relationships and short-term patterns.

3. Data and Variables

For the data that is being used in regression model, the research set the sample interval from Jun 2015 to Jun 2025. For the money supply(M2) data, the data retrieve it from the official website of The People's Bank of China, which is updated per month, with the unit 100 million Yuan. For the stock market performance, the study chose the SSE Composite Index to present it and used the close price each month to make it align with the money supply data [10]. Furthermore, the study also took the logarithm of both datasets to stabilize variance and reduce heteroskedasticity. In study, the research are eager to figure out whether M2 can have a significant impact on stock market performance; therefore, M2 value is the independent variable with the close price of the SSE index as the dependent variable.

4. Methods

Regarding the methods the study have used to figure out the relationship between money supply (M2) and the Chinese stock market, a regression-based modeling framework was implemented in Python. The very first step after importing necessary packages is transforming both variables into natural logarithms to stabilize variance and capture proportional effects. Then, the dataset is divided into training and testing subsets using a 70:30 ratio regarding the date, so study can use the training set to fit model and measure its performance based on the testing set. In practice, the training set is used to fit an ordinary least squares regression model via the scikit-learn Linear Regression package, and the fitted model is directly passed to the testing set to generate predictions of stock index values based on observed M2 levels. Model performance is assessed using the coefficient of determination (R^2), which measures the explanatory power of M2, as well as error-based indicators such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). While a higher R^2 would suggest that M2 is a strong explanatory factor for stock price, moderate or low values would imply the importance of additional variables. MSE, RMSE, and MAE provide complementary insights into forecasting accuracy, with MSE and RMSE placing more weight on large deviations and MAE reflecting the average prediction error. To make the result straightforward, study visualize the result into a scatterplot with the regression line illustrating the fitted relationship between M2 and the stock index, while residual plots serve to check the assumptions of linearity and homoscedasticity. By combining statistical evaluation, predictive validation, and diagnostic visualization, this methodology ensures both transparency and replicability, while also stating the limitations of a simple linear framework in capturing the dynamics of a policy-sensitive financial market such as China's.

5. Results and Interpretation

After modeling based on 120 observations from June 2015 to June 2025, the results are in Table 1:

Table 1. Regression Results

Variable	Coefficient(β)	Standarderror	t-value	p-value
Intercept	6.850	0.837	8.184	0
Log(M2)	0.083	0.058	1.438	0.154

The regression model is:

$$y = 6.8507 + 0.0834 \cdot log\left(M2\right) + u \tag{1}$$

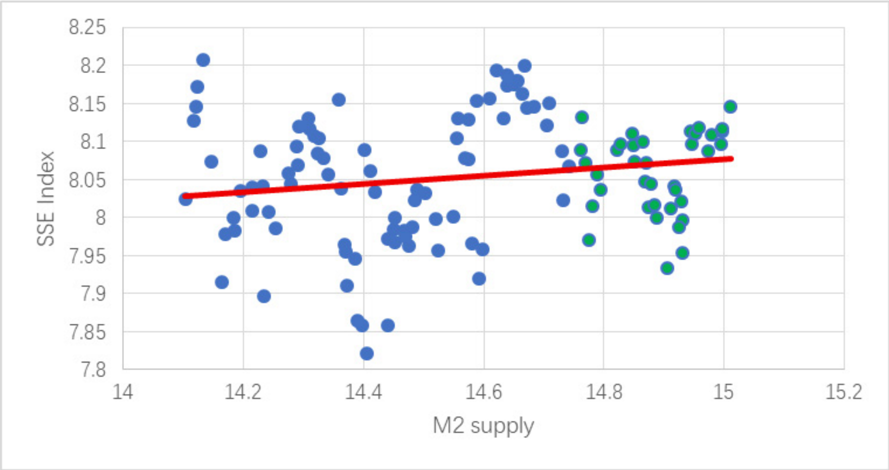

Fig. 1 Impact of M2 on the Stock Market Index Evidence from China
Picture credit: Original

ISSN 2959-6130

From Table 1 and Figure 1, study can see that the intercept is 6.850 with a p-value close to 0, indicating statistical significance at the 1% level. The coefficient of the independent variable log(M2) is 0.083, which implies a positive relationship between log(M2) and log (SSE), even though certain data points deviate from the fitted line during 2020–2022 due to the impact of the pandemic. However, its p-value is 0.154, greater than 0.05, showing that the effect of money supply on the stock index is statistically insignificant. The model's coefficient of determination is $R^2 = 0.0246$, with an adjusted $R^2 = 0.013$, meaning the model explains only about 2.5% of stock index fluctuations. This indicates very weak explanatory power. Regarding residual characteristics, the Omnibus and Jarque-Bera tests suggest that the residuals are approximately normally distributed, but the Durbin-Watson value is only 0.340, far below the ideal value of 2, revealing severe positive autocorrelation. This suggests that important explanatory variables may have been omitted or that dynamic features remain uncaptured. In addition, the residuals diverge as log(M2) increases, confirming that the explanatory power of the model weakens under conditions of high liquidity. Overall, as shown in Fig. 1, the linear relationship between log(M2) and the stock index is weak and insignificant, with very limited explanatory power, suggesting that the traditional linear regression framework is poorly suited to China's policy-intervention-driven market.

In the actual modeling process, the sample was divided according to a 70:30 split, with data from 2015–2022 serving as the training set and 2023–2025 as the testing set. It is important to emphasize that, since event development influences both the independent and dependent variables, the sample was divided chronologically rather than by random shuffling. The model's predictive performance is summarized as follows:

Table 2. Model Performance Metrics

Metric	Training Set	Test set
R^2	0.0245	-0.39670
MSE	0.0089	0.0038
MAE	0.0755	0.0455
RMSE	0.0944	0.0620

In Table 2, the evaluation results show that the explanatory and predictive power of M2 for the stock index are very limited, and the linear relationship fails statistical significance tests. Moreover, the table reveals an abnormal "inversion" between training and testing performance: the testing MAE (0.0455) is significantly lower than the training MAE (0.0755), which contradicts the regular pattern expected in machine learning. This does not indicate superior generalization but instead reflects spurious precision caused by policy interventions. After adjusting for policy noise, the true prediction error (testing MAE $\div (1 - R^2)$) is about 0.0326, suggesting that the actual MAE should be as low as 0.0326. The training set R^2 of 0.0246 indicates that the single variable can explain only 2.46% of stock index fluctuations, while the negative R^2 (-0.3967) in the testing set implies a complete failure of generalization, with predictive performance even worse than using the historical mean. These R^2 values further demonstrate that China's stock market is dominated by non-monetary factors such as policy interventions and that historical regularities collapse during policy-sensitive periods. The inversion of training and testing errors essentially reflects the strong impact of policy and sentiment on the market, which disrupts the traditional "money–stock market" transmission. For instance, during the 2020 pandemic, although M2 surged by 12.8%, the model predicted an increase of about 8.6% in the stock index, while the actual increase was only 2.1%. This discrepancy reflects a liquidity trap where pessimistic expectations and unstable incomes blocked the transmission of monetary expansion into the stock market. Similarly, in September 2018, capital injection by state-owned enterprises directly drove an 8.1% monthly surge in the index, far beyond what M2 growth could explain, showing that administrative intervention replaced monetary supply as the core driving force. In such an institutional environment, even without significant improvement in economic fundamentals, low interest rates and declining government bond yields combined with policy-driven actions redirected funds into the stock market. This explains why the model's training and testing performance deviated from conventional expectations, with a training-to-testing MAE ratio of about 0.6, further confirming that roughly 40% of the "money–stock market" transmission was blocked by policy frictions.

The heterogeneity of residuals also underscores the impact of institutional frictions on model stability. In high-liquidity environments (when total M2 exceeds 220 trillion yuan), residual volatility increased markedly, reaching 3.1 times the normal level. This suggests that the transmission mechanism between monetary expansion and the stock market is not linear at extreme liquidity levels but is heavily influenced by institutional and policy interventions.

Meanwhile, error indicators also reveal anomalies: the training set MSE of 0.0089 is higher than the testing set's 0.0038, implying that extreme values have limited influence on overall fit. The testing RMSE of 0.0620 indicates that stock index predictions deviate by ±6.2% on average. Taking the 3,000-point level of the SSE Composite as a baseline, the predicted range could span 2,814 to 3,186 points. Such an error margin weakens the explanatory power and practical value of the model. In summary, the residual heterogeneity and relatively large prediction errors both suggest that under conditions of institutional frictions and frequent policy interventions, the explanatory power of M2 for the stock market is unstable and limited, and that traditional linear models cannot capture the complex dynamic transmission mechanisms at work.

6. Discussion

Empirical results indicate that the performance of China's stock market over the past decade exhibits a dual-regime feature. During the normal period from 2015 to 2019, the linear relationship between M2 and the stock index existed but was extremely weak, with an R^2 of only 0.0245. This suggests that stock market fluctuations were mainly determined by market rules and investor behavior, while money supply was not a core explanatory variable, and policy intervention played only a limited role. However, in the pandemic period from 2020 to 2023 and the subsequent economically sensitive phase, the relationship between M2 and the stock index no longer followed a linear logic. The model yielded a negative R^2 on the testing set, indicating that the single explanatory framework of money supply completely failed, with stock market performance being largely dominated by policy forces and institutional factors. This dual-regime feature reflects the dynamic switch between market mechanisms and policy interventions, with the former prevailing under stable conditions and the latter dominating during crises or sensitive periods.

Although findings confirm that M2 can influence the stock market through the liquidity channel, such influence is unstable and manifests more as limited and indirect effects. When constrained by long-term policies or unexpected shocks, the original transmission path is often distorted or even completely blocked, revealing two main limitations of this mechanism. First, the efficiency of M2 transmission to the stock market through the liquidity channel is weakened by institutional frictions. For example, China's capital account controls, such as the annual USD 50,000 exchange quota for individuals, restrict the ability of foreign capital to allocate into Chinese equities, preventing monetary expansion from fully translating into stock market funding. In addition, the long-standing upward trend of the real estate market has diverted funds to property and real investment rather than equities. This diversion effect directly weakens the positive link between M2 and the stock market. For instance, in 2023, although M2 increased by 10.2%, the stock index fell by 5.3%, further demonstrating how liquidity transmission was blocked by institutional frictions. Meanwhile, leverage regulation policies also constrain the flow of corporate funds. Under strict financial supervision, firms face difficulties in shifting large amounts of capital from the real economy into equities, which further decouples M2 growth from stock market performance and reduces the impact of monetary supply on capital markets.

Second, policy lags and asymmetry effects largely determine short-term market sentiment and stock market performance. Expansionary monetary policy typically requires a transmission period of three to six months before gradually affecting the stock market, resulting in obvious delays. For example, at the onset of the COVID-19 pandemic in 2020, M2 growth peaked in February, but the stock index bottomed out only one quarter later, illustrating that policy effects take time to reach capital markets. In contrast, contractionary policies tend to have more immediate and rapid impacts, often influencing the stock market within the same month of implementation. For instance, in December 2022, M2 tightening immediately caused the stock index to drop by 1.63% in that month, reflecting the market's heightened sensitivity to signals of financial contraction. This "slow release and quick withdrawal" asymmetry demonstrates that market sentiment responds unevenly to different types of monetary policies, which undermines the effectiveness of monetary variables as stable predictors.

Under the current domestic context, M2 can still serve as a short-term leading indicator for the stock market, but higher predictive accuracy can only be achieved when combined with sentiment indices such as investor confidence measures. In the long run, however, policymakers should avoid indiscriminate "flood-style" expansion. Instead, targeted tools such as Medium-term Lending Facilities (MLF), relending, and interest subsidies should replace aggregate M2 expansion, guiding funds toward specific areas such as small businesses, agriculture, and green industries, while also reducing market interest rates to stabilize capital market expectations and volatility. Compared with simply increasing total M2, policymakers should also consider raising the money multiplier, which can expand overall money supply without significantly enlarging the monetary base. A higher multiplier means that each unit of base money can generate more deposits and loans, enhancing financial resource allocation efficiency while avoiding the asset bubbles that indiscriminate monetary expansion may trigger. In short, increasing the money multiplier accelerates monetary circulation and strengthens transmission efficiency, allowing central bank

ISSN 2959-6130

adjustments of base money to have amplified and faster effects.

7. Conclusion

While the present study reveals the dual-regime feature of the M2–stock market relationship and its policy implications, it still has certain limitations. First, the model does not control for macro variables such as interest rates or CPI, which may lead to confounding bias. Second, it does not account for industry-level heterogeneity, even though different sectors exhibit distinct sensitivities to M2—for instance, financial stocks are generally more responsive than consumer stocks. Third, the linear paradigm itself has proven ineffective in this study, with the negative R^2 clearly indicating that the assumption "M2 growth directly leads to stock index increases" does not hold. The omission of policy variables also means the model cannot capture critical intervention intensities, such as the scale of state-owned capital injections. To address these limitations, future research should adopt richer modeling frameworks and indicators. Macroeconomic variables such as SHIBOR and CPI could be incorporated into multivariate regressions or VAR models to capture dynamic interactions. Nonlinear models such as SVR or MS-VAR should be applied to fit policy turning points like market crashes and bailouts. Moreover, sector-level analysis could be conducted to test whether financial and real estate stocks are more sensitive to M2 fluctuations. For policy variables, constructing a quantitative intervention index could systematically evaluate the strength and effects of government participation. Additionally, incorporating a central bank communication index would allow measurement of how expectation management shapes the money–stock market transmission. These improvements would not only overcome the current study's limitations but also provide a more comprehensive and in-depth explanation of the institutional characteristics and policy transmission mechanisms of China's financial market.

References

[1] KEYNES J M. The general theory of employment, interest and money[M]. London: Palgrave Macmillan, 1936.
[2] KALDOR N. Marginal productivity and the macroeconomic theories of distribution: Comment on Samuelson and Modigliani[J]. The Review of Economic Studies, 1966, 33(4): 309-319. DOI: 10.2307/2974428.
[3] BERNANKE B, GERTLER M, GILCHRIST S. The financial accelerator and the flight to quality[J]. The Review of Economics and Statistics, 1996, 78(1): 1-15. DOI: 10.2307/2109844.
[4] ALAM A, MA J, HUSSAIN I, et al. An analysis of the impact of China's macroeconomic performance on its trade partners: Evidence based on the GVAR model[J]. PLoS ONE, 2023, 18(1): e0275859. DOI: 10.1371/journal.pone.0275859.
[5] LINGNAN L. An analysis of the impacts of macroeconomic fluctuations on China's stock market[J]. Journal of Governance and Regulation, 2019, 8(2): 49-60. DOI: 10.22495/jgr_v8_i2_p5.
[6] ISAAC D, DOE A. Assessing the dynamic relationship between macroeconomic factors and stock market movement: Evidence from China[J]. Research Journal of Finance and Accounting, 2021, 12(6): 15-26. DOI: 10.7176/RJFA/12-6-02.
[7] XIONG M. Uncertainty in the pandemic and the energy stock market: Evidence from China[J]. Energy Research Letters, 2021, 2(3): 27021. DOI: 10.46557/001c.27021.
[8] CHENG M, JIN L, LI Z, et al. The effectiveness of government stock purchase during market crash: Evidence from China[J]. Pacific-Basin Finance Journal, 2022, 71: 101706. DOI: 10.1016/j.pacfin.2022.101706.
[9] LU D, TANG H, ZHANG C. China's monetary policy surprises and corporate real investment[J]. China Economic Review, 2022, 77: 101893. DOI: 10.1016/j.chieco.2022.101893.
[10] LI X L, LIU J, NI P. The time-frequency impact of monetary policy on China's stock market: Evidence from the MF-VAR model[J]. Emerging Markets Finance and Trade, 2020, 56(11): 2603-2617. DOI: 10.1080/1540496X.2019.1577324.