

Predicting implied volatility surface with a Large Number of Factors and Machine Learning

Myeongsu Choi

Department of Finance, Hanyang University Business School, 222 Wangsimni-ro,
Seongdong-gu, Seoul, Korea 04763; +82 2 787 7349; kidsjjang@hanyang.ac.kr

Sol Kim

College of Business, Hankuk University of Foreign Studies, 107 Imun-ro, Dongdaemun-gu,
Seoul, Korea 02450; +82 2 2173 3124; solkim@hufs.ac.kr

Predicting implied volatility surface with a Large Number of Factors and Machine Learning

Abstract

This study introduces a novel approach to estimating the implied volatility (IV) surface by incorporating macro-financial variables, extending beyond traditional strike and time-to-maturity based models. Utilizing the Fama-Macbeth regression and machine learning techniques, the model integrates a comprehensive set of macro-financial data to enhance predictive accuracy. We evaluate the model's performance during the 2007-2010 financial crisis, demonstrating its robustness and reliability under extreme market conditions. The findings highlight the importance of macro-financial factors in financial modeling, providing valuable insights for risk management and financial forecasting.

1. Introduction

This paper introduces an innovative model for estimating implied volatility (IV) surfaces that advances beyond the confines of traditional financial models, which are often critiqued for their lack of economic interpretability and their reliance on static parameters such as strike prices and time-to-maturity. The implied volatility of options serves as an indicator of market participants' expectations and outlook, providing insights into the future volatility of the underlying asset prices. It plays a crucial role in option pricing, enabling market participants to predict future volatility and assess related risks. Moreover, the implied volatility surface visualizes the structure of volatility in the options market across various strike prices and maturities, offering a nuanced understanding of the microstructure of the options market and providing valuable economic significance for option premium valuation and portfolio risk management. Implied volatility surface plays a pivotal role in option pricing models, particularly when examined through the lenses of volatility smile and volatility term structure. The volatility smile, a phenomenon where implied volatilities tend to increase as options move out-of-the-money or in-the-money, underscores the market's anticipation of extreme price movements. This asymmetry in implied volatilities suggests that investors are willing to pay a premium for options that protect against adverse price movements, reflecting their perception of risk and uncertainty. Understanding and interpreting the volatility smile is crucial for pricing exotic options accurately and gauging market sentiment regarding potential future price fluctuations. On the other hand, the volatility term structure provides insights into the relationship between implied volatilities and the time to maturity of options. By observing how implied volatilities vary across different expiration dates, analysts can discern market expectations regarding future volatility levels. A steep upward-sloping term structure implies anticipation of increased volatility in the future, while a flat or downward-sloping structure may signal stability or declining volatility expectations. This information is invaluable for constructing trading strategies, managing risk, and assessing market conditions. In essence, the implied volatility surface encapsulates the collective wisdom of market participants, reflecting their views on risk, uncertainty, and future market dynamics. Analyzing this surface through the prisms of volatility smile and volatility term structure enhances our understanding of option pricing dynamics and facilitates more informed decision-making in financial markets. The research question at the heart of this study aims to provide economic interpretations of IV surface fluctuations through the integration of a comprehensive set of macrofinancial factors,

whilst maintaining or even enhancing the out-of-sample predictive power relative to currently popular models (Cont & Fonseca, 2002) This inquiry is crucial in the realm of financial modeling, where understanding the economic determinants of volatility fluctuations is essential for making informed investment choices and for effective risk management (Mixon, 2002).

The integration of an extensive array of macrofinancial factors into the estimation of volatility, however, introduces substantial challenges, including increased model complexity, demands on computational efficiency, and the need for clear interpretability (Gu et al., 2020). To navigate these challenges, this paper proposes innovative empirical methodologies that harness advanced machine learning techniques and generative algorithms.

Our model extends the traditional analytical framework by incorporating economic indicators known to influence market volatility, thus providing a more robust framework for IV predictions. By assimilating macro-financial data, our model not only captures volatility variability more accurately but also aligns more closely with actual market behaviors, thereby enhancing the efficacy of pricing and risk management strategies. Notably, this model is the inaugural attempt in the literature to articulate economic interpretations of IV surface variations grounded in a broad spectrum of macrofinancial factors and their nonlinear interactions. This represents a significant step forward, allowing for a more nuanced understanding of how diverse economic forces interplay and impact volatility dynamics.

Methodologically, we employ the empirical rigor of the Fama-Macbeth regression (Fama & MacBeth, 1973) and leverage machine learning techniques to adeptly select and incorporate these macro-financial variables into our volatility estimation framework. This approach not only elevates the precision and predictive capacity of the IV estimates but also deepens our understanding of the underlying economic factors that drive market movements. Our model is pioneering in its capability to incorporate a substantial number of macrofinancial factors through the use of generative algorithms and to exploit machine learning algorithms based on augmented data and complex factor relationships. These innovations facilitate the efficient handling of high-dimensional data and the elucidation of intricate, nonlinear relationships between macrofinancial factors and implied volatility surfaces.

To rigorously evaluate the effectiveness of our model, we assess its performance during the 2007-2010 global financial crisis—a period characterized by significant economic upheaval and market instability. This challenging historical context serves as a critical test for the model's robustness and its ability to perform under conditions of heightened volatility and stress. The

model's success in providing accurate and interpretable IV estimates during this tumultuous period underscores its practical relevance and superiority over existing models.

The findings from this research have significant implications for the field of financial modeling. By addressing critical gaps in traditional methods of volatility estimation, our model paves the way for more informed and strategic decision-making in financial markets. It serves as a methodological bridge between theoretical finance and practical market applications, thereby enriching the toolkit available to practitioners for navigating complex market environments. Additionally, the insights gained from our model's economic interpretations of IV surface changes can assist policymakers, regulators, and market participants in better understanding and responding to the underlying drivers of market volatility.

In conclusion, this paper presents a groundbreaking model for estimating implied volatility surfaces that effectively incorporates a vast array of macrofinancial factors and offers nuanced economic interpretations of volatility fluctuations, while maintaining robust out-of-sample prediction capabilities. Through the utilization of advanced machine learning techniques and generative algorithms, our model adeptly addresses the challenges posed by high-dimensional data and complex market dynamics. Its performance during the 2007-2010 global financial crisis not only demonstrates its robustness but also establishes its potential for real-world application. This research contributes significantly to the field of financial modeling by providing a powerful tool for understanding and navigating the intricate relationships between macroeconomic factors and market volatility, ultimately leading to more informed and strategic decision-making in financial markets. Our work opens new avenues for future research, inviting further exploration of the interplay between macrofinancial factors and implied volatility surfaces, as well as the development of even more sophisticated models that can adapt to the continually evolving landscape of financial markets.

2. Literature review

The accurate estimation of the implied volatility surface is paramount for option pricing and risk management. The existing research aimed at estimating the implied volatility surface of options is being conducted along two main avenues. Firstly, there are efforts to expand the assumptions underlying the Black-Scholes option pricing model to better reflect reality, in order to avoid discrepancies with observed implied volatility surfaces. Studies in this direction seek to adjust the assumptions of the Black-Scholes option pricing model to ensure that the

estimated implied volatility varies consistently with moneyness and time-to-maturity. Such inconsistencies suggest a misalignment between the assumptions of the Black-Scholes model and empirical observations, which is interpreted as evidence that the model is inadequate in describing the options market accurately. Consequently, there is an endeavor to amend the assumptions used in the Black-Scholes model to better align with reality.

Previous studies have tried to improve the usefulness of the option pricing model and solve the volatility smile and surface phenomena by additionally reflecting reality in the basic assumptions of the BS model. Stochastic interest rate model (Amin & Jarrow, 1991, 1992; Ho et al., 1997; Rindell, 1995), stochastic volatility model (Hull & White, 1987; Johnson & Shanno, 1987; Melino & Turnbull, 1990; Rindell, 1995; Scott, 1987; Stein & Stein, 1991; Wiggins, 1987), jump diffusion model (Merton, 1976; Naik, 1993), variance gamma model (Madan et al., 1998), regime switching model (Kou, 2002; Kou & Wang, 2004; Naik & Lee, 1990) that assumes a sharp jump in volatility, the model that assumes generalized autoregressive conditional heteroscedasticity (GARCH) (Duan, 1995; Duan & Zhang, 2001; Heston & Nandi, 2000; Hsieh & Ritchken, 2005) as the underlying asset process are examples. Bakshi et al. (1997, 2000), Bates (2003), Kim & Kim (2004, 2005) have tried to find out which option pricing model is most useful for pricing and hedging options. They show that the stochastic volatility model most greatly improves the disadvantages of the BS model and is the best for pricing and hedging options. Furthermore, it has been confirmed that stochastic interest rate factors have significant explanatory power in the long-term options market. Utilizing such models reveals that certain portions of the implied volatility surface vanish. Secondly, there are studies aimed at acknowledging the existence of the implied volatility surface, refining explanations of the current volatility surface, and predicting future volatility surfaces. There is an Ad-Hoc Black Scholes (AHBS) model in which regression analysis is used to estimate the implied volatility surface of the BS model. This is called the Traders' rule, mainly used by practitioners in the options market. The AHBS model is implemented assuming that the implied volatility for each option is related to the exercise price and time-to-maturity. In other words, it is assumed that the volatility of the underlying assets entered into the BS model is a function of the exercise price and time-to-maturity or a combination thereof. Recent studies (Y. Choi et al., 2012; Y. Choi & Ok, 2012; Dixit & Singh, 2018; Dumas et al., 1998; Jackwerth & Rubinstein, 2012; S. Kim, 2009, 2014, 2017; Li & Pearson, 2004, 2007) show that the pricing and hedging performance of the AHBS model is superior to mathematically sophisticated models such as the stochastic volatility model. The AHBS model, ultimately, is a model akin

to conventional statistical models, which aims to understand the characteristics of time series itself to explain and predict the present and future volatility surface. The AHBS model employs a form of linear interpolation. Linear extrapolation is also applied to estimate implied volatility beyond the upper and lower bounds of traded option prices (Jiang & Tian, 2005; Lee & Ryu, 2024). Also there are the semi-parametric spline interpolation (Fengler, 2009; Figlewski, 2008), a non-parametric kernel regression (Aït-Sahalia & Lo, 1998; OptionMetrics, n.d.; Ulrich & Walther, 2020) and the parametric Gram–Charlier expansion (Beber & Brandt, 2006).

Up to now, efforts to explain and predict the volatility surface using control variables outside the options market, such as macroeconomic variables, have not existed. At this juncture, the value of this study can be appreciated.

3. Methodology

3.1. Empirical methods

In this section, we describe our empirical strategy step by step.

Step 1: Construct a surface based on moneyness and time-to-maturity using daily S&P 500 Index (SPX) data.

Step 2: Download 127 monthly macro-financial data from FRED. Then, use the *tcode* from (McCracken & Ng, 2016) for data transformation. Change the unbalanced panel into balanced panel data following the method described in Appendix I.

Step 3: Use the obtained 127 macro-financial variables for another PCA.

Step 4: Using the final data, produce innovation terms. Start with v_{t+1} , a K-dimensional vector of state variables, i.e., macro-financial PCA variables at time $t + 1$. Our simplified equation is as follows although one can use machine learning to produce innovation terms instead:

$$v_{t+1} = \gamma \cdot v_t + \alpha.$$

Then, generate dv_{t+1} and $dz_{v,t}$, which denotes the source of risk, using the following equations:

$$dv_{t+1} = v_{t+1} - (\hat{\gamma}v_t + \hat{\alpha}).$$

$$dz_{t+1} \equiv \Sigma_t^{-0.5} dv_t \sim N(0, I)\sqrt{dt}.$$

Σ_t is the covariance matrix of dv_t , and becomes similar to I when dv_t is obtained as a result of conducting PCA on the time-series data.

Step 5: Collect the sigma and beta (the regression coefficients) using the following equations.

$$\sigma_{IV,n,t}^2 \equiv std(dIV_{n,t})/dt.$$

$$\vec{\beta}_{n,t} dt \equiv cov(dz_{v,t}, dIV_{n,t})/\sigma_{IV,n,t}^2.$$

Step 6: Formulate a regression equation (1), which is rearranged as equation (2) as follows:

$$(1) df_{n,t}\Delta = -\lambda'_t dz_{v,t} + \sigma_{f,n,t}^2 \left(\frac{\Delta}{2} + \lambda'_t \vec{\beta}_{n,t} \right) \Delta dt + \sigma_n \Delta dz_{n,t}$$

$$(2) df_{n,t} = constant + \lambda'_t (\sigma_{f,n,t}^2 \vec{\beta}_{n,t} dt - dz_{v,t}/\Delta) + \epsilon_{n,t}$$

$$constant = \sigma_{f,n,t}^2 \Delta dt / 2$$

λ_t denotes the price of risks and is optionally modeled as a linear function of risks, v_t .

Step 7: Use equation (2) to obtain λ_t for each t in Fama-MacBeth regression (Fama & MacBeth, 1973) (possibly Lasso regression when selecting macro-financial factors). One can model λ_t as an affine function of risks and estimate it.

The empirical method presented has several distinct advantages over traditional models that use the strike and time to maturity (TTM). Firstly, the application of the Fama-Macbeth regression technique offers a more intuitive and straightforward estimation process by linearizing the price kernel. This approach has been demonstrated to have superior predictive performance and computational efficiency when compared with other multivariate volatility

models (Ku et al., 2011; Nyberg & Vilhelmsson, 2008). Additionally, our method's compatibility with machine learning techniques enhances its application, allowing for more robust and intuitive use in both the preliminary and final stages of volatility estimation (Frank & Yang, 2021).

Furthermore, the incorporation of 127 observable macro-financial factors significantly aids in the accurate estimation of the volatility surface. This comprehensive factor inclusion facilitates a deeper understanding and more precise forecasting of market behaviors (Green & Thomas, 2019). As a result, this empirical method not only simplifies the estimation process but also enhances the predictive accuracy and applicability of the volatility surface analysis, marking a significant improvement over classic models.

3.2. Data

The SPX (S&P 500 index) option data used in this paper are from the IVY DB database of OptionMetrics LLC. The data include end-of-day bid and ask quotes, implied volatilities, open interest, and daily trading volume for the SPX options traded on the Chicago Board Options Exchange from January 4, 1996 through August 31, 2015. We examine SPX option contracts based on several advantages. First, it is easy to compare our results with those of the existing papers. The S&P 500 option is one of the leading advanced options markets that several previous studies have analyzed. Implications can be easily derived by comparing our results with those of previous studies. Second, the options with various time-to-maturities are traded. Although the trading volumes in emerging options markets are plentiful, they are concentrated in the short-term options. A fully traded options market with a variety of maturities is required to conduct our study on the volatility surface. Third, the S&P 500 option is a European option. A number of option pricing models have been derived to price European options. Evaluating the American options requires additional tweaking.

The price of the option is calculated as the average of the best bid and ask quotes at the end of each trading day. We use all out-of-the-money (OTM)¹ calls and puts options that satisfy the following criteria. (Bakshi et al., 1997; Y. Choi et al., 2012; Y. Choi & Ok, 2012; I. J. Kim & Kim, 2004, 2005). Any duplicates or missing values in bid or ask quote, strike price or time-

¹ In-the-money (ITM) options are not considered because its trading volumes are significantly low and thus any information regarding ITM can be doubtful. Moreover, it may cause possible duplicates from double counting since ITM calls and OTM puts are equivalent for a specified strike price according to the put-call parity.

to-maturity are excluded. Options with less than 7 days are removed to avoid any liquidity bias. The prices lower than $3/8$ are excluded to control for price discreteness. Finally, prices that do not satisfy the no-arbitrage conditions are eliminated. In other words, prices outside the theoretical upper and lower bounds of option prices are not considered.

The option data used in this analysis is based on monthly figures. This methodical approach ensures consistency in data collection and processing, facilitating a more reliable assessment of option price behaviors over an extended period. Such a dataset provides a robust foundation for examining market trends and verifying theoretical models under different market conditions. This monthly analysis helps in capturing broader market movements and understanding the impact of macroeconomic variables on option prices more effectively.

For macro-financial variables, we collect monthly macro-financial data from the Federal Reserve Economic Data (FRED). We use *tcode* to transform the data before generating macro-financial factors on McCracken and Ng (2016); more specifically, due to frequent missing values, we use the five-step procedure to balance the unbalanced panel of macro-financial variables (see Appendix I for details).

In line with the approach demonstrated by Ludvigson and Ng (2009), we commence our analysis by employing a pool of 127 macro-financial variables to collect comprehensive macro-financial data. We utilize Principal Component Analysis (PCA) to derive 115 principal components from these variables, effectively reducing the dimensionality of our dataset while retaining significant information. The complete list of these variables is provided in Appendix II. The strategic exploitation of these observable variables under no-arbitrage conditions constitutes one of the primary contributions of this research. PCA, as applied in our study, aligns with established methods in the literature, where it is recognized for its capacity to simplify the analysis and interpretation of complex datasets by identifying orthogonal directions of maximum variance (Armeanu & Lache, 2008). This approach facilitates a robust framework for the predictive modeling and analysis of financial markets.

Table 1 compares the summary statistics on the model-generated and observed $dlog(IV)$. The sample period is from 1996:2 to 2015:8.

###Insert Table 1 about here###

The findings presented in Table 1 demonstrate that expanding the number of factors from 5 (comprising Strike, TTM, Strike², TTM², and Strike*TTM) to 127 enhances the fit of the

implied volatility (IV) surface. The statistical metrics of the model-implied surfaces, using 127 factors, closely match the observed returns. For instance, the average observed $d\log(IV)$ ranges between -2.93 and -0.19, aligning with the average model-implied IV. Similarly, the variability in IV, quantified by the standard deviation across different maturities and moneyness, varies from -2.93 and -0.19, consistent across both observed and model-implied IVs. Figures 1 and 2 further illustrate these results, showing a perfect overlap between the observed and model-implied IVs.

What if incorporating a machine-learning model which our IV surface can utilize easily? Machine learning can make the 127-factor model simpler, more intuitive and more powerful. For instance, one can employ simple machine learning to select relevant combinations of macro-financial variables. Lasso regression selects only relevant factors while zeroing out irrelevant ones. Autoencoder selects important nonlinear combinations of the factors. We illustrate that Lasso regression identifies an intuitive set of 19 macro-financial variables out of 127 macro-financial variables or their interactions while not compromising the fit. A discussion of the results is included in subsection 4.3.

4. Empirical analysis

4.1. IV Surface fitting

To evaluate the accuracy of the implied volatility (IV) surface derived from our model, we conduct time series and cross-sectional analyses covering the period from January 1996 to August 2015. The data set comprises implied volatility values from S&P 500 options throughout this timeframe. In the visual representations, solid lines denote the observed implied volatilities, while dashed lines, which are almost indistinguishable from the solid lines, represent the IVs implied by the model. To eliminate any forward-looking bias, macroeconomic variables are lagged by one month.

###Insert Figure 1 about here###

Figure 1 displays four graphs corresponding to different penalties, showing results based on $\log(IV)$ observations with 10, 34, 65, and 115 macro factors. In each graph, the close

alignment between the observed data and the model outputs suggests nearly perfect congruence, affirming the model's robustness in replicating the actual IVs landscape.

###Insert Figure 2 here###

Figure 2 illustrates the outcomes of cross-sectional regressions across various moneyness levels. This figure highlights the cross-sectional fit of the implied volatility, modeled with 10, 34, 64, and 115 macro factors. The graph on the left displays the unconditional mean, while the graph on the right depicts the unconditional standard deviation of both the observed and model-implied implied volatilities.

4.2. Forecasting IV surface using macro-financial factors

It is imperative to evaluate whether our proposed model can forecast shifts in the volatility curve during out-of-sample periods. To this end, we employ the subsequent regression analysis to determine if our model can effectively predict future changes in implied volatility:

- 1) without control variables

$$r_{i,t+1} = constant + \beta_i G_{i,t} + \epsilon_{i,t+1}.$$

- 2) with control variables

$$r_{i,t+1} = constant + \beta_i G_{i,t} + (controls) + \epsilon_{i,t+1}.$$

The independent variable ($G_{i,t}$) is the difference between the observed and the model-implied $\log(IV)$ at each moneyness level i and time period t with its coefficient denoted as β . The dependent variable is the difference between the $\log(IV)$ at time period t and $t + 1$, which moneyness level is i . The difference between the IVs at time period t and $t + 1$ is used as a control variable 1 and the coefficient is denoted *control 1*. The $\log(IV)$ at time period $t - 1$ and moneyness level i is used as a control variable 2 and the coefficient is denoted *control 2*. The sample data is constructed using S&P500 implied volatility data. The sample period is from 1996:1 to 2015:8.

###Insert Table 2 here###

Table 2 presents the regression outcomes using 10, 34, 64, and 115 factors. The model's predictive capability varies with the options' moneyness and time-to-maturity; however, performance generally enhances as the number of factors increases. This improvement suggests a more detailed capture of the volatility surface dynamics with an increasing number of explanatory factors.

###Insert Figure 3 here####

Figure 3 displays the t-values in absolute values obtained from the same regression analysis. The difference between Figure 3 and Table 2 is that Table 2 reports the results of the regression performed using the control variables, while Figure 2 displays the results both with and without the control variables. It can be seen from Figure 2 that the proposed model has significant predictive power for $d\log(IV)$. The statistical significance of the predictive power of the model is almost the same whether we include or exclude the control variables, and is similar to a model with the same number of factors. However, our model allows us to investigate the impact of macro-financial variables in the absence of arbitrage or to incorporate a capital market perspective into the forecast. This is discussed in more detail in the next section.

###Insert Figure 4 about here####

Figure 4 illustrates the in-sample performance of our model, showcasing the predictive accuracy through percentage valuation errors and root mean squared errors (RMSEs). This visualization supports our assertion that the inclusion of macro-financial factors significantly enhances the model's ability to capture the dynamics of the implied volatility surface.

The analysis presented in Figure 4 underscores the robustness of our model during typical market conditions and highlights its superior performance in managing and predicting variations in implied volatility. These results not only validate our methodological approach but also confirm the practical applicability of incorporating macro-financial variables in financial modeling.

4.3. Selection of macro-financial variables

Our model employs 127 macroeconomic variables to estimate the implied volatility surface, effectively navigating the intricacies of the options market. Given the extensive array of factors involved in option pricing, which can be daunting for quantitative analysts, we utilize lasso regression alongside dimensionality reduction. This approach streamlines the factor selection process, retaining only the most pertinent factors to facilitate clearer model interpretation and execution.

Insert Table 3 about here

Initially, Lasso regression, rather than Ordinary Least Squares (OLS), is utilized to generate lambda (λ) coefficients, as detailed in Appendix 1:

We utilize the capacity of data augmentation techniques to populate the implied volatility surface on a monthly basis. Without this augmentation, the lasso coefficient tends to become insignificant, emphasizing the importance of a rich dataset in identifying key macro-financial variables that have a meaningful impact on implied volatility.

Utilizing three L1 hyperparameters, "Lasso alpha" values of 0.0008, 0.00082, and 0.00084, we compute RMSEs for out-of-sample tests and derive Lasso coefficients (λ) for 127 macro-financial variables. Each variable is adjusted by its corresponding Lasso coefficient, and the time-series average and standard deviation are utilized to compute the (Fama-MacBeth) t – $value = mean / std / \sqrt{T}$. We retain only those variables that exhibit (Fama-MacBeth) t – values exceeding 1.960. This process effectively narrows the set from 127 to 19 macro-financial variables, essential for forming robust economic deductions about the implied volatility surface.

In our analysis, we observed that several macroeconomic variables significantly influence implied volatility, as evidenced by their t -values. Particularly, labor market indicators such as All Employees: Construction with a t -value of 1.998 and the Help-Wanted Index for United States with a t -value of 2.067 positively correlate with implied volatility, suggesting that increased construction activity and higher demand for labor may reflect economic expansion, subsequently influencing market volatility. Conversely, negative relationships were noted in other labor market variables such as All Employees: Trade, Transportation & Utilities and All Employees: Financial Activities, which exhibit t -values of -1.983 and -1.996, respectively, indicating that higher employment in these sectors is associated with decreased volatility, potentially due to perceived economic stability.

Interest rate and exchange rate fluctuations also play a crucial role. The 10-Year Treasury Rate and the 1-Year Treasury Rate demonstrated inverse and direct relationships with implied volatility, respectively, highlighting how changes in interest rates, which impact bond prices, can affect market uncertainty. Additionally, a negative correlation with the Nominal Major Currencies U.S. Dollar Index (t-value: -1.992) suggests that a stronger dollar, often viewed as a safe haven, tends to reduce volatility.

In terms of output and income, the IP Index (t-value: 2.055) was found to correlate positively with implied volatility, indicating that heightened industrial production, signaling robust economic activities, could lead to increased market volatility through higher investment and trading volumes.

Price indices such as the Producer Price Index for both Intermediate and Final Demand and the CPI: Apparel also show a positive correlation with volatility (t-values of 2.067, 2.002, and 2.082, respectively), implying that inflationary pressures from rising prices may contribute to heightened market volatility as markets adjust to changing economic conditions.

Furthermore, the M2 Money Stock (t-value: 1.997) suggests that fluctuations in the money supply, which can influence inflation and interest rates, are also significant drivers of financial market conditions and volatility.

Interestingly, the CBOE S&P 100 Volatility Index: VXO showed a negative t-value of -2.020, indicating that periods of higher historical volatility might lead to reduced future volatility, possibly due to a mean reversion effect or the calming of previous market turbulence.

Finally, labor market conditions reflected through Avg Hourly Earnings: Manufacturing and Civilian Employment (t-values of 2.018 and 2.042, respectively) demonstrate that increases in earnings and employment levels might not only suggest economic strengthening but could also lead to increased volatility due to concerns over inflation or shifts in monetary policy responses.

These findings underscore the complex interplay between macroeconomic variables and implied volatility, offering critical insights for financial analysts and economists in forecasting market movements and formulating investment strategies amidst economic changes.

Overall, this approach enables data-driven practitioners to employ a comprehensive set of macro-financial variables effectively, streamlining the dataset after iterative testing to enhance economic interpretations and forecasts of implied volatility. This method allows for efficient

adaptation of economic insights into actionable investment strategies under no-arbitrage conditions.

4.4. Model comparison with classical model

In this chapter, we conduct a comparative analysis between our proposed macro-financial model and the classical model for estimating implied volatility (IV). The classical model, as traditionally defined, represents IV as a linear function of strike price, time-to-maturity (TTM), and their respective squared and interaction terms. This model is referred to as the “absolute smile” approach. In classical model, the independent variable is sometimes set as moneyness rather than absolute exercise price, which is called the “relative smile” approach. However, in this study, the absolute smile approach is considered. Jackwerth and Rubinstein (2012), Li and Pearson (2007), Kim (2009), and Choi and Ok (2012) compare these two approaches and show that the “absolute smile” approach performs better than the “relative smile” approach. The mathematical expression is given by:

$$IV = constant + \beta_1 strike + \beta_2 ttm + \beta_3 strike^2 + \beta_4 ttm^2 + \beta_5 ttm \cdot strike + \epsilon$$

The analysis includes evaluations based on the fit and forecasting accuracy of the classical model. Figure 5 and Figure 6 from the provided documents illustrate the time-series fit and performance comparison of the classical model, respectively. These figures show that while the classical model aligns closely with observed data, it tends to slightly underperform in capturing the tails of the distribution, as indicated by the RMSE and percentage valuation errors (Sun et al., 2018).

###insert Figure 5 and Figure 6###

In Figure 5, the classical model's implied IVs closely follow the observed IVs across different moneyness levels (0.97, 1.00, and 1.03), with both observed and model-implied IVs represented by solid and dashed lines. This visual alignment suggests a decent model fit within standard market conditions (Zeng, 2005).

Figure 6 further quantifies this relationship, comparing the in-sample model performance measured by the percentage valuation errors and root mean squared errors (RMSEs) of the

predicted values. The RMSE and percentage valuation errors provide a quantitative measure of the model's accuracy in predicting IVs, defined as the difference between actual values and model estimates (Ncube, 1996).

The classical model, while robust in its simplicity and widely used in historical contexts, shows limitations when compared with our enhanced model that incorporates macro-financial variables. Specifically, our model addresses the interaction effects and nonlinear relationships that the classical model overlooks. As demonstrated in Figure 7 of the provided documents, the predictive power of our model, incorporating an expanded set of macro-financial variables, significantly enhances forecasting capabilities across various economic scenarios.

The empirical tests conducted, as detailed in the document, involve refining the model through Lasso regression, further distinguishing our approach from the classical methodology by adjusting for multicollinearity and overfitting, particularly when a large number of predictors are involved.

While the classical model provides a foundational approach to estimating IV, our analysis underscores the benefits of integrating macro-financial variables to capture a more comprehensive picture of market dynamics. This integration not only improves predictive accuracy but also offers deeper insights into the economic and financial determinants of options market behavior.

4.5. Model performance during the financial crisis

We conducted an evaluation of our proposed model's effectiveness during the 2007–2010 global financial crisis to determine its robustness under extreme market conditions. Specifically, Figure 7 presents the Root Mean Square Error (RMSE) of our model during this tumultuous period, focusing on out-of-sample tests. For these tests, we utilized data spanning from February 1996 to December 2006—covering 73% of the entire sample period—to generate forecasts using a rolling method, similar to the approach outlined in Figure 5.

###Insert Figure 7 here###

The analysis period for performance comparison is defined from 2007 to 2010, which serves as the out-of-sample period for assessing the model's predictive accuracy (J. W. Choi, 2013). In Figure 7, the upper line indicates the out-of-sample RMSE for the period from

February 1996 to December 2006, while the lower line shows the RMSE specifically for the financial crisis period from January 2007 to December 2010.

The graphical representation in Figure 7 suggests that the expansion of the rolling window likely contributed to enhanced learning due to a larger dataset, which may have resulted in a lower RMSE during these years. These findings underscore the model's robust performance, confirming that its predictive capability remained effective even amid the significant market disruptions during the financial crisis. This resilience supports the model's utility in forecasting under a variety of economic conditions.

4.6. Lasso regression and out-of-sample tests

To further check the robustness of our results, we test our model using Lasso regression. We performed a 7:3 split over the initial sample period [1996:2, 2015:8] to define training and validation sets to modify hyperparameters (Lasso alpha) across papers. The same steps introduced in Figure 8 produce predicted values, denoted $dlog(IV)$, for the next training set period.

Insert Figure 8 here

Figure 8 shows a graph plotting the RMSE estimates to compare the performance of the models with 10, 34, 64, and 115 factors. Model construction also follows the same steps used to generate the results in Figure 4. The difference is that it uses Lasso regression instead of OLS.

Overall, the results indicate that the performance of our proposed model is not impaired during the out-of-sample period. The results in Panel A show that our model performs well when Lasso alpha is set to 0.01. The results in Panel B confirm that the predictability of the proposed model is not impaired after data augmentation for out-of-sample periods.

4.7. Nonlinear relationship between macro-financial variables

In reality, macro-financial indicators are often reported on similar dates, so their impact on asset prices cannot be completely separated from each other. To reflect this, we run

additional empirical tests and show which macro financial variables interacting with other variables have the greatest impact on forward rates.

More specifically, 8,128 independent variables were created, including 127 macro financial variables and 8,001 interaction variables for the empirical testing. Then, we apply steps 4 to 6 introduced in 3.1 Empirical methods with slight modifications. For example, in Step 4, we calculate the inverse of the diagonal matrix of Σ_t to generate $dz_{v,t}$, which denotes the source of risk. In step 6, we perform a Lasso regression with a Lasso alpha of 0.00082. The results are summarized and reported in Table 4.

Insert Table 4 here

As a result, all 466 variables are reported with t-values greater than 1.960, which is used to select the previous 19 macro financial variables reported as most important in Table 3. Table 4 reports the top 20 most statistically significant variables. The results can lead to numerous economic interpretations, and the following are just a few examples.

Three interaction variables, in particular, have demonstrated noteworthy predictive power, suggesting nuanced insights into the economic forces shaping market dynamics.

First, the interaction between civilians unemployed for less than five weeks and employment levels in the construction sector, denoted as UEMPLT5_USCONS, yielded a t-value of 1.996. This variable captures early economic shifts, where short-term unemployment reflects immediate economic conditions and employment in construction signals both current and future economic expectations. The significant influence of this variable on implied volatility underscores the sensitivity of option markets to labor market conditions in sectors directly impacted by economic cycles.

Second, the product of the total business inventory-to-sales ratio and average hourly earnings in manufacturing, labeled ISRATIOx_CES3000000008, also exhibited a t-value of 1.996. This interaction likely highlights the dual impact of supply chain dynamics and cost pressures within the manufacturing sector. High inventory levels relative to sales can signal overstock and potential slowing demand, while changes in wages might indicate cost-driven inflationary pressures. Together, these factors contribute to greater market volatility, reflecting broader economic uncertainties that affect investment and pricing decisions in financial markets.

Lastly, the combination of initial unemployment claims and average weekly hours in goods-producing sectors, referred to as CLAIMSx_CES0600000007, provides significant insights with a t-value of 1.996. This variable serves as a barometer for labor market health within core industries, where an increase in claims and alterations in working hours can signal shifts in production capabilities and economic output. Such labor market fluctuations are critical indicators of economic health, influencing investor sentiment and market volatility due to their direct ties to production and economic activity.

These interaction variables enrich our understanding of the complex relationships between macroeconomic indicators and market behavior, particularly in how they collectively influence the volatility perceived by investors in the options market. This enhanced understanding assists in better anticipating market movements and refining risk management strategies in response to economic indicators.

5. Conclusion

This study has demonstrated that the integration of a broad array of macro-financial factors significantly enhances the predictive accuracy of implied volatility (IV) models, both in-sample and out-of-sample. Traditional no-arbitrage models have faced challenges in accommodating large sets of observable macro-financial factors due to their inherent complexities and computational constraints. To overcome these limitations, we have developed a novel model that simplifies the inclusion of numerous factors, including the application of advanced machine learning techniques, while maintaining the critical no-arbitrage framework.

The empirical evidence suggests that the implied volatility of our sample is markedly influenced by key economic indicators such as housing market dynamics, short-term interest rates, stock market fluctuations, and inflationary trends (Bielinski & Broby, 2021). By optimizing the number of macro-financial variables, our model not only facilitates sharper economic insights but also reduces the cumbersome effort involved in data collection, allowing users to focus on economically significant variables without reliance on arbitrary choices (Barillas & Nimark, 2017).

Moreover, the adoption of data augmentation techniques has proven effective in mitigating the challenges posed by small sample sizes, thereby enhancing the robustness of machine learning applications within our modeling framework (Chen et al., 2019). This approach

ensures that the forecasting capabilities of our model remain robust, even when subjected to rigorous testing during periods of market volatility.

The incorporation of machine learning into asset pricing models presents several advantages, including increased computational efficiency and precision. However, it also raises concerns about potential overfitting and the adherence to foundational economic principles such as the no-arbitrage condition (Gu et al., 2018). Our model addresses these concerns by integrating machine learning in a manner that respects and reinforces these essential theoretical constraints.

Looking forward, there are several avenues for further research. Expanding the application of the proposed model to international markets could provide broader validation of its effectiveness. Exploring its applicability to different asset classes, such as equities and interest-rate derivatives, could further enhance its utility and relevance across the financial sector.

Additionally, future studies might explore a variety of machine learning techniques not examined in this study. While our application of Lasso regression and PCA has been fruitful in identifying relevant macro-financial variables, investigating other methodologies like deep learning or autoencoders could offer new insights and potentially improve the model's performance and interpretability.

In conclusion, this paper extends the traditional methodologies used in financial modeling by effectively incorporating macro-financial variables into the estimation of implied volatility surfaces. This not only bridges the gap between theoretical models and practical market applications but also equips financial practitioners with a more robust and adaptable tool for navigating complex market environments.

Reference

- Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *The Journal of Finance*, 53(2), 499–547. <https://doi.org/10.1111/0022-1082.215228>
- Amin, K. I., & Jarrow, R. A. (1991). Pricing Foreign Currency Options under Stochastic Interest Rates. *Journal of International Money and Finance*, 10(3), 310–329. [https://doi.org/10.1016/0261-5606\(91\)90013-A](https://doi.org/10.1016/0261-5606(91)90013-A)
- Amin, K. I., & Jarrow, R. A. (1992). Pricing Options On Risky Assets In A Stochastic Interest Rate Economy1. *Mathematical Finance*, 2(4), 217–237. <https://doi.org/10.1111/j.1467-9965.1992.tb00030.x>
- Armeanu, D., & Lache, L. (2008). Application of the Model of Principal Components Analysis on Romanian Insurance Market. *Theoretical and Applied Economics*, 11–20.
- Bakshi, G., Charles, C., & Chen, Z. (1997). Empirical Performance of Alternative Option Pricing Models. *Journal of Finance*, 52(5), 2003–2049. <https://doi.org/10.1111/j.1540-6261.1997.tb02749.x>
- Barillas, F., & Nimark, K. P. (2017). Speculation and the Bond Market: An Empirical No-Arbitrage Framework. *ERN: Expectations in Economic Theory & Markets (Topic)*. <https://doi.org/10.2139/ssrn.2109721>
- Bates, D. S. (2003). Empirical Option Pricing: A Retrospection. *Journal of Econometrics*, 116(1–2), 387–404.
- Beber, A., & Brandt, M. W. (2006). The Effect of Macroeconomic News on Beliefs and Preferences: Evidence from the Options Market. *Journal of Monetary Economics*, 53(8), 1997–2039. <https://doi.org/10.1016/j.jmoneco.2006.05.013>
- Bielinski, A., & Broby, D. (2021). *Machine Learning Methods in Asset Pricing*. <https://doi.org/10.2139/ssrn.3950524>
- Chen, L., Pelger, M., & Zhu, J. (2019). Deep Learning in Asset Pricing. *Research Methods & Methodology in Accounting eJournal*. <https://doi.org/10.2139/ssrn.3350138>
- Choi, J. W. (2013). The 2007–2010 U.S. Financial Crisis: Its Origins, Progressions, and Solutions. *The Journal of Economic Asymmetries*, 10(2), 65–77.
- Choi, Y., Jordan, S. J., & Ok, S. (2012). Dividend-Rollover Effect and the Ad Hoc Black-Scholes Model. *Journal of Futures Markets*, 32(8), 742–772. <https://doi.org/10.1002/fut.21541>
- Choi, Y., & Ok, S. (2012). Effects of Rollover Strategies and Information Stability on the Performance Measures in Options Markets: An Examination of the KOSPI 200 Index Options Market. *Journal of Futures Markets*, 32(4), 360–388. <https://doi.org/10.1002/fut.20522>
- Cont, R., & Fonseca, J. da. (2002). Dynamics of Implied Volatility Surfaces. *Quantitative Finance*, 2(1), 45. <https://doi.org/10.1088/1469-7688/2/1/304>

- Dixit, A., & Singh, S. (2018). Ad-Hoc Black–Scholes vis-à-vis TSRV-based Black–Scholes: Evidence from Indian Options Market. *Journal of Quantitative Economics*, 16(1), 57–88. <https://doi.org/10.1007/s40953-017-0078-3>
- Duan, J.-C. (1995). The Garch Option Pricing Model. *Mathematical Finance*, 5(1), 13–32. <https://doi.org/10.1111/j.1467-9965.1995.tb00099.x>
- Duan, J.-C., & Zhang, H. (2001). Pricing Hang Seng Index Options around the Asian Financial Crisis – A GARCH Approach. *Journal of Banking & Finance*, 25(11), 1989–2014. [https://doi.org/10.1016/S0378-4266\(00\)00166-7](https://doi.org/10.1016/S0378-4266(00)00166-7)
- Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied Volatility Functions: Empirical Tests. *The Journal of Finance*, 53(6), 2059–2106. <https://doi.org/10.1111/0022-1082.00083>
- Fama, E., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81, 607–636. <https://doi.org/10.1086/260061>
- Fengler, M. R. (2009). Arbitrage-free Smoothing of the Implied Volatility Surface. *Quantitative Finance*, 9(4), 417–428. <https://doi.org/10.1080/14697680802595585>
- Figlewski, S. (2008). *Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio* (SSRN Scholarly Paper 1256783). <https://papers.ssrn.com/abstract=1256783>
- Frank, M. Z., & Yang, K. (2021). Predicting Firm Profits: From Fama-MacBeth to Gradient Boosting. *Capital Markets: Market Efficiency eJournal*. <https://doi.org/10.2139/ssrn.3919194>
- Green, B., & Thomas, S. (2019). Inference and Prediction of Stock Returns using Multilevel Models. *SPGMI: Compustat Fundamentals (Topic)*. <https://doi.org/10.2139/ssrn.3411358>
- Gu, S., Kelly, B., & Xiu, D. (2018). Empirical Asset Pricing Via Machine Learning. *NBER Working Paper Series*. <https://doi.org/10.2139/ssrn.3159577>
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), 2223–2273. <https://doi.org/10.1093/rfs/hhaa009>
- Heston, S. L., & Nandi, S. (2000). A Closed-Form GARCH Option Valuation Model. *The Review of Financial Studies*, 13(3), 585–625. <https://doi.org/10.1093/rfs/13.3.585>
- Ho, T. S., Stapleton, R. C., & Subrahmanyam, M. G. (1997). The Valuation of American Options with Stochastic Interest Rates: A Generalization of the Geske—Johnson Technique. *The Journal of Finance*, 52(2), 827–840. <https://doi.org/10.1111/j.1540-6261.1997.tb04823.x>
- Hsieh, K. C., & Ritchken, P. (2005). An Empirical Comparison of GARCH Option Pricing Models. *Review of Derivatives Research*, 8(3), 129–150. <https://doi.org/10.1007/s11147-006-9001-3>
- Hull, J., & White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, 42(2), 281–300. <https://doi.org/10.2307/2328253>
- Jackwerth, J. C., & Rubinstein, M. (2012). Recovering Stochastic Processes from Option Prices. *Contemporary Studies in Economic and Financial Analysis*, 123–153.

- Jiang, G. J., & Tian, Y. S. (2005). The Model-Free Implied Volatility and Its Information Content. *The Review of Financial Studies*, 18(4), 1305–1342. <https://doi.org/10.1093/rfs/hhi027>
- Johnson, H., & Shanno, D. (1987). Option Pricing when the Variance Is Changing. *Journal of Financial and Quantitative Analysis*, 22(2), 143–151. <https://doi.org/10.2307/2330709>
- Kim, I. J., & Kim, S. (2004). Empirical Comparison of Alternative Stochastic Volatility Option Pricing Models: Evidence from Korean KOSPI 200 Index Options Market. *Pacific-Basin Finance Journal*, 12(2), 117–142. [https://doi.org/10.1016/S0927-538X\(03\)00042-8](https://doi.org/10.1016/S0927-538X(03)00042-8)
- Kim, I. J., & Kim, S. (2005). Is it Important to Consider the Jump Component for Pricing and Hedging Short-term Options? *Journal of Futures Markets*, 25(10), 989–1009.
- Kim, S. (2009). The Performance of Traders' Rules in Options Market. *Journal of Futures Markets*, 29(11), 999–1020. <https://doi.org/10.1002/fut.20403>
- Kim, S. (2014). Are Traders' Rules Useful for Pricing Options? Evidence from Intraday Data. *Journal of Risk*, 17(1), 63-84. <https://doi.org/10.21314/JOR.2014.277>
- Kim, S. (2017). Pricing and Hedging Options with Rollover Parameters. *Journal of Risk*, 19(5), 1-40. <https://doi.org/10.21314/JOR.2017.352>
- Kou, S. G. (2002). A Jump-Diffusion Model for Option Pricing. *Management Science*, 48(8), 1086–1101.
- Kou, S. G., & Wang, H. (2004). Option Pricing Under a Double Exponential Jump Diffusion Model. *Management Science*, 50(9), 1178–1192. <https://doi.org/10.1287/mnsc.1030.0163>
- Ku, Y.-C., Bloomfield, P., & Ghosh, S. (2011). A Flexible Observed Factor Model with Separate Dynamics for the Factor Volatilities and their Correlation Matrix. *Statistical Modelling*, 14, 1–20. <https://doi.org/10.1177/1471082X13490016>
- Lee, G., & Ryu, D. (2024). Linear Extrapolation and Model-free Option Implied Moments. *Borsa Istanbul Review*. <https://doi.org/10.1016/j.bir.2024.01.009>
- Li, M., & Pearson, N. D. (2004). *Price Deviations of S&P 500 Index Options from the Black-Scholes Formula Follow a Simple Pattern* (SSRN Scholarly Paper 686803). <https://doi.org/10.2139/ssrn.686803>
- Li, M., & Pearson, N. D. (2007). *A "Horse Race" Among Competing Option Pricing Models Using S&P 500 Index Options* (SSRN Scholarly Paper 952770). <https://doi.org/10.2139/ssrn.952770>
- Ludvigson, S. C., & Ng, S. (2009). Macro Factors in Bond Risk Premia. *The Review of Financial Studies*, 22(12), 5027–5067. <https://doi.org/10.1093/rfs/hhp081>
- Madan, D. B., Carr, P. P., & Chang, E. C. (1998). The Variance Gamma Process and Option Pricing. *Review of Finance*, 2(1), 79–105. <https://doi.org/10.1023/A:1009703431535>

- McCracken, M. W., & Ng, S. (2016). FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, 34(4), 574–589. <https://doi.org/10.1080/07350015.2015.1086655>
- Melino, A., & Turnbull, S. M. (1990). Pricing Foreign Currency Options with Stochastic Volatility. *Journal of Econometrics*, 45(1), 239–265. [https://doi.org/10.1016/0304-4076\(90\)90100-8](https://doi.org/10.1016/0304-4076(90)90100-8)
- Merton, R. C. (1976). Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics*, 3(1), 125–144. [https://doi.org/10.1016/0304-405X\(76\)90022-2](https://doi.org/10.1016/0304-405X(76)90022-2)
- Mixon, S. (2002). Factors Explaining Movements in the Implied Volatility Surface. *Derivatives eJournal*. <https://doi.org/10.1002/FUT.10039>
- Naik, V. (1993). Option Valuation and Hedging Strategies with Jumps in the Volatility of Asset Returns. *The Journal of Finance*, 48(5), 1969–1984. <https://doi.org/10.1111/j.1540-6261.1993.tb05137.x>
- Naik, V., & Lee, M. (1990). General Equilibrium Pricing of Options on the Market Portfolio with Discontinuous Returns. *The Review of Financial Studies*, 3(4), 493–521. <https://doi.org/10.1093/rfs/3.4.493>
- Ncube, M. (1996). Modelling Implied Volatility with OLS and Panel Data Models. *Journal of Banking and Finance*, 20, 71–84. [https://doi.org/10.1016/0378-4266\(94\)00109-X](https://doi.org/10.1016/0378-4266(94)00109-X)
- Nyberg, P. M., & Vilhelmsson, A. (2008). Volatility Risk Premium, Risk Aversion and the Cross-Section of Stock Returns. *European Finance eJournal*. <https://doi.org/10.2139/ssrn.1024158>
- OptionMetrics, W. R. D. (n.d.). *OptionMetrics*. WRDS. Retrieved April 19, 2024, from <https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/optionmetrics/>
- Rindell, K. (1995). Pricing of Index Options when Interest Rates are Stochastic: An Empirical Test. *Journal of Banking & Finance*, 19(5), 785–802. [https://doi.org/10.1016/0378-4266\(94\)00087-J](https://doi.org/10.1016/0378-4266(94)00087-J)
- Scott, L. O. (1987). Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application. *Journal of Financial and Quantitative Analysis*, 22(4), 419–438. <https://doi.org/10.2307/2330793>
- Stein, E. M., & Stein, J. C. (1991). Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *The Review of Financial Studies*, 4(4), 727–752. <https://doi.org/10.1093/rfs/4.4.727>
- Sun, H., Wang, M., Du, B., & Wang, J. (2018). A Weighted Strike-Related Implied Volatility Model. *2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)*, 212–215. <https://doi.org/10.1109/DCABES.2018.00062>
- Ulrich, M., & Walther, S. (2020). Option-implied Information: What’s the Vol Surface Got to Do with It? *Review of Derivatives Research*, 23(3), 323–355. <https://doi.org/10.1007/s11147-020-09166-0>
- Wiggins, J. B. (1987). Option Values under Stochastic Volatility: Theory and Empirical Estimates. *Journal of Financial Economics*, 19(2), 351–372. [https://doi.org/10.1016/0304-405X\(87\)90009-2](https://doi.org/10.1016/0304-405X(87)90009-2)

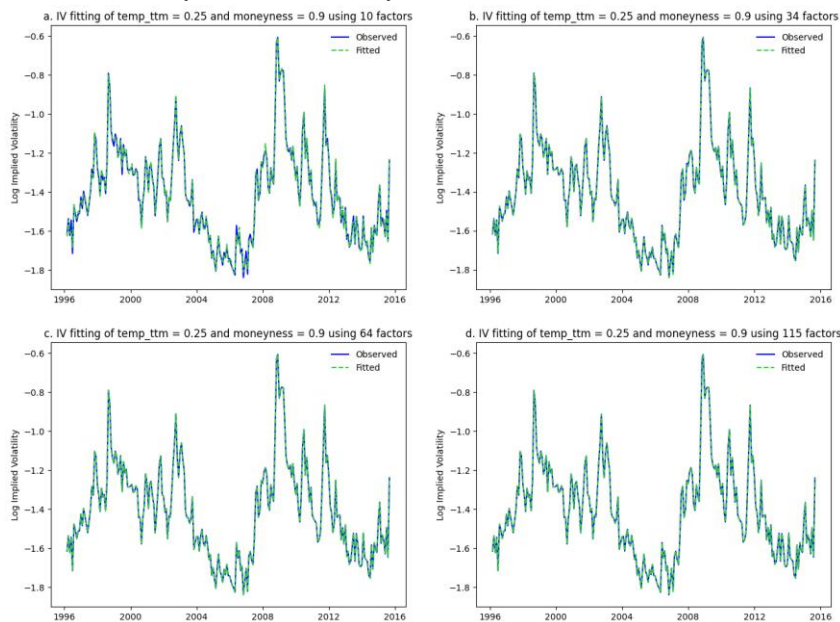
Zeng, J. (2005). *Application of Smoothing Techniques to Implied Volatility*.
<https://doi.org/10.18452/14056>

Tables And Figures

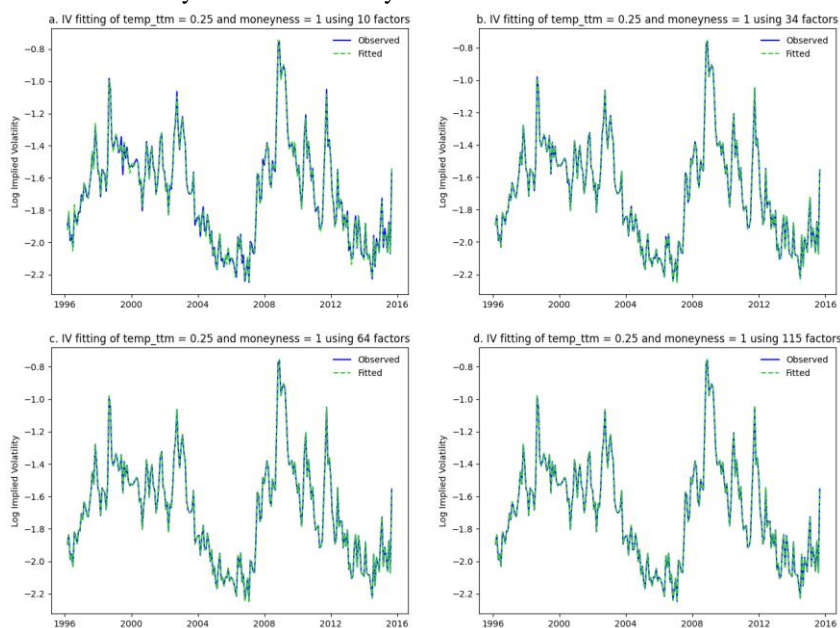
Figure 1. Time-series fit of model-implied IV using 10, 34, 64, and 115 macro (PCA) factors

This figure plots the implied volatility fitting for moneyness with 0.94, 1.0, and 1.025, as observed and implied by our model with 10, 34, 64, and 115 macro factors. The sample data are constructed using the implied volatility data of S&P 500 options. The sample period ranges from 1996:1 to 2015:8. In all graphs, solid lines represent observed implied volatilities and dashed lines (nearly indistinguishable from solid lines) represent model-implied IVs. To prevent forward-looking bias, we lag macro variables by one month.

Panel A. Maturity = 0.25 and Moneyness = 0.9



Panel B. Maturity = 0.25 and Moneyness = 1.0



Panel C. Maturity = 0.25 and Moneyness = 1.1

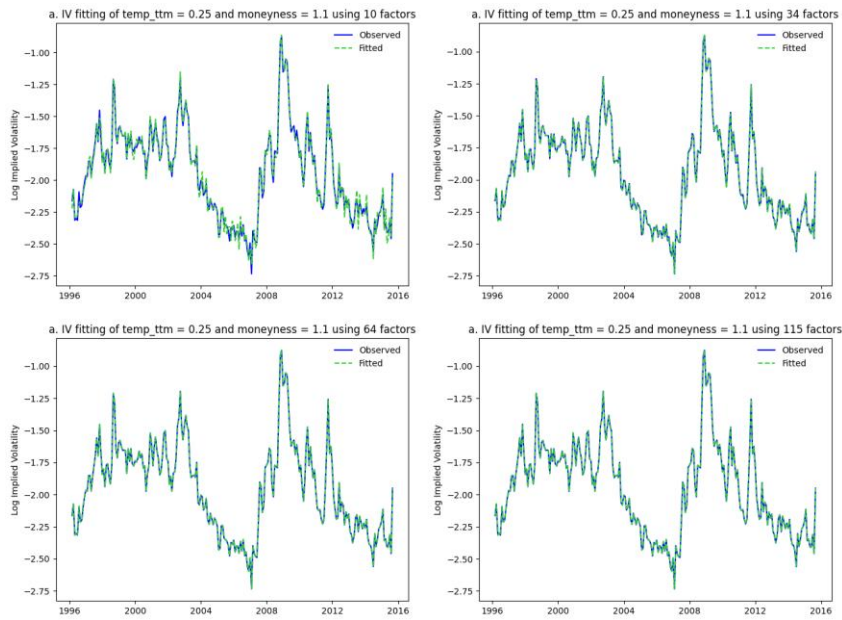
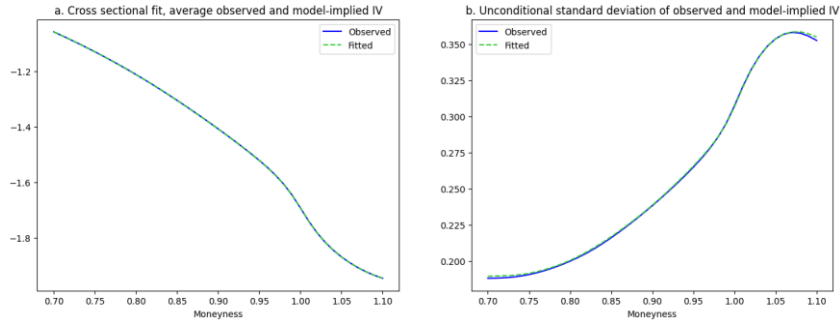


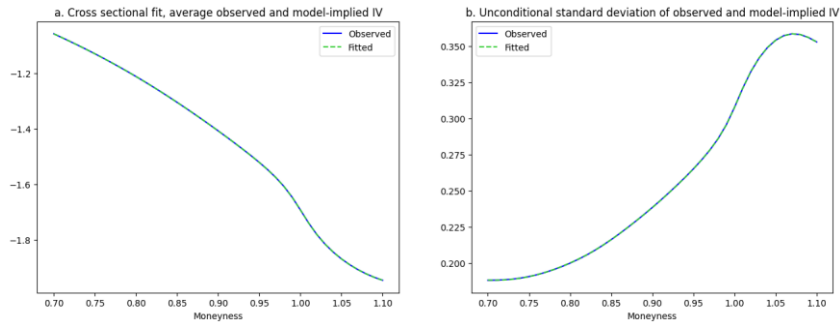
Figure 2. Cross-sectional diagnostics of the model-implied IV using 10, 34, 64, and 115 macro factors

This figure presents graphs exhibiting the cross-sectional fit of the implied volatilities generated using our model with 10, 34, 64, and 115 macro factors. The sample data are constructed using the implied volatility data of the S&P500 option. The sample period ranges from 1996:1 to 2015:8. The graph on the left plots the unconditional means, whereas on the right side plots the unconditional standard deviations of the observed and model-implied IVs.

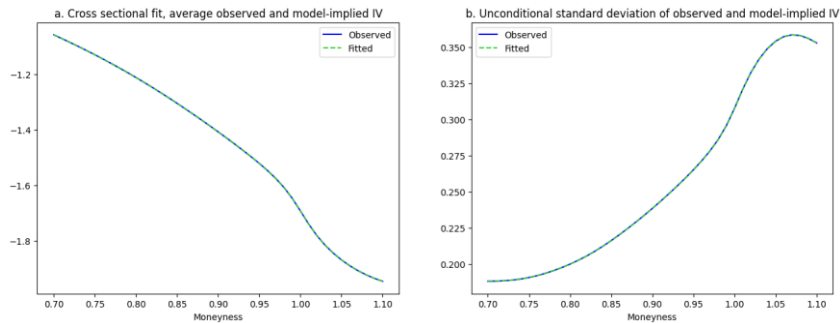
Panel A. Model-implied IV using 10 macro factors



Panel B. Model-implied IV using 34 macro factors



Panel C. Model-implied IV using 64 macro factors



Panel D. Model-implied IV using 115 macro factors

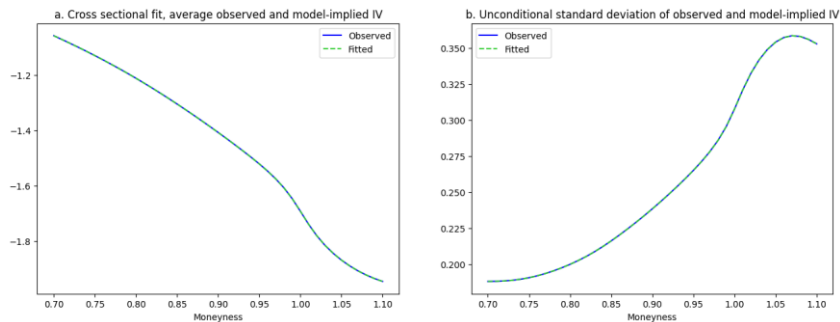


Figure 3. Predictive power across different moneyness

This figure plots absolute t -values, $|t|$, for β generated from our regression to test the predictive power of our proposed model. The following equations are used:

- 1) without control variables

$$r_{i,t+1} = \text{constant} + \beta_i G_{i,t} + \epsilon_{i,t+1} .$$

- 2) with control variables

$$r_{i,t+1} = \text{constant} + \beta_i G_{i,t} + (\text{controls}) + \epsilon_{i,t+1} .$$

The independent variable ($G_{i,t}$) is the difference between the observed and the model-implied IVs at each moneyness level i and time period t with its coefficient denoted as β . The dependent variable is the difference between the IVs at time period t and $t + 1$, which moneyness level is i . The difference between the IVs at time period t and $t + 1$ is used as a control variable 1 and the coefficient is denoted *control 1*. The log of IV at time period $t - 1$ and moneyness level i is used as a control variable 2 and the coefficient is denoted *control 2*. The sample data is constructed using S&P500 implied volatility data. The sample period is from 1996:1 to 2015:8.

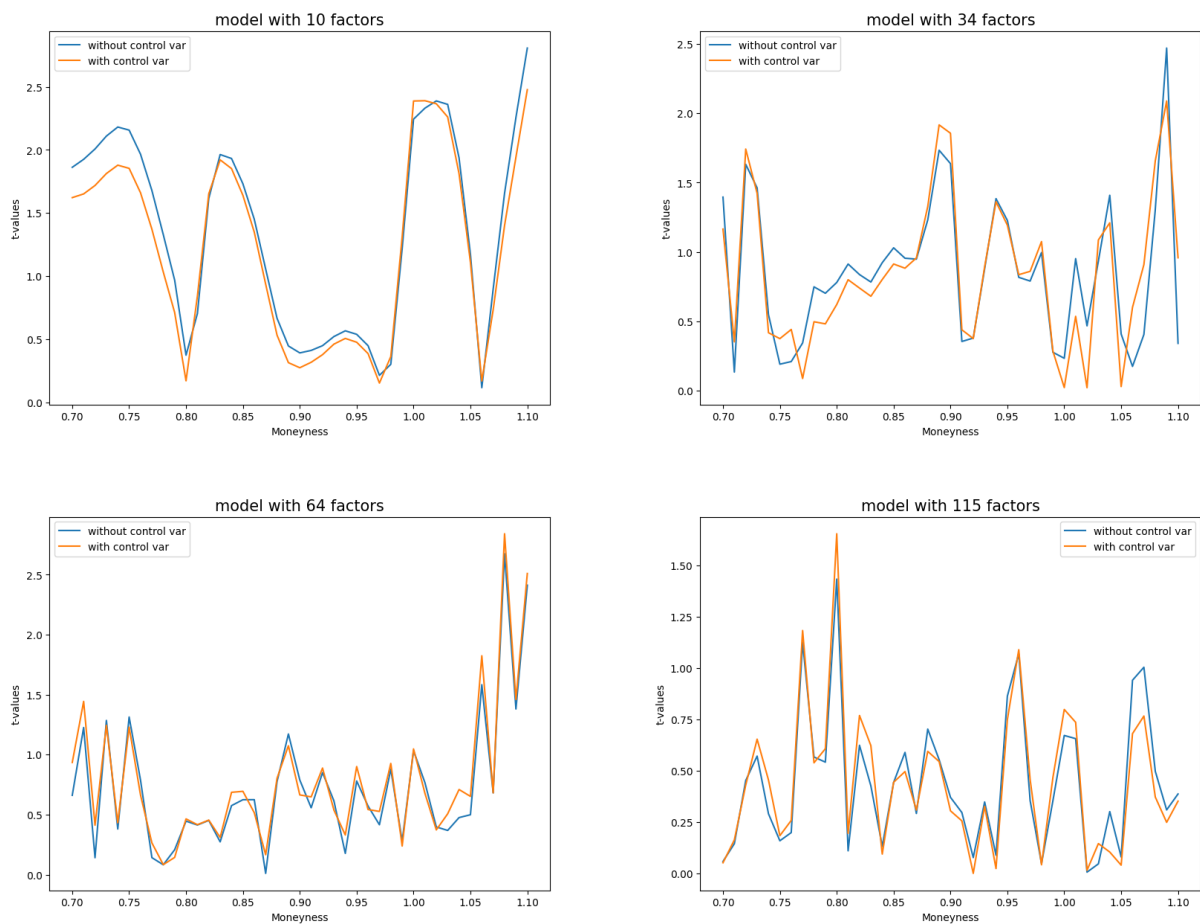


Figure 4. In-sample model performance comparison

This figure plots in-sample model performance for the comparison measured by the percentage valuation errors (ϵ) of the predicted value of $d\log(IV)$ and root mean squared errors (RMSEs). The percentage valuation errors measure the accuracy in predicting $d\ln(IV)$ and are defined as $\epsilon \equiv d\log(\widehat{IV}) / d\log(IV) - 1$ where $d\log(IV)$ is the difference of implied volatilities and $d\log(\widehat{IV})$ is the corresponding model estimate. RMSE measures the difference between the actual value and predicted value, and it is defined as

$$RMSE(d\log(\widehat{IV})) \equiv \left[\frac{\left(d\log(\widehat{IV}) - d\log(IV) \right)^2}{N} \right]^{\frac{1}{2}}$$

=

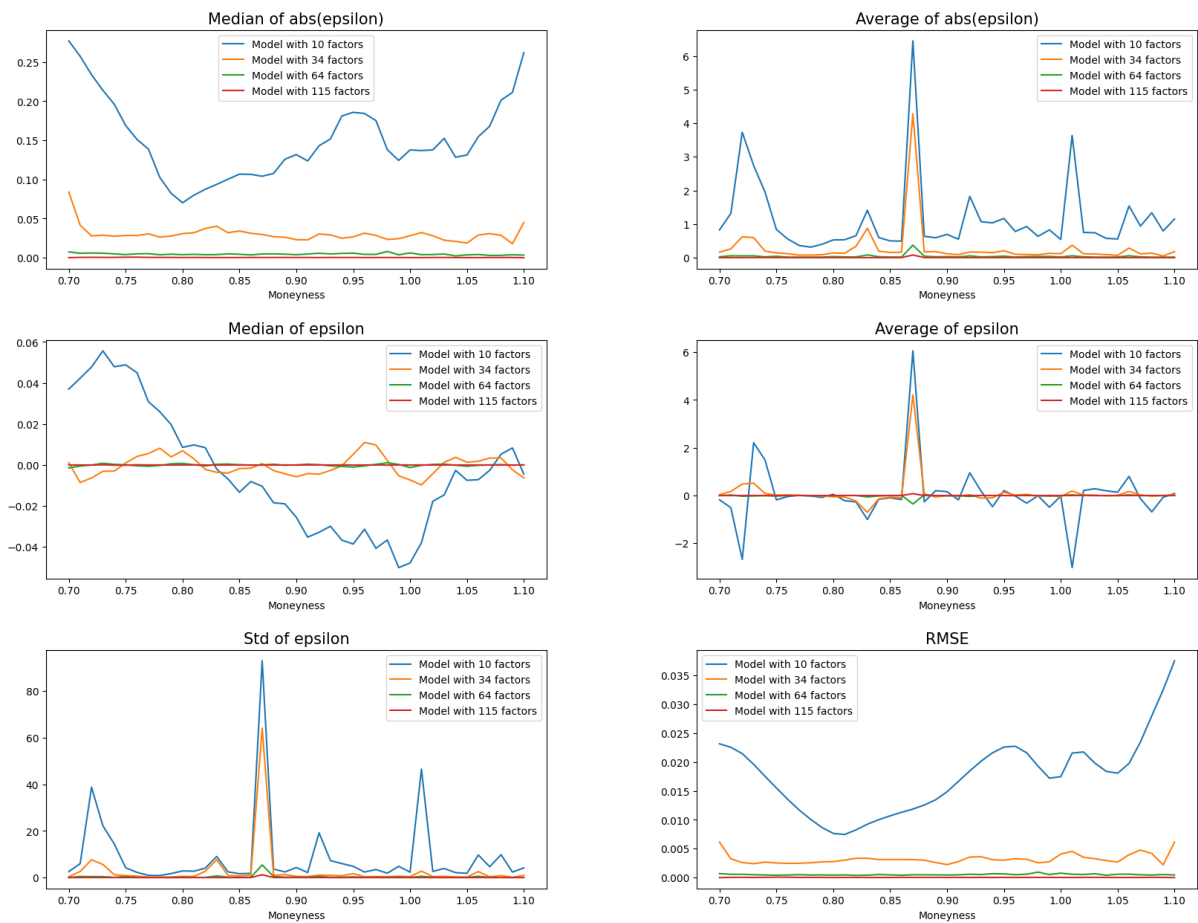


Figure 5. Time-series fit of classic model-implied IV

This figure plots the implied volatility fitting for moneyness with 0.97, 1.0, and 1.03, as observed and implied by classical model with stike and moneyness. The sample data are constructed using the implied volatility data of S&P 500 options. The sample period ranges from 1996:1 to 2015:8. In all graphs, solid lines represent observed implied volatilities and dashed lines (nearly indistinguishable from solid lines) represent classical model-implied IVs.

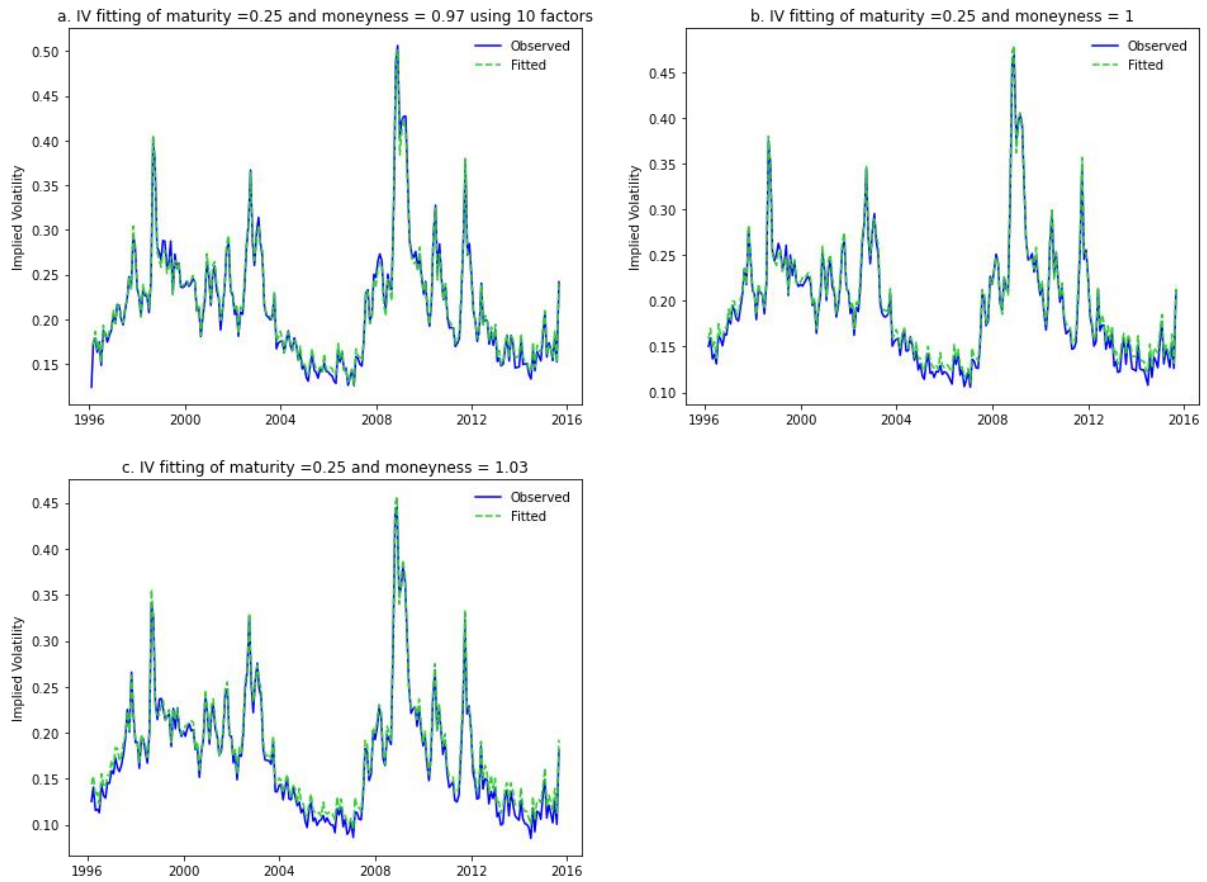


Figure 6. classical model performance comparison

This figure plots classical model performance for the comparison measured by the percentage valuation errors (ϵ) of the predicted value of $d\log(IV)$ and root mean squared errors (RMSEs). The percentage valuation errors measure the accuracy in predicting $d\ln(IV)$ and are defined as $\epsilon \equiv d\log(\widehat{IV})/d\log(IV) - 1$ where $d\log(IV)$ is the difference of implied volatilities and $d\log(\widehat{IV})$ is the corresponding model estimate. RMSE measures the difference between the actual value and predicted value, and it is defined as

$$RMSE(d\log(\widehat{IV})) \equiv \left[\frac{(d\log(\widehat{IV}) - d\log(IV))^2}{N} \right]^{\frac{1}{2}}$$

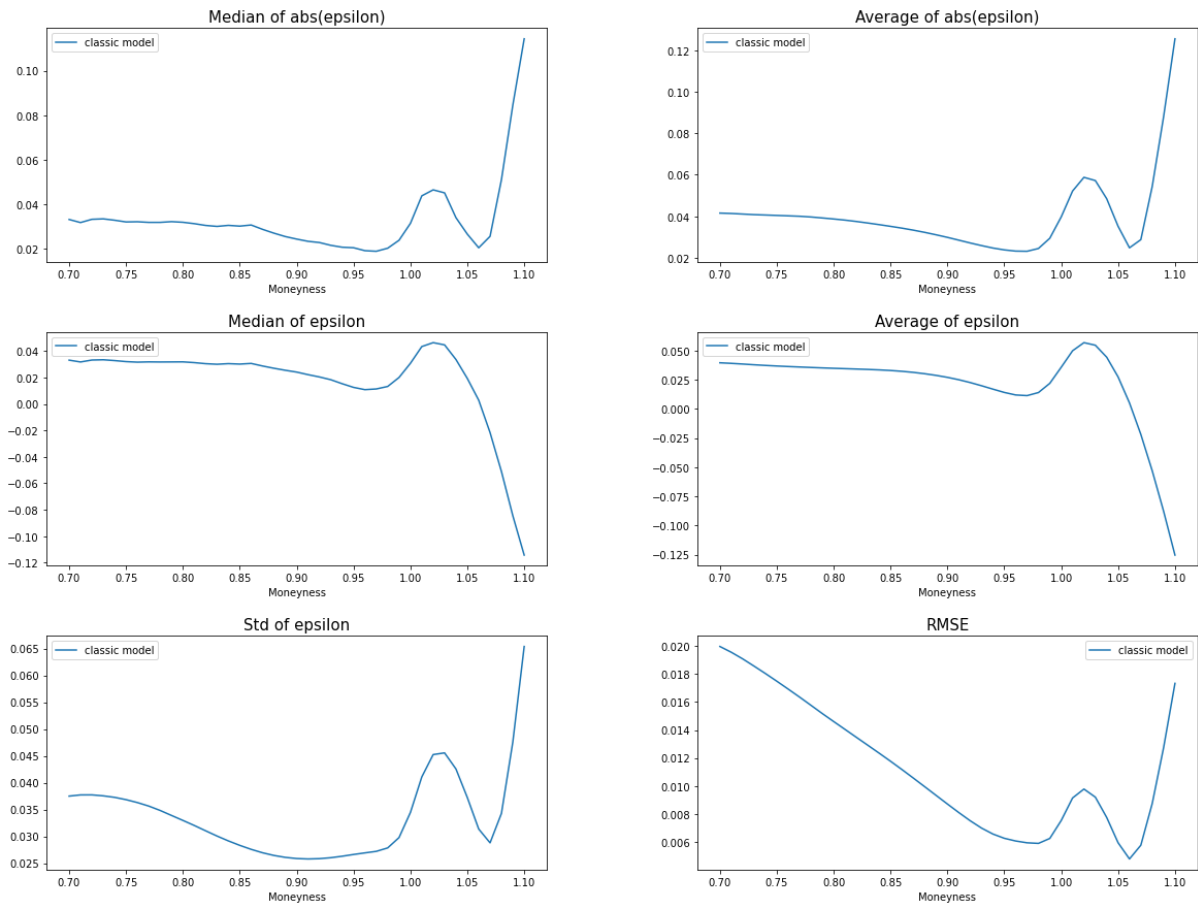


Figure 7. Out-of-sample model performance comparison during the subprime financial crisis.

This figure plots the estimates of root mean squared error (RMSE) during the subprime financial crisis for out-of-sample tests. Using this data from 1996:2 to 2006:12 (i.e., 73% of the entire sample period), we estimate the prediction values in a rolling method. The following steps describe the process:

Step 1. Use the data from 1996:2 to 2006:12 for yield fitting

Step 2. Calculate $dlog(IV)$ by using model-implied and observed IV in the equation,

$$\Delta f_{observed,t+1} = \alpha + \beta_t(f_{observed,t} - f_{model,t}) + \gamma_t \Delta f_{observed,t} + \epsilon_t$$

Step 3. Expand the period for the estimation of prediction values to 1996:2 to 2006:12, which adds an additional month at the end of the previous estimation period. Repeat Steps 1 and 2.

For performance comparison, The analysis period for performance comparison is defined from 2007 to 2010, which serves as the out-of-sample period for assessing the model's predictive accuracy, to set the out-of-sample period. Therefore, the upper five lines in the graph represent the rolling out-of-sample RMSEs during the period of 2007:1 to 2010:12.

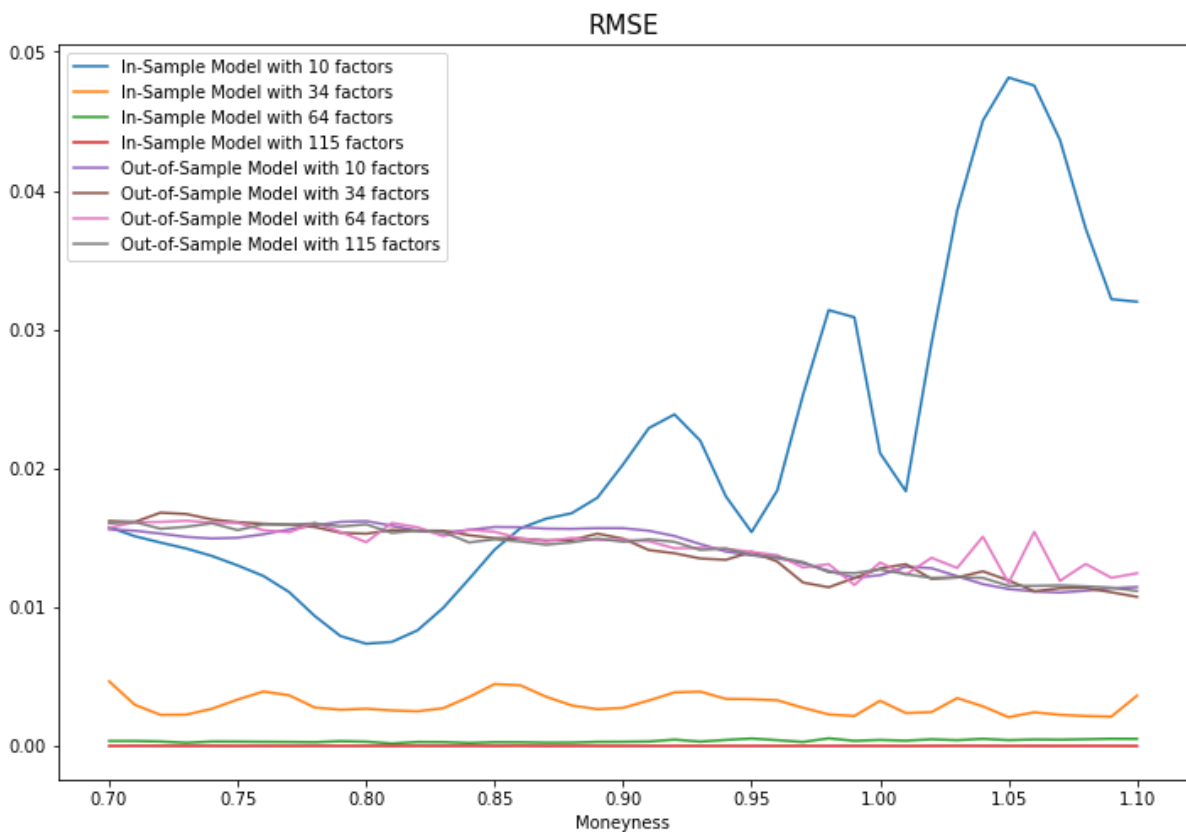


Figure 8. Out-of-sample model performance comparison

This figure presents the out-of-sample results for model performance comparison measured by the percentage valuation errors (ϵ) of the predicted value of $d\log(IV)$ and root mean squared errors (RMSEs). The percentage valuation errors measure the accuracy in predicting $d\log(IV)$ and are defined as $\epsilon \equiv d\log(\widehat{IV}) / d\log(IV) - 1$ where $d\log(IV)$ is the difference of implied volatilities and $d\log(\widehat{IV})$ is the corresponding model estimate. RMSE measures the difference between the actual value and predicted value, and it is defined as

$RMSE(d\log(\widehat{IV})) \equiv \left[\frac{(d\log(\widehat{IV}) - d\log(IV))^2}{N} \right]^{\frac{1}{2}}$. For out-of-sample tests, we use the implied volatility data of S&P500 during the first 164 months, from 1996:2 to 2009:9 (i.e., 70% of the entire sample period), to kickstart the monthly rolling estimation for prediction values. This process is described as follows.

Step 1. Use the data from 1996:2 to 2009:9 for $d\log(IV)$ fitting

Step 2. Calculate $d\log(\widehat{IV})$ using the model-implied and observed IVs using the following equation:

$$r_{i,t+1} = constant + \beta_i G_{i,t} + (controls) + \epsilon_{i,t+1}.$$

Step 3. We expand the period for the estimation of prediction values to [1996:2 and 2009:10], which adds an additional month at the end of the previous rolling estimation period. Repeat Steps 1 and 2.

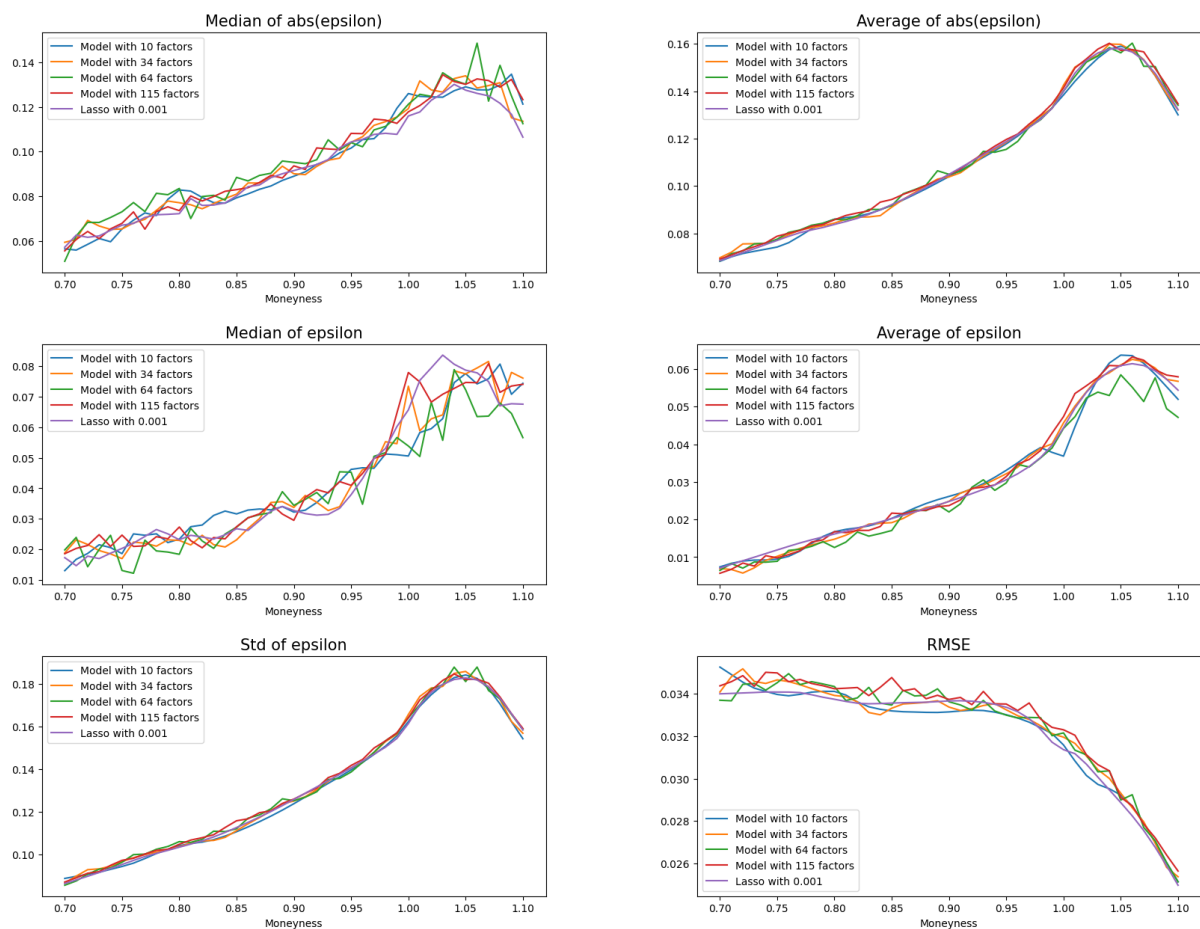


Table 1. Summary statistics of observed and model-implied log IV

This table reports summary statistics of the observed and model-implied IVs. The sample data are constructed using the implied volatility data of S&P500. The sample period ranges from 1996:1 to 2015:8. Panel A shows the summary statistics of implied volatilities observed for the same period. Panel B reports the summary statistics of the implied volatilities generated by the 10 macro factors in the proposed model. Panel C, Panel D and Panel E report the summary statistics of the implied volatilities generated by the 34 macro factors, 64 macro factors, and 115 macro factors in our proposed model. The fitted value of difference between implied volatilities at time t and t + 1 generated by the model are transformed into the log of implied volatilities, which the log of IVs are used as summary statistics to compare with the observed IVs. For all panels, the number of observations (count), average values (mean), standard deviations (std), minimum values (min), 25% (25%), 50% (50%), and 75% (75%) percentile values, and maximum values (max) are reported.

Panel A. Observed IV

Table with 30 columns and 21 rows of summary statistics for observed IVs, including measures like count, mean, std, min, 25%, 50%, 75%, and max.

Table 3. Relative importance of macro-financial variables

This table reports the relative importance of the macro-financial variables using LASSO regression. With eigenvalues obtained from conducting PCA on the data of 127 macro-financial variables, we calculate a Lasso coefficient multiplied by the eigenvector for each macro-financial variable, for which the time-series average and standard deviation are used to calculate the t -value of each macro-financial variable using (Fama-MacBeth) t – $value = mean / std / \sqrt{T}$. The LASSO alpha (L1 hyperparameter) ranges from 0.0001 to 20.

Penal A. alpha = 0.0008

No.	FRED	Category	Description	t-value
1	USCONS	Labor Market	All Employees: Construction	1.998
2	USTPU	Labor Market	All Employees: Trade, Transportation & Utilities	-1.983
3	GS10	Interest rate and Exchange rates	10-Year Treasury Rate	-1.998
4	HWI	Labor Market	Help-Wanted Index for United States	2.067
5	INDPRO	Output and Income	IP Index	2.055
6	UEMP15OV	Labor Market	Civilians Unemployed - 15 Weeks & Over	2.047
7	WPSID62	Prices	Producer Price Index by Commodity: Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand	2.067
8	CPIAPPSL	Prices	CPI : Apparel	2.082
9	M2SL	Money and Credit	M2 Money Stock	1.997
10	USFIRE	Labor Market	All Employees: Financial Activities	-1.996
11	GS1	Interest rate and Exchange rates	1-Year Treasury Rate	1.992
12	TWEXAFEGSMTHx	Others	Nominal Major Currencies U.S. Dollar Index (Goods Only)	-1.992
13	VIXCLSx	Stock Market	CBOE S&P 100 Volatility Index: VXO	-2.020
14	CE16OV	Labor Market	Civilian Employment	2.042
15	WPSFD49502	Prices	Producer Price Index by Commodity: Final Demand: Personal Consumption Goods	2.002
16	CES3000000008	Labor Market	Avg Hourly Earnings : Manufacturing	2.018
17	UNRATE	Labor Market	Civilian Unemployment Rate	-1.961

Penal B. $\alpha=0.00082$

No	FRED	Category	Description	t-value
1	USCONS	Labor Market	All Employees: Construction	1.993
2	RETAILx	Orders and Inventories	Retail and Food Services Sales	-1.961
3	USTPU	Labor Market	All Employees: Trade, Transportation & Utilities	-1.996
4	GS10	Interest rate and Exchange rates	10-Year Treasury Rate	-1.990
5	HWI	Labor Market	Help-Wanted Index for United States	2.065
6	WPSFD49207	Prices	Producer Price Index by Commodity: Final Demand: Finished Goods	-1.962
7	INDPRO	Output and Income	IP Index	2.049
8	UEMP15OV	Labor Market	Civilians Unemployed - 15 Weeks & Over	2.056
9	WPSID62	Prices	Producer Price Index by Commodity: Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand	2.064
10	CPIAPPSL	Prices	CPI : Apparel	2.087
11	M2SL	Money and Credit	M2 Money Stock	1.989
12	USFIRE	Labor Market	All Employees: Financial Activities	-2.011
13	GS1	Interest rate and Exchange rates	1-Year Treasury Rate	1.985
14	TWEXAFEGSMTHx	Others	Nominal Major Currencies U.S. Dollar Index (Goods Only)	-1.985
15	VIXCLSx	Stock Market	CBOE S&P 100 Volatility Index: VXO	-2.017
16	CE16OV	Labor Market	Civilian Employment	2.079
17	WPSFD49502	Prices	Producer Price Index by Commodity: Final Demand: Personal Consumption Goods	2.025
18	CES3000000008	Labor Market	Avg Hourly Earnings : Manufacturing	2.051
19	IPDCONGD	Output and Income	IP: Durable Consumer Goods	-1.963

Penal C. alpha = 0.00084

No	FRED	Category	Description	t-value
1	USCONS	Labor Market	All Employees: Construction	1.970
2	RETAILx	Orders and Inventories	Retail and Food Services Sales	-1.961
3	PCEPI	Prices	Personal Cons. Expend: Chain Index	1.981
4	USTPU	Labor Market	All Employees: Trade, Transportation & Utilities	-1.997
5	GS10	Interest rate and Exchange rates	10-Year Treasury Rate	-1.962
6	HWI	Labor Market	Help-Wanted Index for United States	2.042
7	WPSFD49207	Prices	Producer Price Index by Commodity: Final Demand: Finished Goods	-1.973
8	INDPRO	Output and Income	IP Index	2.022
9	UEMP15OV	Labor Market	Civilians Unemployed - 15 Weeks & Over	2.047
10	WPSID62	Prices	Producer Price Index by Commodity: Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand	2.037
11	CPIAPPSL	Prices	CPI : Apparel	2.065
12	M2SL	Money and Credit	M2 Money Stock	1.960
13	USFIRE	Labor Market	All Employees: Financial Activities	-2.013
14	VIXCLSx	Stock Market	CBOE S&P 100 Volatility Index: VXO	-1.989
15	CE16OV	Labor Market	Civilian Employment	2.082
16	WPSFD49502	Prices	Producer Price Index by Commodity: Final Demand: Personal Consumption Goods	2.032
17	CES3000000008	Labor Market	Avg Hourly Earnings : Manufacturing	2.045
18	IPDCONGD	Output and Income	IP: Durable Consumer Goods	-1.972
19	IPDMAT	Output and Income	IP: Durable Materials	1.962

Table 4. Top 20 nonlinear relationships between macro-financial variables and bond returns

This table reports the t-values for the 20 most significant macro-financial variables obtained from the LASSO regression using 8,128 variables. These variables include 127 macro-financial variables and 8,001 interaction variables. The interaction variables are constructed using 127 macro-financial variables. We follow Steps 4–6 introduced under 3.1 Empirical methods with some variations such as, in Step 4, we calculate the inverse of the diagonal matrix of Σ_t to generate $dz_{v,t}$. With eigenvalues obtained from conducting PCA on the data of 8,128 macro-financial variables, we calculate a Lasso coefficient multiplied by the eigenvector for each macro-financial variable, of which the time-series average and standard deviation are used to calculate a t-value of each macro-financial variable using (Fama-MacBeth) $t - value = mean / std / \sqrt{T}$. The LASSO alpha (L1 hyperparameter) is 0.00082

No	Variable Name	Description	t-value
1	UEMPLT5_USCONS	(Civilians Unemployed - Less Than 5 Weeks) * (All Employees: Construction)	1.996
2	UEMP15OV_BUSINVx	(Civilians Unemployed - 15 Weeks & Over) * (Total Business Inventories)	1.996
3	AMDMUOx_AAA	(Unfilled Orders for Durable Goods) * (Moody's Seasoned Aaa Corporate Bond Yield)	1.996
4	ISRATIOx_CES3000000008	(Total Business: Inventories to Sales Ratio) * (Avg Hourly Earnings : Manufacturing)	1.996
5	AWHMAN_NONREVSL	(Avg Weekly Hours : Manufacturing) * (Total Nonrevolving Credit)	1.996
6	CLAIMSx_CES0600000007	(Initial Claims) * (Avg Weekly Hours : Goods-Producing)	1.996
7	USGOOD_GS5	(All Employees: Goods-Producing Industries) * (5-Year Treasury Rate)	1.996
8	IPFPNSS_EXJPUSx	(IP: Final Products and Nonindustrial Supplies) * (Japan / U.S. Foreign Exchange Rate)	-1.996
9	UEMPMEAN_CPITRNSL	(Average Duration of Unemployment (Weeks)) * (CPI : Transportation)	-1.996
10	MANEMP_BAA	(All Employees: Manufacturing) * (Moody's Seasoned Baa Corporate Bond Yield)	1.996
11	DPCERA3M086SBEA_HWI	(Real personal consumption expenditures) * (Help-Wanted Index for United States)	-1.996
12	TB3MS_CES2000000008	(3-Month Treasury Bill) * (Avg Hourly Earnings : Construction)	1.996
13	HOUSTS_NONREVSL	(Housing Starts, South) * (Total Nonrevolving Credit)	1.995
14	REALLN_BAA	(Real Estate Loans at All Commercial Banks) * (Moody's Seasoned Baa Corporate Bond Yield)	-1.995
15	UEMPMEAN_HOUSTS	(Average Duration of Unemployment (Weeks)) * (Housing Starts, South)	1.995
16	IPFPNSS_PERMITS	(IP: Final Products and Nonindustrial Supplies) * (New Private Housing Permits, South (SAAR))	1.995
17	M2SL_CUSR0000SA0L2	(M2 Money Stock) * (CPI : All items less shelter)	1.995
18	UEMPMEAN_HOUST	(Average Duration of Unemployment (Weeks)) * (Housing Starts: Total New Privately Owned)	1.995
19	IPFPNSS_HOUSTW	(IP: Final Products and Nonindustrial Supplies) * (Housing Starts, West)	1.995
20	PAYEMS_PERMITNE	(All Employees: Total nonfarm) * (New Private Housing Permits, Northeast (SAAR))	-1.995

Appendix I. Five-step procedure to balance the panel of macro-financial variables

Step 1. Use the *tcode* from McCracken and Ng (2016) to transform the data.

Step 2. Normalize the outcome from the first step because “observations that are missing are initialized to the unconditional mean based on the non-missing values (which is zero since the data are demeaned and standardized) so that the panel is re-balanced (McCracken and Ng, 2016).”

Step 3. Use the generated panel data to obtain factors and loadings before rewriting the missing values with the estimates of the lambda times factor.

Step 4. Use the standard deviation and mean estimates obtained in the process of normalization in Step 2 to inverse the normalization to obtain the original data form.

Step 5. Repeat Step 2 to 4 until missing values do not change.

Appendix II. List of macro-financial variables

The table lists all 127 macro-financial variables along with the variable names, descriptions, and *tcodes*, following Ludvigson and Ng (2009). The *tcode* column denotes the following data transformation for a series x :

1. No transformation
2. Δx_t
3. $\Delta^2 x_t$
4. $\log(x_t)$
5. $\Delta \log(x_t)$
6. $\Delta^2 \log(x_t)$
7. $\Delta \left(\frac{x_t}{x_{t-1}} - 1.0 \right)$

Category	FRED	Description	tcode
Output and Income	IPDCONGD	IP: Durable Consumer Goods	5
	IPFUELS	IP: Fuels	5
	IPBUSEQ	IP: Business Equipment	5
	IPDMAT	IP: Durable Materials	5
	IPNCONGD	IP: Nondurable Consumer Goods	5
	IPFPNSS	IP: Final Products and Nonindustrial Supplies	5
	IPNMAT	IP: Nondurable Materials	5
	IPCONGD	IP: Consumer Goods	5
	IPMAT	IP: Materials	5
	IPFINAL	IP: Final Products (Market Group)	5
	INDPRO	IP Index	5
	RPI	Real Personal Income	5
	IPB51222S	IP: Residential Utilities	5
	IPMANSICS	IP: Manufacturing (SIC)	5
	W875RX1	Real personal income ex transfer receipts	5
CUMFNS	Capacity Utilization: Manufacturing	2	
Labor Market	UNRATE	Civilian Unemployment Rate	2
	DMANEMP	All Employees: Durable goods	5
	USCONS	All Employees: Construction	5
	AWHMAN	Avg Weekly Hours : Manufacturing	1
	UEMP5TO14	Civilians Unemployed for 5-14 Weeks	5
	USTPU	All Employees: Trade, Transportation & Utilities	5
	PAYEMS	All Employees: Total nonfarm	5
	HWIURATIO	Ratio of Help Wanted/No. Unemployed	2
	CES3000000008	Avg Hourly Earnings : Manufacturing	6
	CES2000000008	Avg Hourly Earnings : Construction	6
	CLF16OV	Civilian Labor Force	5
	NDMANEMP	All Employees: Nondurable goods	5
	CES0600000007	Avg Weekly Hours : Goods-Producing	1
	CE16OV	Civilian Employment	5
	SRVPRD	All Employees: Service-Providing Industries	5
	UEMP27OV	Civilians Unemployed for 27 Weeks and Over	5
	UEMPMEAN	Average Duration of Unemployment (Weeks)	2
	MANEMP	All Employees: Manufacturing	5
	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	5
	CLAIMSx	Initial Claims	5
	UEMP15T26	Civilians Unemployed for 15-26 Weeks	5
	UEMP15OV	Civilians Unemployed - 15 Weeks & Over	5
	USFIRE	All Employees: Financial Activities	5
	USGOOD	All Employees: Goods-Producing Industries	5
	USGOVT	All Employees: Government	5
	USTRADE	All Employees: Retail Trade	5
	CES0600000008	Avg Hourly Earnings : Goods-Producing	6
	USWTRADE	All Employees: Wholesale Trade	5
	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing	2
	CES1021000001	All Employees: Mining and Logging: Mining	5
HWI	Help-Wanted Index for United States	2	

Category	FRED	Description	tcode
Consumption and Orders	HOUSTMW	Housing Starts, Midwest	4
	HOUSTNE	Housing Starts, Northeast	4
	PERMITS	New Private Housing Permits, South (SAAR)	4
	PERMITW	New Private Housing Permits, West (SAAR)	4
	HOUST	Housing Starts: Total New Privately Owned	4
	PERMIT	New Private Housing Permits (SAAR)	4
	HOUSTW	Housing Starts, West	4
	PERMITMW	New Private Housing Permits, Midwest (SAAR)	4
	PERMITNE	New Private Housing Permits, Northeast (SAAR)	4
	HOUSTS	Housing Starts, South	4
Orders and Inventories	UMCSENTx	Consumer Sentiment Index	2
	DPCERA3M086SBEA	Real personal consumption expenditures	5
	RETAILx	Retail and Food Services Sales	5
	AMDMUOx	Unfilled Orders for Durable Goods	5
	BUSINVx	Total Business Inventories	5
	ISRATIOx	Total Business: Inventories to Sales Ratio	2
	ANDENOx	New Orders for Nondefense Capital Goods	5
	ACOGNO	New Orders for Consumer Goods	5
	CMRMTSPLx	Real Manu. and Trade Industries Sales	5
AMDMNOx	New Orders for Durable Goods	5	
Money and Credit	M1SL	M1 Money Stock	6
	DTCTHFNM	Total Consumer Loans and Leases Outstanding	6
	M2REAL	Real M2 Money Stock	5
	INVEST	Securities in Bank Credit at All Commercial Banks	6
	REALLN	Real Estate Loans at All Commercial Banks	6
	M2SL	M2 Money Stock	6
	NONBORRES	Reserves Of Depository Institutions	7
	TOTRESNS	Total Reserves of Depository Institutions	6
	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding	6
	BUSLOANS	Commercial and Industrial Loans	6
	NONREVSL	Total Nonrevolving Credit	6
CONSPI	Nonrevolving consumer credit to Personal Income	2	
Interest rate and Exchange Rates	AAA	Moody's Seasoned Aaa Corporate Bond Yield	2
	EXJPUSx	Japan / U.S. Foreign Exchange Rate	5
	T10YFFM	10-Year Treasury C Minus FEDFUNDS	1
	TB3MS	3-Month Treasury Bill	2
	GS1	1-Year Treasury Rate	2
	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS	1
	EXCAUSx	Canada / U.S. Foreign Exchange Rate	5
	BAA	Moody's Seasoned Baa Corporate Bond Yield	2
	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	5
	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS	1
	CP3Mx	3-Month AA Financial Commercial Paper Rate	2
	GS5	5-Year Treasury Rate	2
	T1YFFM	1-Year Treasury C Minus FEDFUNDS	1
	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS	1
	FEDFUNDS	Effective Federal Funds Rate	2
	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS	1
	GS10	10-Year Treasury Rate	2
	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS	1
	TB6MS	6-Month Treasury Bill	2
	EXUSUKx	U.S. / U.K. Foreign Exchange Rate	5
T5YFFM	5-Year Treasury C Minus FEDFUNDS	1	

Category	FRED	Description	tcode
Prices	CUSR0000SA0L2	CPI : All items less shelter	6
	DDURRG3M086SBEA	Personal Cons. Exp: Durable goods	6
	CPIMEDSL	CPI : Medical Care	6
	WPSFD49207	Producer Price Index by Commodity: Final Demand: Finished Goods	6
	WPSID62	Producer Price Index by Commodity: Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand	6
	CPIAUCSL	CPI : All Items	6
	CPIAPPSL	CPI : Apparel	6
	DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods	6
	CUSR0000SA0L5	CPI : All items less medical care	6
	WPSID61	Producer Price Index by Commodity: Intermediate Demand by Commodity Type: Processed Goods for Intermediate Demand	6
	CUSR0000SAS	CPI : Services	6
	OILPRICEx	Crude Oil, spliced WTI and Cushing	6
	CUSR0000SAD	CPI : Durables	6
	CPITRNSL	CPI : Transportation	6
	PCEPI	Personal Cons. Expend: Chain Index	6
	DSERRG3M086SBEA	Personal Cons. Exp: Services	6
	CPIULFSL	CPI : All Items Less Food	6
	WPSFD49502	Producer Price Index by Commodity: Final Demand: Personal Consumption Goods	6
	PPICMM	PPI: Metals and metal products	6
	CUSR0000SAC	CPI : Commodities	6
Stock Market	S&P: indust	S&P's Common Stock Price Index: Industrials	5
	S&P div yield	S&P's Composite Common Stock: Dividend Yield	2
	VIXCLSx	CBOE S&P 100 Volatility Index: VXO	1
	S&P 500	S&P's Common Stock Price Index: Composite	5
	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio	5
Others	TWEXAFEGSMTHx	Nominal Major Currencies U.S. Dollar Index (Goods Only)	5
	BOGMBASE	St. Louis Adjusted Monetary Base	6