HEC MONTRÉAL

Does Unexpected Change in Climate Risk Concerns Affect IV Surface Dynamics? A VAR Approach Using the Unexpected Media Climate Change Concern Index

by

Shihao Tong

Under the supervision of

David Ardia

Department of Decision Sciences
(Financial Engineering)

In Partial fulfillment of the requirements for
the Degree of Master of Science (M.Sc.)

April 2025

Abstract

This thesis examines the relationship between unexpected changes in climate risk concerns and the implied volatility (IV) surfaces of green and brown stocks. Using the Unexpected Media Climate Change Concern (UMC) index developed by Ardia et al. (2023), the study analyzes the temporal behavior of IV surfaces through a 5-factor model (François et al., 2022) and vector autoregressive (VAR) frameworks. The analysis draws on option data from 10 green and 10 brown firms, selected based on greenhouse gas (GHG) emission intensity and market capitalization, sourced from OptionMetrics. The results indicate no statistically significant linear association between UMC and IV for either group of firms. Additionally, incorporating UMC as an exogenous variable in VAR models associates with reduced predictive performance across several measures, including calibration accuracy and forecasting errors of surface parameters. These findings underscore the complexity of characterizing the connection between sudden climate-related concerns and implied volatility surfaces, and they suggest the need for more nuanced modeling strategies.

Acknowledgments

I would like to express my deepest gratitude to Professor David Ardia for guiding me through my master's study in Financial Engineering. This paper would never have become true without his consistent and thorough support. I am also in debt to all the professors and mentors who ever taught me during my studies at HEC Montréal, especially Professor David Benatia and Pascal François, for their inspirational support. I am also thankful to my beloved family and friends for the company of me during this journey, whether they keep staying or not thereafter. Finally, I want to extend my appreciation to the group Fin-ML, Fondation Boucaro and La Fondation HEC Montréal for their financial support and professional training. These years in Montréal will be remembered forever.

1 Introduction

Climate change has become an important source of systematic risk in the financial market. Its effect and implication have been broadly discussed in the literature (Long et al., 2022). Not only does it directly affect equities' returns (Tankov and Tantet, 2019; Schlenker and Taylor, 2021), but growing evidence also shows the potential relevance between climate change risk and the volatility of returns. Recent studies show that climate change leads to stricter regulations introduced by the government, causing uncertainty in policy, and a significant impact towards volatility has been detected (Lv and Li, 2023; Isah et al., 2023). Most existing research shows evidence of a connection between climate change risk and the realized volatility (RV) of equities.

The option implied volatility (IV) is another standard measure for the return variation. Option IV is the volatility of the underlying asset derived from the price of the option on which it is written. It is typically computed by the inverse of the well-known Black-Scholes equation (Black and Scholes, 1973). In contrast to the backward nature of RV (i.e., computed as the variance of past returns), IV is regarded as forward-looking and reflective of the market's perspective on the underlying stock return volatility, extending from the moment of trade until the option expires (Malliaris and Salchenberger, 1996). As a significant instrument for risk management and a trading indicator, however, its association with climate risk remains under-documented. The existing research is either very specific to the physical aspect of climate risk (Bertolotti et al., 2019; Kruttli et al., 2023) or focuses only on the effect of policy (Ilhan et al., 2021). There is a paucity of research examining the potential association between climate change risk and IV across general contexts.

The potential risks of climate change may not equally distribute across all industries and corporate entities. Regarding equity returns, Pástor et al. (2021) demonstrate that green stocks outperform brown stocks through the discount factor and cash flow channels in unexpected climate deterioration. Furthermore, studies have shown that brown stocks are more sensitive to climate risk change than green stocks, both in the short and long run (Li et al., 2023; Bouri et al. (2022)). This evidence of an imbalanced impact raises questions about whether investment vehicles based

on firms with different levels of 'greenness' could respond to climate risk differently.

Climate change risk is an intangible and complex phenomenon, and in sectors such as infrastructure, the limitation of data has become a barrier to conducting further studies on climate risk (Venturini, 2022). To address this issue, researchers and practitioners have developed various types of indexes to capture climate change risk and implement them successfully in different contexts, such as portfolio management and hedging (Ardia et al., 2023; Engle et al., 2020). Of particular note is the index proposed by Ardia et al. (2023), which is based on textual information collected from a corpus of news outlets at both the aggregate and thematic levels. Moreover, it verifies the hypothesis proposed by Pástor et al. (2021) that green stocks outperform brown stocks in the event of an unexpected escalation in climate change.

This paper investigates the effect of unexpected climate change risk on the IV surface of green and brown stocks, respectively. We select 10 green and 10 brown firms based on their greenhouse gas emission (GHG) intensity and then collect their corresponding option data from OptionMetrics. We then calibrate the IV surface proposed by François et al. (2022) daily for the loadings. Next, we study the temporal evolution of these loadings by estimating daily the AR, VAR, and VAR-X models up to order 2, with the UMC index being included as the exogenous variable and assessing models' performance on various dimensions. Finally, a Bayesian VAR(X) model is run on options from a single green and brown firm for robustness check to alleviate concerns about the curse of dimensionality in the VAR model.

Our predictive analysis indicates the absence of a significant linear relation between option IV and the UMC in both green and brown firms. The investigation reveals that incorporating UMC as an exogenous variable led to a decline in prediction accuracy across diverse metrics, including the prediction of IV surface parameters, the number of days that UMC results in improvement, the number of contracts' IV that have been enhanced, and the IV projected from the predicted surface across all categories based on their moneyness and time to maturity. The study also provides insights into the factors that should be considered in future research to explore the association between IV and climate change risk.

2 Literature review

2.1 Overview

In recent years, there has been a considerable increase in the discourse surrounding climate finance. A study by Long et al. (2022) reveals that most literature and publications were produced after 2016, with a subsequent and accelerated increase in output. Their clustering study indicated that publications related to green bonds and financial markets have dominated the field in recent years. Existing literature demonstrates various solid patterns between climate change risk and both asset returns (Tankov and Tantet, 2019; Schlenker and Taylor, 2021) and its volatility (Lv and Li, 2023; Isah et al., 2023; Bonato et al., 2023). However, the majority of studies on volatility focus on realized volatility. The paucity of literature focusing on the connection between climate change risk and IV on a holistic level is, therefore, striking. This section will review the theory behind the IV surface and existing approaches for its construction. Secondly, the existing literature exploring the connection between climate change risk, specifically with asset volatility, is introduced. Thirdly, we examine how researchers make it possible to quantitatively study the intangible and complex climate change risk by developing indices.

2.2 IV surface

2.2.1 Theory

Implied volatility (IV) usually refers to the volatility solved from the inverse of the Black-Scholes formula (BS) (Black and Scholes, 1973) that compute European options price as $O \equiv f(S_t, T, K, \sigma, r, y)$ where the five inputs are spot price of the underlying asset at time t (S_t), annualized time to maturity (T), the strike price (K), volatility of return of S_t (σ), risk-free rate (r) and dividend yield (y). Since O is an analytical function and, more importantly, IV is strictly increasing in σ, a one-on-one mapping exists between price O and volatility σ. The IV is then calculated by $f^{-1}(O)$ given the rest of the inputs. A notable distinction between European and American options is the former's ability to be exercised at any time prior to the expiry date,

thereby rendering the final payoff path dependent on the evolution of the underlying asset price. This characteristic precludes the existence of a direct, closed-form solution, thus making the inverse function impossible to express in terms of the option price. Nevertheless, given that most options written on a single stock are of the American type, the IV remains a challenging concept to grasp, as demonstrated in Ekström (2004) the monotonic relationship between price and volatility still holds under certain technical assumptions concerning the convexity of payoff functions of American options with underlying asset price following a more general diffusion process than GBM. Consequently, the computation of IV from American options is valid from this perspective.

American options are usually priced using Monte Carlo (MC) simulation or path generation discretely by tree model. For instance, Carriere (1996) proposed the first regression-based simulation method based on MC simulation. The intuition is to determine the option's value at each node by comparing the continuous and intrinsic values and then taking the average discounted option value as the option price at time zero. The most common method for solving for IV is the tree method, and the most prevalent one is the Cox-Ross-Rubinstein (CRR) tree model, as outlined in Cox et al. (1979). This tree model posits that underlying asset prices increase or decrease after each time step to generate the tree. Some paths rejoin at each step by establishing a convenient scale factor. Once a tree has been generated, the option at time zero can be solved backward using the same logic. OptionMetrics (2024) adapts the CRR tree model. Although not explicitly mentioned, they search over discretization of sigma. For each given sigma, the price is computed with a CRR tree. The algorithm repeats until convergence of the series of simulated prices toward market price is observed.

IV surface is a volatility function $\sigma(M, \tau)$ of moneyness M and maturity τ. Once the function has been determined, users can price any options, regardless of whether they are observable. For European options, the underlying asset price is assumed to be a geometric Brownian motion in the BS model with fixed volatility. If this assumption holds, all input into $\sigma = f^{-1}(O_{\text{market}}; S, T, K, r, y)$ will generate the same volatility and consequently lead to a flat

IV surface. In reality, this is hardly true. The IV empirically shows some stylized shapes in the market and are already well documented (Rubinstein, 1985; Goncalves and Guidolin, 2006). First, IV shows convexity over moneyness. This implies that the BM formula underestimates the out-of-the-money (OTM) options and treats the price the same as at-the-money (ATM) options. Second, the convexity could be asymmetric at the ATM level, sometimes showing a blade shape over the deep out-of-the-money (DOTM) call option. More precisely, this reflects a higher volatility for the OTM put than for the OTM call option. Third, the smile curve gradually flattens in maturity. The decrease in convexity is known as smile attenuation. Numerous studies aim to rationalize these facts. First, some researchers suggest that the smile is caused by a distribution of log-return other than normal law. For instance, Eberlein et al. (1998) imposed a hyperbolic Levy motion of the price movement, resulting in a non-normal return and evidence of a smile reduction relative to BS. Chen and Palmon (2005) use the index historical returns as an empirical distribution to price the options and successfully eliminate the smile. Furthermore, Duan (1999) hypothesizes that the underestimation of option price by BS is attributable to the fat-tail distribution of actual underlying returns. Merton (1976) introduces jumps with intensity and frequency into the price process, thereby placing greater emphasis on the tail risk than BS. Furthermore, other evidence has been presented which argues for BS's incompatibility due to the volatility's stochastic nature. For example, Bates (1996) has shown that the smile with extra kurtosis could be explained by modeling the volatility as a mean-reverting square root process embedded in the price moment.

2.2.2 Models

Empirical observations were the foundation upon which researchers build surface models to depict the distribution of IV over maturity and moneyness. Pioneering work was undertaken by modeling volatility as a quadratic form of maturity and strike price Dumas et al. (1998). A series of polynomial functions involving volatility and maturity up to the second order are tested. They show that both time and strike are significant factors, and the model with a quadratic form of

time appears to eliminate the most prediction errors along the maturity dimension. Building on these findings, Goncalves and Guidolin (2006) proposed an alternative functional form for log-IV which, instead of being a function of the strike price, takes moneyness and maturity as the input and models the logarithm of IV to avoid obtaining negative values. Other than the parametric method, Fengler et al. (2007) developed a semi-parametric factor model (SFM) to describe the surface. Innovations of this method lie in estimating the basis function that are solved numerically from least square criteria in a neighborhood of points of maturity and moneyness. To improve interoperability, based on surface form Goncalves and Guidolin (2006), Chalamandaris and Tsekrekos (2011) introduced explicitly seven factors based on maturity and moneyness of options, which are the level, left and right smile, short and medium term structure, and left and right smile attenuation. Moreover, by extending the idea of non-constant volatility to randomness, Heston (1993) introduced the famous stochastic volatility model by setting volatility as a mean-reverting process. Once the price is determined by a stochastic volatility model, the IV surface is obtained by utilizing the inverse of the BS equation. The model we pick to describe the IV surface is the one proposed by François et al. (2022). This factor model explicitly captures the stylized facts discussed above for S&P 500 index options. Several advantageous properties characterize the surface. Primarily, the expression is remarkably uncomplicated. The surface can be constructed by a mere five parameters in a linear form. Moreover, as a factor model, the surface can be calibrated irrespective of the underlying price stochastic process. The five factors are also selected with their distinct meanings, facilitating interpretation and illustrating the potential channels through which external variables could affect the IV surface. Last but not least, it is smooth and twice differentiable with nice asymptotic behavior. The surface is detailed in Section 3.3.

The surfaces under discussion are all static in nature, as they can only be calibrated by collecting observable data from the market at each point in time. Dumas et al. (1998) showed evidence that the strike price and maturity have prediction power of the IV surface. It is convenient to use Goncalves and Guidolin (2006) and Fengler et al. (2007)s' model since if a parametric model

is applied to the data, once the coefficients are determined, users are then able to pin down the entire IV surface and to price any options with given maturity and moneyness. Thus, predicting the surface is equivalent to predicting the coefficients in a factor model. Goncalves and Guidolin (2006) used the vector autoregressive model (VAR) to describe the potential inner structure of the loadings of their surface. The VAR model is also utilized by Fengler et al. (2007) on their factor loadings. Predictable evidence is shown in both studies using VAR. In this paper, we rely also on the VAR formulation to study the temporal dynamic of the factor loadings. For detail, see Section 3.4.

2.3 Climate change risk, asset return and volatility

Effects of climate risk on equity returns are already extensively studied (Tankov and Tantet, 2019; Schlenker and Taylor, 2021; Pástor et al., 2021), and its connection with volatility is also getting noticed. Many studies show effects towards equities RV. Recent studies show that stricter regulations introduced by the government due to climate change bring policy uncertainty to the market, and a significant impact towards volatility has been detected (Lv and Li, 2023; Isah et al., 2023). Bonato et al. (2023) 's study shows evidence of the prediction power of climate risk factor over stock volatility even at the state level across the U.S. IV also saw solid climate change risk effect through multiple channels in the option market. Ilhan et al. (2021)'s study directly uses IV features to show evidence that policy uncertainty caused by climate change has been priced in the option market and raised the hedging cost for carbon-intense business. Regarding physical risk, Kruttli et al. (2023) found that unexpected extreme weather leads to IV surge of companies having business in the impacted region and also evidence its persistent effect on slow IV recovery back to its original level. According to Bertolotti et al. (2019), exposure to hurricanes and wildfire leads to around 6% and 4% shock in option IV for equity in utility sectors.

Climate change issues are often related to the environment, which is easily affected by carbon emissions, and its effects could be unbalanced by industries and firms. Regarding equity returns, Pástor et al. (2021) shows that green stocks outperform brown stocks through discount factors

and cash flow when climate conditions worsen unexpectedly. Investors will lower future cash flow and require higher discount rates from brown firms, leading to an immediate decrease in stock price. This statement is later verified by Ardia et al. (2023). Regarding volatility, Bouri et al. (2022) shows that green stocks have lower volatility relative to brown stocks during times of high policy uncertainty by regressing the climate policy uncertainty index on the green and brown volatility ratio. In the long run, evidence shows volatility for brown stocks significantly increases with climate factors, while this is not observed regarding green stocks, and the negative climate news index has the most significant impact on green and brown return correlation (Li et al., 2023). In this study, we empirically examine the relationship between climate change risk and implied volatility (IV) in green versus brown stocks, testing whether the return differential documented by Pástor et al. (2021) similarly manifests in volatility dynamics.

2.4 Climate risk measurement in financial market

Measuring climate risk presents significant methodological challenges due to the complex system of interdependent variables characterizing climate change, compounded by its inherent externality that complicates economic and financial impact assessments (Dawson, 2015). Sector-specific data limitations further obstruct accurate measurement (Verschuur et al., 2024). The literature broadly classifies climate risks into physical risks—such as operational disruptions from extreme weather—and transitional risks stemming from technological shifts toward renewable energy (Venturini, 2022). Recent methodological advances have developed two distinct approaches to proxy climate risk in financial markets (Guo et al., 2023): Meteorology-based indices (Bressan and Romagnoli, 2021), while valuable, exhibit inherent bias toward physical risks through their reliance on environmental data like temperature records. Conversely, text-based indices (Engle et al., 2020) leverage natural language processing of media content, employing lexicon-filtered frequency analysis to quantify climate concerns. Engle et al. (2020) pioneers this approach using New York Times content to construct climate-hedged portfolios, while Gavriilidis (2021)'s Climate Policy Uncertainty Index later enabled Bouri et al. (2022) to examine policy-volatility

linkages. Building on this paradigm, we utilize the Media Climate Change Concerns (MCCC) index from Ardia et al. (2023) to capture aggregate concern levels from where we derive the UMC via residual analysis as detailed in Section 3.2.

3 Methodology

3.1 Overview

Aiming to empirically investigate the relation between unexpected climate risk change and IV surface dynamic, we introduce the three main building blocks of our work in this section: UMC index (Ardia et al., 2023) for the quantification of unexpected climate change concern, a five-factor model for IV surface description (François et al., 2022) and the AR, VAR model for temporal IV surface projection. The UMC index is construed from MCCC, a textual and thematic news index based on news from leading U.S. newspapers and newswires. The IV surface model proposed by François et al. (2022) is chosen for its interpretable factors and loadings. Finally, a VAR structure allows us to capture the temporal interdependency among the five parameters and predict the IV surface over time.

3.2 Unexpected Climate change concern: the UMC index

Ardia et al. (2023) proposed, in the same spirit as Engle et al. (2020), a novel news text-based climate change index that capture a much broader and accurate news information about climate change. Ardia et al. (2023) expand the sources of text to 10 highly circulated newspapers including *Los Angeles Times, New York Times, Wall Street Journal, USA Today, Washington Post, Houston Chronicle, Chicago Tribune, Arizona Republic, New York Daily News, New York Post* and two other newswires that are Associated Press Newswires and Reuters News. For each article, a concern score is computed using a lexicon-based measure, combining the risk level (by the percentage of risk words) and negativity level (by a ratio between 0, most positive, and 1, most negative). Then, the article scores are aggregated daily by source level. Finally, an increasing

concave function is applied due to the decreasing rate of concern level increment. The MCCC index proxy the overall level of climate change risk concerns. The author runs an AR-X model with X as some control variables over rolling windows to obtain UMC and uses the one-day-ahead prediction error as UMC.

The UMC index has its advantage transited from MCCC. First, as a text-based index, it absorbs information regarding climate change with no bias. The choice of news outlet covers heterogeneous topics. Consequently, the aggregated index captures all transition, physical, and liability risks well. Second, Ardia et al. (2023) already tested the hypothesis of Pástor et al. (2021) using the UMC index that green firms outperform brown firms through the discount factor and cash flow channel when climate change concern increases unexpectedly. This effect can potentially be extended to the context of volatility modeling as well. Let the *change* in climate change concern at time t be ΔCC_t. The UMC can be represented as

$$\text{UMC}_t = \Delta CC_t - \mathbb{E}\big(\Delta CC_t \,|\, \mathcal{I}_{t-1}\big) \tag{1}$$

where we follow Ardia et al. (2023)'s approach to approximate ΔCC_t by the MCCC index that is assumed to follow and $\mathrm{AR}(1)$ process

$$\text{MCCC}_t = \mu + \rho\, \text{MCCC}_{t-1} + \epsilon_t. \tag{2}$$

We estimate (2) over a fixed-length rolling window of 1000 days and carry out a one-day-ahead forecast. The daily prediction error is then defined as the UMC index.

3.3 Static surface depiction: five-factor model

3.3.1 Surface Expression

Motivated by Chalamandaris and Tsekrekos (2011), François et al. (2022) suggest the following linear form of factorization

$$\begin{aligned}\sigma(M,\tau) \equiv\ & \underbrace{\beta_1}_{\text{long-term ATM level}} + \beta_2 \underbrace{\exp\left(-\sqrt{\tau/T_{\text{conv}}}\right)}_{\text{Time-to-maturity slope}} + \beta_3 \underbrace{\left(M\mathbb{1}_{\{M\geq 0\}} + \frac{e^{2M}-1}{e^{2M}+1}\mathbb{1}_{\{M<0\}}\right)}_{\text{Moneyness slope}} \\ & + \beta_4 \underbrace{\left(1-\exp\left(-M^2\right)\right)\log\left(\tau/T_{\max}\right)}_{\text{Smile attenuation}} + \beta_5 \underbrace{\left(1-\exp\left((3M)^3\right)\right)\log\left(\tau/T_{\max}\right)\mathbb{1}_{\{M<0\}}}_{\text{Smirk}}\end{aligned} \tag{3}$$

where τ is the annualized maturity (i.e. in terms of 252 trading days a year $\tau = \text{DTE}/252$ where DTE is the number of trading days util expire) and M is the log moneyness defined as

$$M \equiv \frac{1}{\sqrt{\tau}} \log\left(\frac{F_{t,\tau}}{K}\right) \tag{4}$$

with $F_{t,\tau}$ as the forward price of underlying asset at time t in τ years and K as the strike price, T_{max} and T_{conv} are two constant chosen based empirical observation discussed below. Based on (4), $M > 0$ implies OTM put, and $M < 0$ implies OTM call. Model (3) falls into the class of factor models where the five factors are borrowed from Chalamandaris and Tsekrekos (2011), and each factor is assigned a specific role in the parametrization. The entire surface can be determined once the loadings β_1 to β_5 are estimated.

$$\begin{aligned}
&(a)\ \lim_{M\to 0}\lim_{\tau\to\infty}\sigma(M,\tau) = \beta_1 \\
&(b)\ \lim_{\tau\to 0}\lim_{M\to 0}\sigma(M,\tau) = \beta_1 + \beta_2 \\
&(c)\ \frac{\partial}{\partial M}\sigma(M,\tau) = \begin{cases} \beta_3 + \beta_4(2M)\log\left(\frac{\tau}{T_{\text{max}}}\right)e^{-M^2}, & M > 0 \text{ (OTM Put)} \\ \beta_3 \frac{4e^{2M}}{\left(e^{2M}+1\right)^2} + \beta_4\left(2M\log\left(\frac{\tau}{T_{\text{max}}}e^{-M^2}\right)\right) & \\ \quad -\beta_5\left(3^4M^2\log\left(\frac{\tau}{T_{\text{max}}}\right)e^{(3M)^3}\right), & M < 0 \text{ (OTM Call)} \end{cases} \\
&(d)\ \frac{\partial^2}{\partial M^2}\sigma(M,\tau) = 2\beta_4\log\left(\frac{\tau}{T_{\text{max}}}\right)e^{-M^2}\left(1-2M^2\right),\ M > 0 \text{ (OTM Put)}
\end{aligned} \tag{5}$$

The role of βs can be interpreted by looking at its derivatives and asymptotic behavior. As shown in (5), (a) β_1 measures the long-term ATM time-to-maturity level, (b) β_2 measures the time-to-maturity slope of ATM IV, (c) β_3 alone can measure the moneyness slope for ATM put options (d) β_4 contribute to the smile convexity. β_5 specifically captures the loading of possible convexity of deep OTM call. Moreover, $\partial^3/\partial^2 M\partial\tau \propto 1/\tau$ so the surface reflects that moneyness convexity is more pronounced for short-term options. One dimension of time-to-maturity is convexity, which is usually more apparent for short-term options. Scaling τ with T_{conv} thus provides more convexity to options with maturity less than T_{conv}. For T_{max}, it serves for the temporal extrapolation purpose since as $\tau \to T_{\text{max}}$, both the smile attenuation and smirk (i.e. the last two terms in (3)) fade away thus no stylized fact could be observed. In this paper, we follow François et al. (2022) take $T_{\text{conv}} = 0.25$ and $T_{\text{max}} = 5$ which is bigger than the average maximum of the 20 firms option maturity 2.2 years.

3.3.2 Calibration

François et al. (2022) tested the validity of model (3) for options written on S&P500 index which are all of European style. However, our analysis is carried out on single stocks, and the corresponding options

are American style. Thus, we presume that the surface also works on American options. In Section 5.1, we will show that its performance is comparable to the results proposed by François et al. (2022) on S&P 500 index European options. Our calibration procedure closely follows François et al. (2022) using Bayesian linear regression for daily calibration for regularization by assuming a normal and independent likelihood and prior. The prior structure is implemented for regularization and to avoid identification issues (i.e., different combinations of loadings might result in the same shape of the IV surface). Let $\boldsymbol{Y}$ be the collection of observed $\sigma(M,\tau)$, $\boldsymbol{\beta}_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t}, \beta_{4,t}, \beta_{5,t})^\top$ for loadings in model (3), $\mathbf{X}$ be the information matrix for factors. The Bayesian regression is set up as $\mathbf{Y} \mid \boldsymbol{\beta} \sim \mathrm{N}\left(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}_\epsilon = \sigma^2 \boldsymbol{I}\right)$, $\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\beta}_{\text{prior}}, \boldsymbol{\Sigma}_\delta)$ which is equivalent to

$$\begin{bmatrix} \mathbf{Y} \\ \boldsymbol{\beta}_{\text{prior}} \end{bmatrix} = \begin{bmatrix} \mathbf{X} \\ \mathbf{R} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \epsilon \\ \delta \end{bmatrix}, \begin{bmatrix} \epsilon \\ \delta \end{bmatrix} \sim \mathrm{N}(\mathbf{0}, \Omega), \Omega \equiv \begin{bmatrix} \boldsymbol{\Sigma}_\epsilon & 0 \\ 0 & \boldsymbol{\Sigma}_\delta \end{bmatrix}. \tag{6}$$

In equation (6), $\boldsymbol{\beta}_{\text{prior}}$ and Σ_δ are the hyper-parameters to be defined. $\boldsymbol{\beta}_{\text{prior},t}$ are based on the $\boldsymbol{\beta}_{t-1}$ and the diagonals of Σ_δ are chosen based on construction of factor variance. More precisely, they are

$$\boldsymbol{\beta}_{\text{prior}} = \begin{bmatrix} \mathrm{ATM}_{1\mathrm{y},t} \\ \text{Slope}_{\ t} \\ \beta_{3,t-1} \\ \beta_{5,t-1} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{\Sigma}_\delta = \begin{bmatrix} \sigma^2_{\beta_1} & 0 & 0 & 0 \\ 0 & \sigma^2_{\beta_2} & 0 & 0 \\ 0 & 0 & \sigma^2_{\beta_3} & 0 \\ 0 & 0 & 0 & \sigma^2_{\beta_5} \end{bmatrix} \times 10^{-4} \tag{7}$$

with $\mathrm{ATM}_{1\mathrm{y},t}$ to be the observed one-year ATM IV and the Slope_t is the time-to-maturity slope of ATM option approximated by the one-month ATM IV observed ($\mathrm{ATM}_{1\mathrm{m},t}$) and $\mathrm{ATM}_{1\mathrm{y},t}$

$$\text{Slope}_t = \frac{\partial}{\partial \tau}\sigma(\tau, M=0) \approx \frac{\mathrm{ATM}_{1\mathrm{y},t} - \mathrm{ATM}_{1\mathrm{m},t}}{\exp(-\sqrt{4/12})} \tag{8}$$

For prior variance $\boldsymbol{\Sigma}_\delta$, first the σ_1^2 and σ^2 are simply the sample analog estimation of $\mathrm{ATM}_{1\mathrm{y},t}$ and Slope_t computed over the entire period. Second, the σ_3^2 is defined as the variance of the moneyness slope variable, which is approximated by the sample variance of the quantity (9)

$$\frac{\sigma(M=0, \tau=20/252) - \sigma(0.4, \tau=20/252)}{0.4} \tag{9}$$

Finally, no prior is assigned to β_4, and σ_5 is set to half the parameters level. For the rationale, see appendix of François et al. (2022). Since the system no longer has homogeneous variance due to the prior, $\boldsymbol{\beta}$ is estimated by the general least square (GLS) estimator

$$\widehat{\boldsymbol{\beta}}_{\text{GLS}} = \left(\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{R} \end{bmatrix}^\top \Omega^{-1} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{R} \end{bmatrix} \right)^{-1} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{R} \end{bmatrix}^\top \Omega^{-1} \begin{bmatrix} \boldsymbol{Y} \\ \boldsymbol{\beta}_{\text{prior}} \end{bmatrix}. \tag{10}$$

The Σ_δ is fixed over the entire calibration period while $\boldsymbol{\beta}_{\text{prior}}$ varies since it's bounded around the $\boldsymbol{\beta}_{\text{t-1}}$. To initialize the iterating of prior mean, we run a simple OLS model without any prior, which is $\widehat{\boldsymbol{\beta}}_1 = \widehat{\boldsymbol{\beta}_{\mathbf{1},\text{OLS}}}$, then the Bayesian linear regression is carried out daily for each firm with its prior settings.

3.4 Surface dynamic: AR, VAR, and VAR-X

We use the following notations throughout this paper. We assume to have $\mathbf{y} \in \mathbb{R}^{k\times 1}$, p and p_x to be the number of lags for endogenous and exogenous variables, $\mathbf{x} \in \mathbb{R}^{q\times 1}$ and our data has maximum T days of observations.

3.4.1 General form

The autoregressive model (AR) of order p, denoted AR(p), is a common approach for time series modeling. Its general form is given by:

$$y_t = \mu + \sum_{i=1}^{p} \alpha_i y_{t-i} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2) \tag{11}$$

For a stationary AR(p) process, future realizations can be inferred from the variable's own lagged values. To ensure stationarity, the model requires all roots of its characteristic polynomial greater than 1, that is,$\forall\lambda$ s.t. $\Phi(\lambda) = 1 - \sum_{i=1}^{p} \lambda^i = 0$, its norm greater than 1. The coefficients α can be efficiently estimated via ordinary least squares (OLS), as the model is linear in parameters. When analyzing multiple time series jointly, cross-dependencies between processes require explicit consideration. The vector autoregressive (VAR) model addresses this by combining multiple AR processes into a vectorized formulation, assuming a multivariate normal distribution for the innovation terms. The VAR(p) model is

$$\boldsymbol{y}_t = \boldsymbol{\theta} + \sum_{i=1}^{p} \boldsymbol{\Phi_i} \boldsymbol{y}_{t-i} + \boldsymbol{\epsilon}_t, \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \Sigma) \tag{12}$$

where $\boldsymbol{\Phi}$ are the matrix includes lags from other variables in a linear form, Σ is the positive semi-definite covariance matrix. If the independence of innovation is assumed between variables, Σ is diagonal. To maintain stationarity, similar to AR(p), the eigenvalue of $\boldsymbol{\Phi}(\mathrm{L}) = \mathbf{I}_\mathrm{N} - \sum_{i=1}^{p} \boldsymbol{\Phi}_i \, \mathrm{L}^i$ lie outside of the unit circle of the complex plane. As can be seen from (12), only variables' own lags are included in the system, which is endogenous. Exogenous variables can be added to study the effect of other variables external to the system but not vice versa, leading to the vector autoregressive with exogenous variable (VAR-X)

model. A full setting of the VAR-X model is

$$\boldsymbol{y}_t = \boldsymbol{\theta} + \sum_{i=1}^{p} \boldsymbol{\Phi_i} \boldsymbol{y}_{t-i} + \sum_{j=1}^{p_x} \boldsymbol{C} \boldsymbol{x}_{t-j} + \boldsymbol{\epsilon}_t, \;\; \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \Sigma) \tag{13}$$

where $\boldsymbol{x}_{t-j}$ is the vector of exogenous variables of lag j and $\boldsymbol{C}$ is its coefficient vector. The exogenous variables $\boldsymbol{x}$ are stochastic. The model can also be extended with deterministic variables like time trends.

3.4.2 Estimator

We estimate the AR model using maximum likelihood estimator (MLE) and the VAR and VAR-X models using a generalized least square (GLS) estimator.

First, the conditional MLE for AR(p) models coincide with OLS estimator asymptotically by assuming normal innovations. More precisely, rewrite model (11) into $\mathbf{y}_t = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\epsilon}$ where $\mathbf{X}$ is the information matrix and $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_p)^\top$. Then $\widehat{\boldsymbol{\alpha}} = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y}$. We carried out the MLE estimation above using the `arima` module from package `statsmodels` (Seabold and Perktold, 2010). For details on derivation, see Hamilton (2020). For VAR and VAR-X models, we need to rewrite the equations to show how the GLS works. Here, we show the formulation for the VAR-X model, and the VAR is simply by discarding the exogenous variable. Consider the transpose of equation (13). By staking all the row vectors together, we obtain the system

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\Phi} + \boldsymbol{\epsilon} \tag{14}$$

More precisely the system is

$$\underbrace{\begin{pmatrix} \mathbf{y}_T^\top \\ \mathbf{y}_{T-1}^\top \\ \cdots \\ \mathbf{y}_{p+1}^\top \end{pmatrix}}_{(T-p)\times k} = \underbrace{\begin{pmatrix} 1 & \mathbf{y}_{T-1}^\top & \cdots & \mathbf{y}_{T-p}^\top & \mathbf{x}_{T-1}^\top \\ 1 & \mathbf{y}_{T-2}^\top & \cdots & \mathbf{y}_{T-p+1} & \mathbf{x}_{T-2}^\top \\ & & \cdots & & \\ 1 & \mathbf{y}_P^\top & \cdots & \mathbf{y}_1^\top & \mathbf{x}_p^\top \end{pmatrix}}_{(T-p)\times(1+kp+q)} \underbrace{\begin{pmatrix} \boldsymbol{\mu}^\top \\ \boldsymbol{\Phi}_1^\top \\ \cdots \\ \boldsymbol{\Phi}_p^\top \\ \boldsymbol{C}^\top \end{pmatrix}}_{(1+kp+q)\times k} + \underbrace{\begin{pmatrix} \boldsymbol{\epsilon}_T^\top \\ \boldsymbol{\epsilon}_{T-1}^\top \\ \cdots \\ \boldsymbol{\epsilon}_{p+1}^\top \end{pmatrix}}_{(T-p)\times k} \tag{15}$$

This is a so-called seemingly unrelated regression (SUE) equation. The SUE under the case of a VAR model means the innovations are independent across time while they could correlate simultaneously. That's been said, we assume in the (15) that $\boldsymbol{\epsilon}_i \sim \mathrm{N}(\mathbf{0}, \Sigma), \forall i = 1, 2, \cdots, (p+1)$ and simultaneously $\boldsymbol{\epsilon}_i \perp \boldsymbol{\epsilon}_j \; \forall i \neq j$. Considering the vectorization of system (14), the assumption results in

$$\begin{aligned} \mathbf{y} &= (\mathbf{I} \otimes \mathbf{Z}) \operatorname{vect}(\boldsymbol{\Phi}) + \operatorname{vect}(\boldsymbol{\epsilon}) \\ \operatorname{vect}(\boldsymbol{\epsilon}) &\sim \mathrm{N}\left(\mathbf{0}, \Sigma \otimes \mathbf{I}_{T-p}\right) \end{aligned} \tag{16}$$

Then (16) allows for a neat form of GLS estimation which is

$$\widehat{\text{vect}(\mathbf{\Phi})} = \left[\mathbf{Z}_*' \left(\widehat{\mathbf{\Sigma}}^{-1} \otimes \mathbf{I}_T\right) \mathbf{Z}_*\right]^{-1} \mathbf{Z}_* \left(\widehat{\mathbf{\Sigma}}^{-1} \otimes \mathbf{I}_T\right) \mathbf{y} \tag{17}$$

where $\mathbf{Z}_* = \mathbf{I} \otimes \mathbf{Z}$. Each element in the covariance matrix can be estimated by their sample analogue estimator

$$\widehat{\Sigma}_{ij} = \frac{\widehat{\boldsymbol{\epsilon}}_i' \widehat{\boldsymbol{\epsilon}}_j}{T - (kp + q) - 1} \tag{18}$$

where $\widehat{\boldsymbol{\epsilon}}_\cdot$ can be estimated from

$$\hat{\boldsymbol{\epsilon}}_i = \mathbf{y}_i - \mathbf{Z}\hat{\boldsymbol{\phi}}_i \tag{19}$$

where $\hat{\boldsymbol{\phi}}_i$ is the estimation of ith column of $\mathbf{\Phi}$ from equation (14). The GLS approach is the same as estimating each equation separately by OLS due to the same information matrix used for each equation, even though no assumption of the diagonal covariance matrix is made. Consequently, its asymptotic behavior mimics the OLS estimator. For a detailed discussion about the GLS and OLS estimator on the VAR model, see A. Zellner (1962) and Moon and Perron (2016). We carry out GLS estimation using `vector_ar.VAR` module in `statsmodels` (Seabold and Perktold, 2010) package.

3.4.3 Application

For the series of options quoted for each of the 20 firms, we estimate the following model on the first difference of each calibrated β from (10) which are $\Delta\beta_{i,t} = \beta_{i,t} - \beta_{i,t-1}, i = 1, 2, \cdots, 5$ since the raw βs are not stationary for example the illustration for a single green and brown firm in Table 5. We conduct an ex-ante analysis using VAR-X models in (20) over a fixed-length 1-year (252 days) rolling window. We use the UMC index's first leg to predict the loadings change. The highest lag we consider up to order two because first, as shown in Table 5, the Ljung–Box test (Box and Pierce, 1970) shows the loadings' own lags are still significant even up to order 20 for all firms. Second, for VAR models, one more lag leads to an entire extra five-by-five matrix to be estimated—the curse of dimensionality results in vast amounts of unstable estimation results. The Bayesian VAR (BVAR) is another common approach to dealing with dimension issues in estimation. We perform the BVAR on two single firms for extra exploration, detailed in Section 5.5. The AR models are estimated using MLE, while VAR and VAR-X are estimated by OLS and

GLS, respectively. After the $\Delta\beta$s are estimated, we restore the estimation of beta by $\widehat{\beta}_{i,t} = \beta_{i,t-1} + \Delta\widehat{\beta}_{i,t}$.

$$\begin{cases} \text{AR(1)} & \Delta\beta_{i,t} = \mu_i + \alpha_i \Delta\beta_{i,t-1}, \quad i = 1, 2, \cdots, 5 \\ \text{VAR(1)} & \Delta\boldsymbol{\beta}_t = \boldsymbol{\theta} + \boldsymbol{\Phi}\Delta\boldsymbol{\beta}_{t-1} + \boldsymbol{\epsilon}_t \\ \text{VAR(1)-X} & \Delta\boldsymbol{\beta}_t = \boldsymbol{\theta} + \boldsymbol{\Phi}\Delta\boldsymbol{\beta}_{t-1} + \boldsymbol{D}\text{UMC}_{t-1} + \boldsymbol{\epsilon}_t \\ \text{AR(2)} & \Delta\beta_{i,t} = \mu_i + \alpha_{i,1}\Delta\beta_{i,t-1} + \alpha_{i,2}\Delta\beta_{i,t-2}, \quad i = 1, 2, \cdots, 5 \\ \text{VAR(2)} & \Delta\boldsymbol{\beta}_t = \boldsymbol{\theta} + \boldsymbol{\Phi}_1\Delta\boldsymbol{\beta}_{t-1} + \boldsymbol{\Phi}_2\Delta\boldsymbol{\beta}_{t-2} + \boldsymbol{\epsilon}_t \\ \text{VAR(2)-X} & \Delta\boldsymbol{\beta}_t = \boldsymbol{\theta} + \boldsymbol{\Phi}_1\Delta\boldsymbol{\beta}_{t-1} + \boldsymbol{\Phi}_2\Delta\boldsymbol{\beta}_{t-2} + \boldsymbol{D}\text{UMC}_{t-1} + \boldsymbol{\epsilon}_t \end{cases} \tag{20}$$

With the predicted loadings, we perform the analysis in the following directions. First, we show the raw calibration performance of the surface and loadings in Section 5.1 to verify the validity of (3) over American equity options. Second, we show in Section 5.2 the $\widehat{\beta}$ prediction performance of models in (20). Since the approach is to price the options by constructing the IV surface in the first place, improvement or any changes in loadings forecasting may not necessarily transfer to a better performance in IV prediction. Then, investigate the daily performance of IV across the models and focus more on comparing the VAR and VAR-X models in Section 5.3. Finally, shown in Section 5.4, we measure the performance of models on various buckets of maturity and moneyness to explore if UMC has an unbalanced effect on the sub-group of options.

4 Data

This section introduces in detail the data we use in the paper. Section 4.1 shows the MCCC and UMC index we computed. Section 4.2 gives the scope of brown and green stocks on which the options are written. Finally, Section 4.3 describes the synthesized option contracts we constructed after the necessary filtration was applied.

4.1 MCCC and UMC Indices

Figure 1 shows the MCCC and UMC indices. The top figure shows the raw MCCC index (the black curve) quoted from Ardia et al. (2020), which is already available, and the daily rolling forecast of MCCC (the

orange curve) index using the AR(1) model (2). The bottom plot shows the UMC index using (1), which is the daily difference between MCCC and its rolling window prediction.

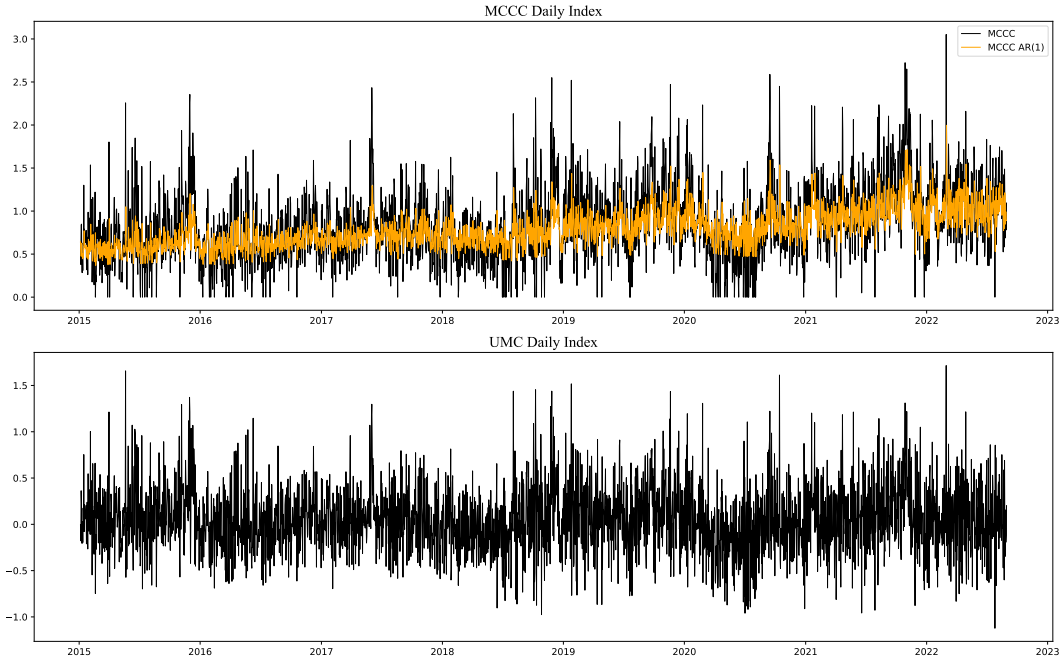

Figure 1: Daily MCCC and UMC index
The first figure shows the raw MCCC index (in black) and the one-day-ahead predicted MCCC index (in orange) by an AR(1) model. The second figure illustrates the UMC index obtained by taking the difference between the two series in the first figure.

Table 1: Summary statistics of MCCC and UMC index

NOTE: This table presents the distribution statistics of MCCC and UMC indices from 2015 Jan 5th to 2022 Aug 30th. $\mathbf{N}_+$, $\mathbf{N}_-$ and $\mathbf{N}_0$ are the number of positive, negative, and zero values observed over the period.

	Count	Mean	Std	Min	Median	Max	$\mathbf{N}_+$	$\mathbf{N}_-$	$\mathbf{N}_0$
MCCC	2795	0.85	0.43	0.00	0.81	3.05	2724	n/a	71
UMC	2795	0.06	0.38	-1.12	0.03	1.71	1495	1300	0

It is worth noticing from Table 1 that the raw MCCC sometimes hits 0. However, this does not necessarily imply zero concerns about climate change. The index itself does not differentiate between no news and entirely positive news since the concern score on the article level is zero in both cases (Ardia et al., 2023).

4.2 Green and brown firms

To investigate the differential response of options on green and brown stocks to unexpected changes in climate risk, we pick 10 green and 10 brown stocks, respectively. We follow the definition proposed

by Ardia et al. (2023), who measures the firms' greenness level by the tons of greenhouse gas (GHG) emission per 1 million dollars of revenue, which is called the GHG intensity. The data is available from the ASSET4/Refinitivon database. We first rank all S&P 500 Index constituents firms by their GHG emission intensity and pick the top (for brown) 20 and bottom 20 (for green) firms. Second, among the 20 green (brown) firms, we rank them by their market capitalization. Finally, the 10 largest ones are adopted to ensure sufficient liquidity. As seen from Table 2, most green firms are from the financial industry, and

Table 2: Green and brown stocks utilized for analysis

NOTE: The table shows the top 10 largest firms by market capitalization from the 20 firms with the lowest (for green stocks) or highest (for brown stocks) greenhouse gas (GHG) emissions intensity. GHG Emission Intensity is defined as the tons of GHG emission per 1 million $ revenue. $\overline{\mathbf{N}}_t$ is the daily average number of contracts observed to calibrate the IV surface after the filtration be applied.

Ticker	**Sector**	**Division**	**GHG Emission Intensity**	**Market Cap.** (M$)	$\overline{\mathbf{N}}_t$
		Green Firms			
AIZ	Multi-line Insurance	Insurance	27.6	5194.7	43.4
TRV	Property & Casualty	Insurance	27.6	34735.8	48.8
HUM	Managed Health	Health Care Equipment	27.4	30462.7	329.9
AFL	Life & Health Insurance	Insurance	27.2	27029.0	91.1
CNC	Managed Health	Health Care Equipment	27.1	3261.1	131.7
FRC	Regional Banks	Banks	26.8	14009.3	39.7
LNC	Life & Health Insurance	Insurance	26.3	7029.3	66.6
ELV	Managed Health	Health Care Equipment	26.2	64715.9	263.8
SIVB	Regional Banks	Banks	25.9	8970.0	117.7
ZION	Regional Banks	Banks	25.0	7128.1	70.1
		Brown Firms			
AEP	Electric Utilities	Utilities	6409.3	30958.2	66.5
NRG	Electric Utilities	Utilities	5719.4	9073.1	93.8
SO	Electric Utilities	Utilities	4838.5	45404.0	84.1
CF	Fertilizers & Agrichem	Materials	4588.0	9924.0	153.1
DUK	Multi-Utilities	Utilities	4520.7	58877.0	75.1
PPL	Electric Utilities	Utilities	4227.0	21683.2	33.6
FE	Electric Utilities	Utilities	3437.4	16418.8	46.6
ETR	Electric Utilities	Utilities	3318.3	14692.4	33.5
APD	Industrial Gases	Materials	3282.1	32676.5	88.6
D	Multi-Utilities	Utilities	2777.1	48098.5	61.1

most brown firms are from the utility sector.

4.3 Option contracts

For each of the firms in Table 2, we quote their corresponding option data from OptionMetrics database[1]. OptionMetrics employs an advanced numerical algorithm based on the tree method developed by Cox et al. (1979) to compute IV for American options. This algorithm performs a discrete search over σ, generating multiple pricing trees until convergence to market prices is achieved, as documented in OptionMetrics, 2024. We apply the following procedures to build the working data. First, we quote the daily forward price of underlying stocks for each option contract. We compute their moneyness by (4) and merge with the options data, including the trade execution date, contract expiration date, call and put indicator, lowest ask and highest bid price. The mid-price of the best ask and bid is used as the execution price for each option. Second, we filter out options according to the same rules applied by François et al. (2022), which are borrowed from G. Bakshi et al. (1997) and Azzone and Baviera (2022). More specifically, options of the following type are filtered out: (1) In-the-money (ITM) options (2) expire in 6 trading days (3) price less than $\frac{3}{8}$\$ (4) bid-ask spread larger than 175% of the mid-price (5) option with zero bid price. For each

Table 3: Descriptive statistics of green and brown OTM options data

NOTE: Summary statistics of the implied volatility of options listed in Table 2, spanning from January 5th, 2015 to Aug 30th, 2022 across different buckets of expiration days τ (in days) and log moneyness M defined in (4). A positive M indicates ITM call or OTM put. A negative M indicates OTM call or ITM put.

	Expiration Days (τ)			**Log Moneyness** (M)				
	$(-\infty, 60]$	$(60, 180]$	$(180, \infty)$	$(-\infty, -0.2]$	$(-0.2, 0]$	$(0, 0.2]$	$(0.2, 0.6]$	$(0.6, \infty)$
				Green Options				
Mean IV (%)	34.43	34.89	32.92	32.00	28.39	30.24	37.45	53.84
Standard deviation IV (%)	14.57	11.74	9.31	10.95	9.17	9.13	9.77	15.64
Median IV(%)	31.22	32.87	31.77	29.72	27.06	28.91	36.00	50.09
Number of contracts	831,578	697,844	485,212	381,989	458,455	438,726	558,727	176,737
Total number of contracts				2,014,634				
				Brown Options				
Mean IV (%)	35.36	28.51	27.84	34.19	24.83	26.65	35.77	57.16
Standard deviation IV (%)	16.47	11.80	10.22	13.85	10.79	10.32	11.22	17.38
Median IV(%)	33.09	25.45	25.38	32.37	21.15	23.39	33.66	53.98
Number of contracts	373,091	447,432	412,247	175,327	368,880	353,577	294,585	40,601
Total number of contracts				1,232,770				

[1] OptionMetrics does not provide European options data for single stocks. Only indices have European-style options available on the exchange.

firm, the average number of observed options after the above filters applied are shown in column $\overline{\mathbf{N}}_t$ in Table 2.

Although we select large-cap stocks to ensure liquidity, the number of available option contracts does not increases directly with market capitalization, and substantial variation exists in data availability for surface calibration across individual firms. The minimum daily contract count of 10 still exceeds our six estimated parameters (e.g. five β coefficients and one σ parameter in (6)), while the normal prior further safeguards against feasibility issues, ensuring reliable estimation. When constructing the β_2 prior mean in (7), occasional unobservable one-year or one-month ATM IV values—likely due to our data filters—necessitated bilinear interpolation, with missing spatial reference points (e.g., clustered ATM vol observations) addressed by carrying forward the previous day's ATM IV. The combined data for green and brown data is shown in Table 3 in different buckets of moneyness and maturity. Classification of maturity τ is based on Goncalves and Guidolin (2006) borrowed from G. S. Bakshi and Chen (1997) with day 60 and 180 to separate short, medium, and long-term time-to-maturity options. For assortment in moneyness M, we follow François et al. (2022) while we adjust the threshold from 0.8 to 0.6 to avoid too few contracts in DOTM put. As seen from Table 3, we have much more green options than brown options. Also, there is more OTM put than call options.

Observed IV on three selected days are shown in Figure 2, which shows a similar stylized shape as European option IV. In plots [1] and [3], the asymmetric IV smile (the smirk) is evident and more pronounced for short-term options. Plots [2] and [3] show that the convexity over time-to-maturity is also more prominent for short-maturity options and decreases as maturity gets longer. Plot [3] reflects the shape of the hook on OTM call options, and finally [2] shows the almost eliminated smirk (the smile attenuation) on the long-term options that expire in more than a year. We did not detect any evident jump between spots by examining some other days. This can be crucial since the surface (3) is designed to be smooth and twice differentiable. Model (3) is a sufficient and solid choice for American options IV surfaces. We will show calibration performance in Section 5.1 to verify its validity.

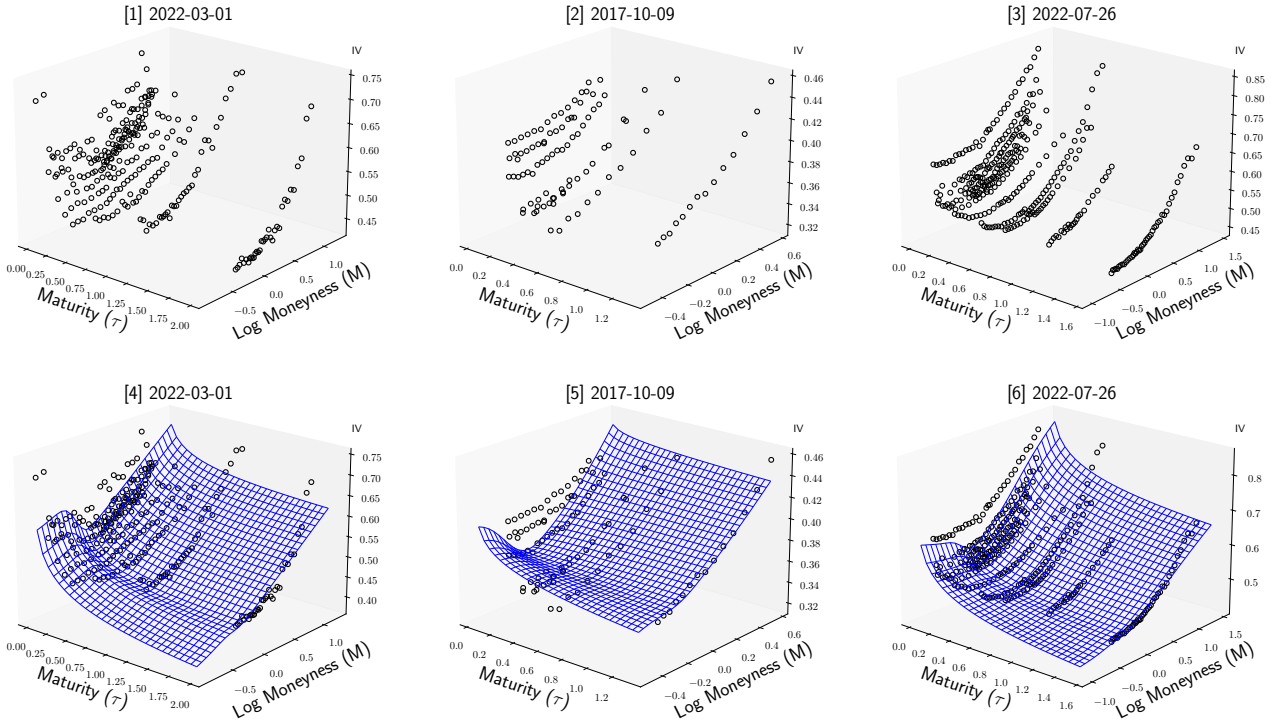

Figure 2: Observed American Option IV on 3 days with different level of UMC
The figure shows observed IVs on three days of CF, the brown firm with the highest number of average daily observations. [1] 2022-03-01 is the day with the highest UMC equals 1.715 [2] 2017-10-09 is the day with almost no unexpected climate change concern, and UMC equals 10^{-4} [3] 2022-07-26 is the day with the lowest UMC equals -1.22. The lower three figures show the calibrated surface of model (3) on the same days with observed IVs.

5 Empirical Results

5.1 Raw parameter series: β and $\Delta\beta$ calibration on American options

Recall (3) is designed according to the empirical shape of S&P 500 European options, we show evidence it also captures well the IV derived from American equity options.

The calibration performance of model (3) over options on the 20 green and brown stocks are outlined in Table 4 in terms of absolute mean percentage error (AMPE) and root mean square error (RMSE) defined as

$$\text{AMPE}_i = \frac{1}{N_i}\sum_{j=1}^{N_i}\left|\frac{\widehat{\sigma}_j - \sigma_j^{\text{obs}}}{\sigma_{\text{obs}}}\right|, \quad \text{RMSE}_i = \sqrt{\frac{1}{N_i}\sum_{j=1}^{N_i}\left(\widehat{\sigma}_j - \sigma_j^{\text{obs}}\right)^2} \tag{21}$$

where i indexes the 20 firms. $\widehat{\sigma}_j$ and σ_j^{obs} are the jth calibrated and observed IV over total N_i options across the entire period. Overall, the aggregate RMSE is 2.37%

The lower panel of Table 4 shows the mean and variance of the calibrated βs over time. We noticed that the sign of means is relatively consistent across firms, whether green or brown, except for β_2, which measures the time-to-maturity slope of ATM IV. Furthermore, it is easy to detect the loadings, which could

Table 4: Calibration performance of IV and raw β series of individual option

NOTE: The upper panel shows the RMSE and AMPE of the IV derived from the daily calibrated $\boldsymbol{\beta}$ in model (3). Agg refers to the performance of all option contracts discarding the tickers. The lower panel concerns the mean (variance) of individual options' daily calibrated βs.

	Green Options											**Brown Options**										
	AFL	AIZ	CNC	ELV	FRC	HUM	LNC	SIVB	TRV	ZION	**Agg**	AEP	APD	CF	D	DUK	ETR	FE	NRG	PPL	SO	**Agg**
AMPE (%)	4.27	4.45	4.81	3.92	8.97	4.45	2.67	2.81	3.05	3.65	4.15	2.72	2.90	3.89	3.00	3.18	3.62	3.62	3.95	3.74	3.58	3.46
RMSE (%)																						
Overall	2.36	1.96	3.35	1.75	5.30	2.01	2.17	1.99	1.56	2.46	2.37	1.01	1.35	3.34	1.17	1.20	2.93	1.95	3.32	1.82	1.74	2.33
OTM C	1.83	1.66	2.79	1.61	6.93	1.89	1.56	1.66	1.36	1.98	2.24	0.78	1.30	2.72	0.88	0.89	2.66	1.67	2.73	1.60	1.30	1.84
DOTM C	3.89	2.14	3.57	1.79	4.85	2.05	2.24	1.79	2.02	2.74	2.48	1.36	1.34	3.40	1.39	1.38	3.71	2.70	3.69	2.76	3.30	2.95
OTM P	1.63	1.55	2.70	1.45	5.45	1.74	1.44	1.38	1.26	1.93	1.99	0.72	1.25	2.57	0.86	0.82	2.47	1.54	2.76	1.49	1.20	1.75
MOTM P	2.56	2.33	3.48	1.66	4.00	2.02	1.81	1.64	1.40	2.73	2.28	1.27	1.22	3.56	1.49	1.49	2.94	2.41	3.59	2.36	2.18	2.53
DOTM P	5.72	3.44	5.32	2.54	5.22	2.69	4.56	3.01	2.45	4.86	3.36	2.65	2.14	5.89	2.88	3.13	5.46	4.01	7.97	5.40	5.13	4.79
	Calibrated $\boldsymbol{\beta}$																					
β_1	28.44	23.10	19.89	37.96	30.49	33.92	23.47	34.23	28.54	22.76		20.06	20.49	19.57	24.76	20.02	20.88	23.99	21.50	39.16	37.54	
	(.21)	(.41)	(.23)	(.47)	(.37)	(1.05)	(.25)	(.15)	(.18)	(.42)		(.11)	(.14)	(.20)	(.18)	(.12)	(.22)	(.39)	(.51)	(.38)	(.68)	
β_2	-.38	-5.22	5.91	-2.74	-1.30	-.96	1.95	-2.23	-4.04	-6.74		-5.98	-5.04	-3.10	-6.10	-4.41	-3.24	-4.22	-3.20	1.52	.04	
	(2.16)	(1.11)	(.69)	(2.05)	(.69)	(1.34)	(1.11)	(1.81)	(1.36)	(.29)		(.75)	(.65)	(.72)	(.60)	(.61)	(.71)	(.49)	(.42)	(1.58)	(1.49)	
β_3	15.29	22.71	17.99	13.69	14.92	15.00	12.40	12.46	16.34	20.97		22.26	20.13	19.30	16.86	19.39	16.95	16.05	13.04	10.22	12.92	
	(.19)	(.26)	(1.17)	(.28)	(.21)	(.14)	(.26)	(.21)	(.09)	(.16)		(.14)	(.30)	(.16)	(.18)	(.19)	(.34)	(.17)	(.23)	(.08)	(.23)	
β_4	-4.82	-8.87	-17.53	-5.22	-6.94	-5.99	-12.10	-6.00	-3.35	-5.75		-4.88	-6.73	-10.16	-4.05	-6.52	-11.59	-10.09	-9.89	-4.18	-6.48	
	(.11)	(.32)	(10.57)	(.19)	(.41)	(.14)	(1.69)	(.19)	(.05)	(.08)		(.16)	(.19)	(.42)	(.09)	(.21)	(2.03)	(.57)	(.29)	(.18)	(.31)	
β_5	-1.27	-5.57	-4.11	-0.57	-1.62	-0.45	-0.77	-0.88	-1.52	-3.25		-6.37	-4.32	-4.72	-1.74	-3.87	-3.70	-2.45	-0.31	-0.70	-0.88	
	(.01)	(.14)	(.18)	(.02)	(.04)	(.02)	(.03)	(.02)	(.01)	(.04)		(.11)	(.07)	(.12)	(.02)	(.08)	(.35)	(.07)	(.03)	(.01)	(.02)	

vary significantly for individual firms, even if they are all green (brown). To illustrate these patterns, Figure 3 displays the daily evolution of calibrated loadings $\boldsymbol{\beta}$ and their first differences $\Delta\boldsymbol{\beta}$ for one representative green stock and one brown stock. Additionally, the lower panel of Figure 2 exhibits the calibrated volatility surfaces for three characteristic trading days. Table 5 displays the summary statistics of $\Delta\boldsymbol{\beta}$ and comprehensive diagnostic testing. It reveals that all $\Delta\boldsymbol{\beta}$ series exhibit pronounced long-memory while satisfying stationarity conditions, as confirmed by Dickey-Fuller tests that uniformly reject the unit root hypothesis at the 1% significance level. Furthermore, the persistence of autocorrelation patterns established through Ljung-Box tests that maintain statistical significance up to 20 lags. These robust empirical findings collectively justify adopting a first-order autoregressive specification in modeling the $\Delta\boldsymbol{\beta}$ series.

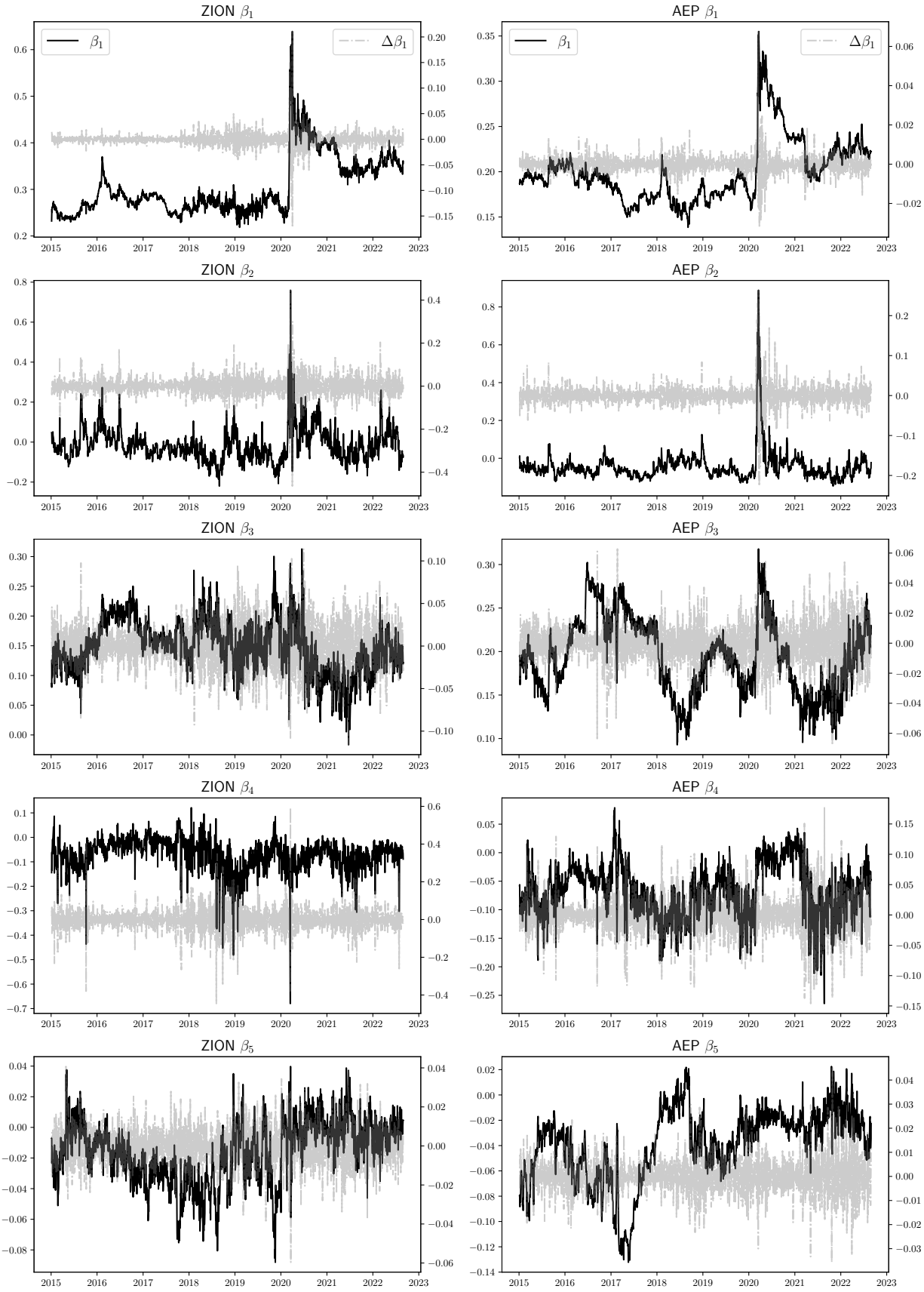

Figure 3: Calibrated $\boldsymbol{\beta}$ and $\Delta\boldsymbol{\beta}$ of ZION and AEP
Figure shows the calibrated $\boldsymbol{\beta}$ (black curve by the left axis) and $\Delta\boldsymbol{\beta}$ (gray curve by the right axis) of the brownest firm AEP and the greenest firm ZION.

Table 5: Summary statistics of $\Delta\beta$ calibration

NOTE: Tables shows the calibrated $\Delta\beta_t = \beta_t - \beta_{t-1}$ from January 6th, 2016 to Aug 30th, 2022. The first 5 measures concern the distribution of 10 mean $\Delta\beta$ of individual firms. DF is the mean of 10 Dickey-Fuller test statistics of individual firms. LB(p) averages 10 order p Ljung–Box test statistics. The $*$ suggests that $\Delta\beta$ is significant under the Ljung-Box test across all 10 firms at 1% significance level.

	Mean (10^{-4})	**Variance** (10^{-8})	**Min** (10^{-4})	**Max** (10^{-4})	**Skew**	**DF**	**LB(1)**	**LB(2)**	**LB(10)**	**LB(20)**
Green Options										
$\Delta\beta_1$	0.43	0.19	-0.01	1.49	1.67	-13.92^*	193.84^*	244.74^*	302.49^*	363.06^*
$\Delta\beta_2$	-0.28	0.05	-0.66	0.07	-0.35	-12.49^*	165.72^*	172.79^*	292.33^*	358.82^*
$\Delta\beta_3$	0.22	0.13	-0.52	0.82	-0.47	-18.33^*	245.14^*	250.85^*	268.30^*	284.56^*
$\Delta\beta_4$	0.11	4.68	-4.33	4.79	0.21	-14.02^*	343.84^*	354.11^*	396.76^*	428.11^*
$\Delta\beta_5$	0.11	0.25	-0.41	1.47	2.59	-15.06^*	139.16^*	152.88^*	167.78^*	181.33^*
Brown Options										
$\Delta\beta_1$	0.22	0.08	-0.23	0.90	1.30	-11.13^*	110.91^*	140.70^*	256.48^*	322.31^*
$\Delta\beta_2$	-0.23	0.04	-0.42	0.30	2.09	-11.15^*	147.97^*	156.31^*	260.29^*	307.74^*
$\Delta\beta_3$	0.17	0.06	-0.26	0.46	-0.61	-18.31^*	205.42^*	211.01^*	232.34^*	244.24^*
$\Delta\beta_4$	-0.28	0.31	-1.75	0.05	-2.43	-14.58^*	349.48^*	354.58^*	375.29^*	392.45^*
$\Delta\beta_5$	0.19	0.09	-0.12	0.90	1.50	-19.25^*	65.76^*	71.84^*	86.84^*	99.60^*

5.2 β : overall performance

Our first prediction study focuses on the performance concerning the loadings β. It is worth noting initially that the prediction errors associated with $\Delta\beta$ and β are identical since $\beta_t - \widehat{\beta}_t = \beta_t - (\beta_{t-1} + \widehat{\Delta\beta_t}) = \Delta\beta_t - \widehat{\Delta\beta_t}$. The prediction performance, measured in terms of RMSE and AMPE, is presented in Table 10. The results show that since the model with the best performance is either the AR(1) or AR(2) model, the inter-variable structure does not contribute to accuracy. Second, the prediction performance is better on brown firms except for $\widehat{\beta}_4$. Third, the model with the UMC index does not outperform the pure VAR model for orders 1 and 2. The performance difference between VAR and VAR-UMC models is relatively minor, suggesting that the UMC index does not enhance the prediction of factor loadings in a linear framework For illustration, we show the predicted surface on 4 days in Figure 4 for the greenest and brownest firms by VAR1 and VAR1 UMC model. The first 3 days correspond to those shown in Figure 2. while the fourth day was selected for its high UMC value while remaining pre-pandemic. The surface predicted using whether or not UMC is visually indistinguishable. Notably, for AEP (the brownest firm), the VAR and VAR-UMC surfaces become nearly identical post-pandemic, even during periods of high UMC values, as evident from the top right plot.

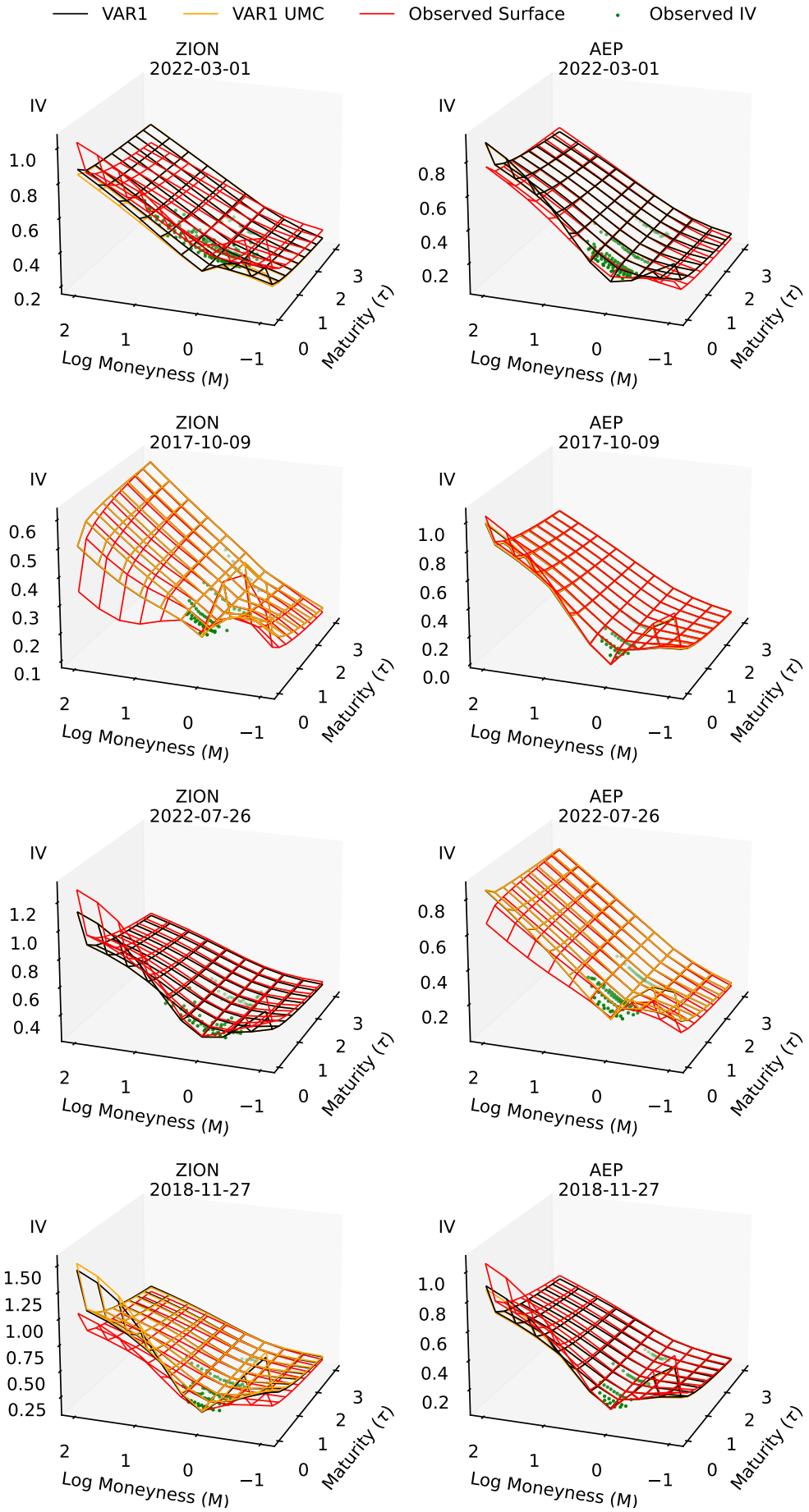

Figure 4: Calibrated and predicted IV surface on selected days

The figure shows the predicted and calibrated IV surface on four selected days. The first three rows are the days with the highest, lowest, and almost zero UMC index. The last day, Nov 27th, 2018, has a large UMC but before the start of the pandemic.

Table 6: Aggregated forecasting performance of β across models

NOTE: The table shows the prediction performance of β from January 6th, 2016, to August 30th, 2022. Each cell represents the average RMSE (AMPE) across the 10 stock options. For each $\widehat{\beta}$, the model that produces the lowest RMSE (AMPE) is highlighted to be bold.

RMSE (10^{-2})

	Model	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$
Green						
	AR1	1.33	**4.72**	**2.18**	3.82	0.87
	VAR1	1.35	4.85	2.21	3.89	0.89
	VAR1 UMC	1.36	4.86	2.21	3.90	0.89
	AR2	**1.33**	4.75	2.15	**3.71**	0.86
	VAR2	1.39	5.01	2.22	3.85	0.89
	VAR2 UMC	1.39	5.03	2.22	3.86	0.89
Brown						
	AR1	**0.96**	**3.42**	1.83	4.20	0.76
	VAR1	0.98	3.54	1.87	4.27	0.76
	VAR1 UMC	0.98	3.55	1.88	4.27	0.77
	AR2	0.97	3.46	**1.81**	**4.09**	**0.75**
	VAR2	1.08	3.73	1.89	4.20	0.77
	VAR2 UMC	1.08	3.74	1.89	4.21	0.77

AMPE

	Model	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$
Green						
	AR1	2.05	**2.20**	2.97	**4.56**	3.05
	VAR1	2.51	2.78	4.50	6.00	4.99
	VAR1 UMC	2.57	2.77	4.49	6.14	4.63
	AR2	**2.03**	2.24	**2.82**	6.16	**2.78**
	VAR2	3.01	2.94	4.12	7.12	5.70
	VAR2 UMC	3.08	2.92	4.18	6.99	5.31
Brown						
	AR1	**2.17**	3.27	**2.23**	**2.70**	**2.37**
	VAR1	2.89	5.80	2.99	2.87	2.82
	VAR1 UMC	2.87	6.71	3.01	2.88	3.14
	AR2	2.43	**3.12**	2.41	3.06	2.86
	VAR2	3.47	5.71	2.99	3.38	3.90
	VAR2 UMC	3.43	6.31	3.03	3.39	4.19

5.3 IV: Daily and overall performance

We present the results of IV prediction performance. Figure 7 displays the initial findings regarding daily performance. The gray and black bar (with scales on the right axis) represent

$$d_t = \text{RMSE}_t^{\text{VAR(p)}} - \text{RMSE}_t^{\text{VAR(p)-UMC}}, \quad p = 1, 2 \tag{22}$$

the daily difference in RMSE between the VAR and VAR-UMC models computed over IV. A positive value of d_t indicates superior performance of the VAR-UMC model on day t, while a negative value favors the VAR model. The total number of days and option contracts, denoted as N_t and N along with their percentages, are summarized in Table 7. To illustrate the cumulative trend, we show the blue curve defined by

$$s_t = \sum_{i=1}^{t} d_i \tag{23}$$

which tracks the accumulated RMSE difference between the models. The first two plots compare the VAR(1) and VAR(1)-UMC models for green and brown options, while the bottom two plots present results for the VAR(2) and VAR(2)-UMC models. For options written on green firms, the UMC contribution to

VAR(1) prediction accuracy is minimal, though slightly more apparent in VAR(2). In contrast, brown firms demonstrate greater sensitivity to UMC, with the enhanced model outperforming VAR(1) on 44.3% of days and VAR(2) on 45% of days. This consistent improvement across both model orders suggests UMC contains particularly relevant information for brown firm volatility prediction, while its value for green firms remains limited.

The comprehensive performance metrics are presented in Table 7, which include the mean and variance of prediction errors $\bar{\epsilon}$ for each model. To facilitate the Diebold-Mariano test, we additionally report transformed error metrics: $\bar{\epsilon}_S$ (mean squared error) and $\bar{\epsilon}_A$ (mean absolute error) computed across all option contracts. Our analysis reveals that the VAR-UMC model does not demonstrate statistically significant superiority over the VAR model for either green or brown options, regardless of model order (1 or 2) or error metric (RMSE or AMPE). Notably, while the number of options demonstrating improvement for green firms shows only marginal growth, we observe a substantial 30% increase in the number of days with enhanced performance.

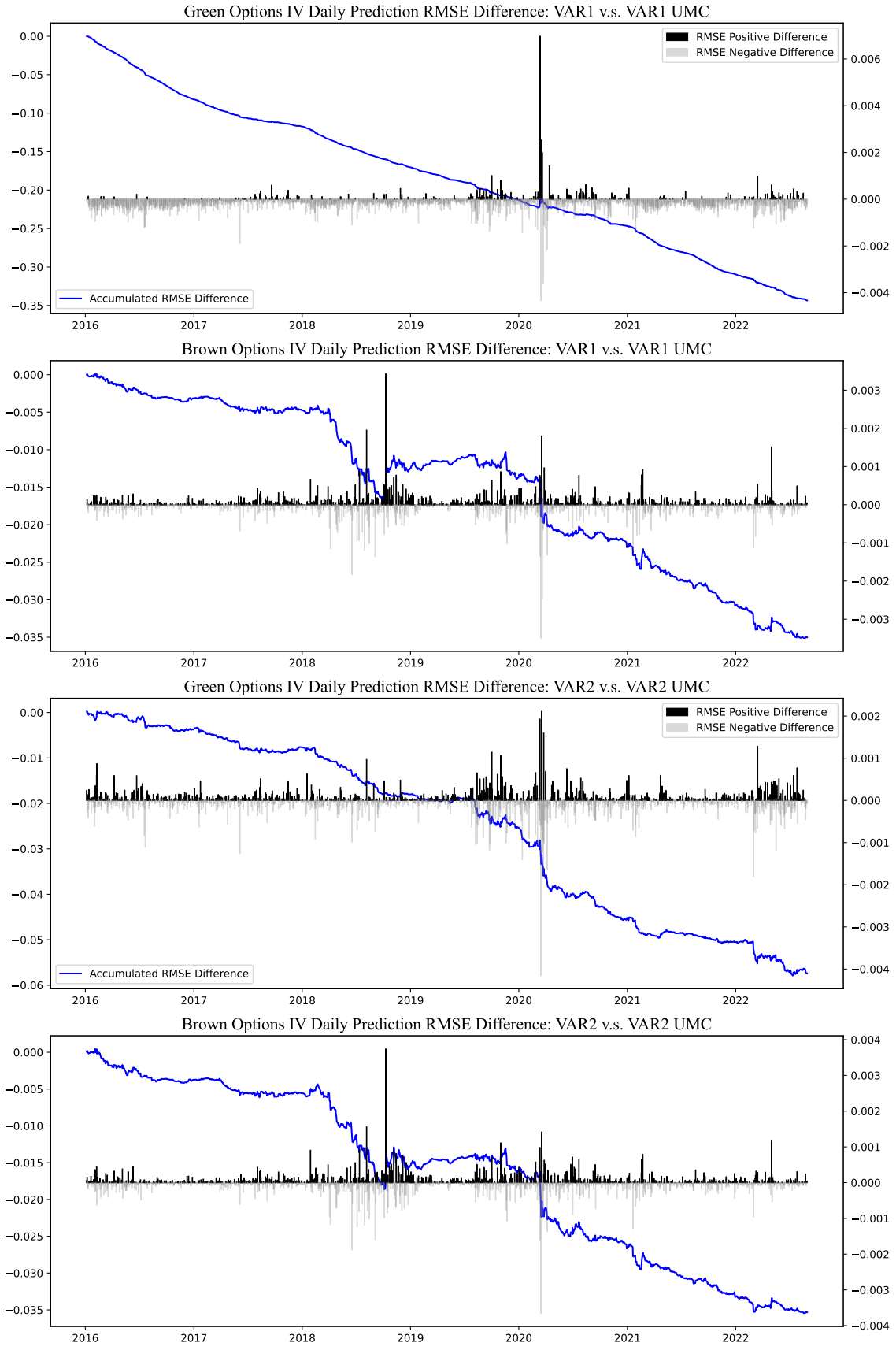

Figure 5: Daily RMSE difference between models
Accumulated RMSE difference is defined in (22). A positive difference implies a better performance of the model that includes UMC as an external regressor over the simple VAR model. The first two figures concern the VAR(1) model, and the bottom two concern the VAR(2), all estimated by OLS.

Table 7: Out of sample prediction performance across difference models

NOTE: Prediction performance comparison across models over the entire sample of green and brown options from January 6th, 2016 to Aug 30th, 2022. $\bar{\epsilon}$ is the error mean. $\bar{\epsilon}_s$ and $\bar{\epsilon}_A$ are then mean of square and absolute error respectively. N ($N\%$) is the number (percentage) of contracts that model outperform its benchmark, where we compare VAR1 to VAR1 UMC and VAR2 to VAR2 UMC. Similarly, N_T ($N_T\%$) is the number of days the model with UMC beat the pure VAR model. DM is the Diebold-Mariano statistics to justify the performance regarding root mean square error (RMSE) and mean absolute error (MAE).

Model	$\bar{\epsilon}\,(10^{-4})$	$\bar{\epsilon}_S\,(10^{-4})$	$\bar{\epsilon}_A(10^{-2})$	N	N_T	RMSE(10^{-2})	MAE (10^{-2})
	$\sigma^2(10^{-4})$	$\sigma^2(10^{-4})$	$\sigma^2(10^{-2})$	N%	$N_T\%$	DM	DM
		Green					
AR1	-9.098	9.703	1.838			3.115	1.838
	(9.695)	(.553)	(.063)				
VAR1	-8.734	9.934	1.853			3.152	1.853
	(9.927)	(.561)	(.065)				
VAR1 UMC	-8.785	9.960	1.856	989,913	221	3.156	1.856
	(9.953)	(.554)	(.652)	(.491)	(.132)	(16.338)	(41.154)
AR2	-9.418	9.735	1.840			3.120	1.840
	(9.726)	(.565)	(.063)				
VAR2	-8.729	10.163	1.865			3.187	1.866
	(10.155)	(.608)	(.067)				
VAR2 UMC	-8.699	10.192	1.869	990,041	733	3.193	1.869
	(10.185)	(.603)	(.067)	(.491)	(.438)	(19.784)	(43.061)
		Brown					
AR1	-2.980	9.170	1.475			3.028	1.475
	(.917)	(.120)	(.699)				
VAR1	-2.090	9.480	1.493			3.079	1.493
	(.948)	(.131)	(.725)				
VAR1 UMC	-2.000	9.500	1.495	605,464	744	3.082	1.495
	(.950)	(.131)	(.726)	(.490)	(.443)	(7.17)	(24.05)
AR2	-3.360	9.290	1.480			3.047	1.480
	(.929)	(.122)	(.710)				
VAR2	-2.200	9.880	1.511			3.144	1.511
	(.988)	(.145)	(.760)				
VAR2 UMC	-2.030	9.900	1.513	605,328	748	3.147	1.513
	(.990)	(.145)	(.761)	(.491)	(.450)	(7.07)	(23.91)

5.4 IV: RMSE in different maturity & log-moneyness buckets

Our next step takes a closer look at the RMSE performance of options in different buckets categorized by maturity and (OTM) log-moneyness level. Results are summarized in Table 8. The results are consistent with the performance of loadings' prediction. First, the AR(1) and AR(2) specifications demonstrate superior performance across all maturity-moneyness buckets. Notably, incorporating the UMC index fails to

Table 8: Prediction RMSE (10^{-2}) across difference buckets

NOTE: This table presents the performance (measured by RMSE) comparison across models for options on green/brown stocks from January 6th, 2016, to August 30th, 2022 by maturity and log-moneyness buckets. The lowest RMSE is in bold.

Option Type	log Moneyness (M)	Maturity (τ)	Count	AR1	VAR1	VAR1 UMC	AR2	VAR2	VAR2 UMC
			Green Options						
Call	$-0.2 < M \leq 0$	$180 < \tau$	117,116	**1.476**	1.487	1.490	1.487	1.505	1.507
		$60 < \tau \leq 180$	141,583	2.260	2.275	2.275	**2.257**	2.290	2.290
		$\tau \leq 60$	199,755	**3.558**	3.593	3.597	3.568	3.628	3.633
	$M \leq -0.2$	$180 < \tau$	81,631	**2.207**	2.236	2.243	2.222	2.278	2.282
		$60 < \tau \leq 180$	143,695	2.442	2.458	2.468	**2.423**	2.448	2.497
		$\tau \leq 60$	156,663	**4.889**	4.955	4.956	4.904	5.011	5.015
Put	$0 < M \leq 0.2$	$180 < \tau$	109,174	**1.558**	1.567	1.570	1.563	1.580	1.582
		$60 < \tau \leq 180$	133,216	1.970	1.988	1.988	**1.968**	2.005	2.005
		$\tau \leq 60$	196,337	**3.223**	3.260	3.263	3.230	3.299	3.304
	$0.2 < M \leq 0.6$	$180 < \tau$	145,417	**2.008**	2.020	2.022	2.010	2.034	2.036
		$60 < \tau \leq 180$	195,538	2.132	2.152	2.156	**2.127**	2.170	2.172
		$\tau \leq 60$	217,772	**3.837**	3.877	3.881	3.838	3.903	3.909
	$M > 0.6$	$180 < \tau$	31,875	**3.430**	3.439	3.444	3.431	3.466	3.470
		$60 < \tau \leq 180$	83,812	**3.176**	3.227	3.236	3.170	3.260	3.267
		$\tau \leq 60$	61,050	**6.181**	6.292	6.308	6.198	6.439	6.455
			Brown Options						
Call	$-0.2 < M \leq 0$	$180 < \tau$	125,240	**1.238**	1.256	1.259	1.244	1.299	1.302
		$60 < \tau \leq 180$	128,550	**1.454**	1.471	1.473	1.454	1.551	1.552
		$\tau \leq 60$	114,890	**3.689**	3.722	3.721	3.719	3.764	3.763
	$M \leq -0.2$	$180 < \tau$	59,916	**2.094**	2.123	2.126	2.137	2.213	2.219
		$60 < \tau \leq 180$	61,289	**2.846**	2.867	2.875	2.883	3.088	3.101
		$\tau \leq 60$	54,122	**6.529**	6.601	6.598	6.572	6.659	6.659
Put	$0 < M \leq 0.2$	$180 < \tau$	106,518	**1.356**	1.374	1.376	1.360	1.419	1.419
		$60 < \tau \leq 180$	128,490	**1.418**	1.444	1.451	1.414	1.490	1.500
		$\tau \leq 60$	118,569	**3.279**	3.317	3.318	3.300	3.336	3.337
	$0.2 < M \leq 0.6$	$180 < \tau$	109,641	**1.970**	1.988	1.991	1.974	2.030	2.033
		$60 < \tau \leq 180$	110,835	1.997	2.045	2.050	**1.995**	2.108	2.113
		$\tau \leq 60$	74,109	**5.181**	5.247	5.252	5.215	5.306	5.310
	$M > 0.6$	$180 < \tau$	10,932	**4.531**	4.570	4.580	4.547	4.620	4.614
		$60 < \tau \leq 180$	18,268	**4.391**	4.647	4.662	4.397	4.907	4.919
		$\tau \leq 60$	11,401	**9.592**	10.070	10.090	9.639	10.342	10.360

enhance the predictive accuracy of VAR models for either green or brown options. Second, the prediction errors for all models are much higher for options with shorter time-to-maturity than those that expire in more than 180 days. The prediction error gets bigger on the moneyness dimension when options are more OTM. This is more pronounced for deep OTM (i.e., with $M > 0.6$) put options on brown firms than on

green firms. All models perform better in each bucket of brown options than green options with very few exceptions.

5.5 A Bayesian VAR trail on options of two stocks

We present the empirical results from implementing a Bayesian VAR model for the firms with the highest and lowest GHG intensity in our sample - AEP (brownest) and ZION (greenest). The methodological framework and theoretical justification for this approach are thoroughly discussed in Appendix A.

The overall performance of IV is shown in Table 9. Due to computational efficiency, we only test the model of order 1. We compare the BVAR and BVAR UMC models to the simple VAR(1) model estimated by GLS. It can be seen that none of the VAR(1) UMC, BVAR(1) and BVAR(1)-UMC models are significantly better than the VAR(1) model in terms of either RMSE or MAE regardless of options on green or brown stock. Especially for the options on green stock ZION, the RMSE increased dramatically after adding a normal-diffuse structure (e.g., compare BVAR1 UMC to VAR1 UMC, RMSE increases by about 4% from 3.083 to 3.204).

Table 10 presents the β prediction performance across models. Key findings reveal that while the AR(1) specification maintains superior RMSE performance for both ZION and AEP, the BVAR(1)-UMC model demonstrates notable advantages. It achieves the lowest AMPE for all β coefficients in ZION, and it shows better predictive accuracy for β_2 and β_3 in AEP. This AMPE advantage aligns with the model's hyperparameter calibration objective. Figure 6 illustrates the cumulative prediction differences for β_1 to β_5 across three competing specifications (VAR(1), VAR(1)-UMC, and BVAR(1)-UMC). The COVID-19 market shock in March 2020 significantly impacted all models' β_1 and β_2 predictions, with the BVAR-UMC model exhibiting stronger post-shock recovery for both firms. However, the Bayesian prior structure yields mixed results for AEP - while substantially improving β_5 predictions, it simultaneously degrades β_3 estimation accuracy relative to alternative specifications.

Table 9: Out of sample prediction performance on AEP and ZION

NOTE: Prediction performance comparison of the greenest firm, ZION and brownest firm, AEP January 6th, 2016 to Aug 30th, 2022. $\bar{\epsilon}$ is the error mean. $\bar{\epsilon}_s$ and $\bar{\epsilon}_A$ are then mean of square and absolute error respectively. N ($N\%$) is the number (percentage) of contracts that model outperform the benchmark VAR1. Similarly, N_T ($N_T\%$) is the number of days each model beats the VAR1 model. DM is the Diablo Mariano statistics to justify the performance regarding root mean square error (RMSE) and mean absolute error (MAE).

Model	$\bar{\epsilon}\,(10^{-4})$	$\bar{\epsilon}_S\,(10^{-4})$	$\bar{\epsilon}_A(10^{-2})$	N	N_T	RMSE(10^{-2})	MAE (10^{-2})
	$\sigma^2(10^{-4})$	$\sigma^2(10^{-4})$	$\sigma^2(10^{-2})$	N%	$N_T\%$	DM	DM
ZION (green)							
AR1	-.825	9.428	1.675			3.070	1.675
	(9.428)	(.980)	(.066)				
VAR1	-.277	9.480	1.686			3.079	1.686
	(9.480)	(.768)	(.066)				
VAR1 UMC	-.310	9.504	1.689	57,519	771	3.083	1.689
	(9.504)	(.787)	(.067)	(.490)	(.460)	(5.681)	(8.176)
BVAR1	-.718	10.392	1.748	52,424	648	3.224	1.748
	(10.392)	(1.309)	(.073)	(.462)	(.387)	(5.989)	(20.154)
BVAR1 UMC	-.786	10.265	1.742	54,562	647	3.204	1.742
	(10.265)	(1.538)	(.072)	(.464)	(.386)	(4.139)	(18.436)
AEP (brown)							
AR1	-.579	2.339	.847			1.530	0.847
	(2.339)	(.039)	(.016)				
VAR1	.037	2.423	.855			1.557	0.855
	(2.423)	(.038)	(.017)				
VAR1 UMC	.199	2.429	.856	54,993	814	1.559	0.856
	(2.430)	(.038)	(.017)	(.494)	(.486)	(4.146)	(7.763)
BVAR1	-.878	2.459	.867	53,604	728	1.568	0.867
	(2.459	(.038)	(.017)	(.481)	(.435)	(2.048)	(8.870)
BVAR1 UMC	-.649	2.591	.898	50,804	609	1.610	0.898
	(2.591)	(.048)	(.017)	(.456)	(.364)	(7.674)	(26.180)

Table 10: Aggregated forecasting performance of β across models

NOTE: This table presents the out-of-sample prediction performance of factor loadings β for the period January 6, 2016 to August 30, 2022, comparing options on green stock ZION and brown stock AEP. For each $\widehat{\beta}$, the model that produces the lowest RMSE (AMPE) is highlighted in boldface.

RMSE (10^{-2})

Model	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$
ZION (green)					
AR1	**1.33**	**4.29**	**2.45**	**5.39**	**0.94**
VAR1	1.34	4.42	2.51	5.46	0.95
VAR1 UMC	1.34	4.43	2.52	5.47	0.95
BVAR1 UMC	1.46	4.65	2.70	6.14	0.99
AEP (brown)					
AR1	**0.47**	**2.50**	**1.24**	**2.83**	**0.72**
VAR1	0.47	2.58	1.25	2.85	0.73
VAR1 UMC	0.48	2.59	1.26	2.86	0.73
BVAR1 UMC	0.51	2.71	1.40	3.28	0.77

AMPE

Model	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$
ZION (green)					
AR1	2.47	2.98	2.08	2.16	2.02
VAR1	3.57	2.97	2.22	2.25	2.35
VAR1 UMC	3.78	2.98	2.03	2.31	2.36
BVAR1 UMC	**1.69**	**1.75**	**1.74**	**1.94**	**1.85**
AEP (brown)					
AR1	**1.58**	10.61	2.32	**3.32**	**3.27**
VAR1	2.09	19.27	2.40	3.65	4.01
VAR1 UMC	2.00	22.01	2.36	3.63	4.58
BVAR1 UMC	2.75	**9.04**	**2.29**	4.12	4.49

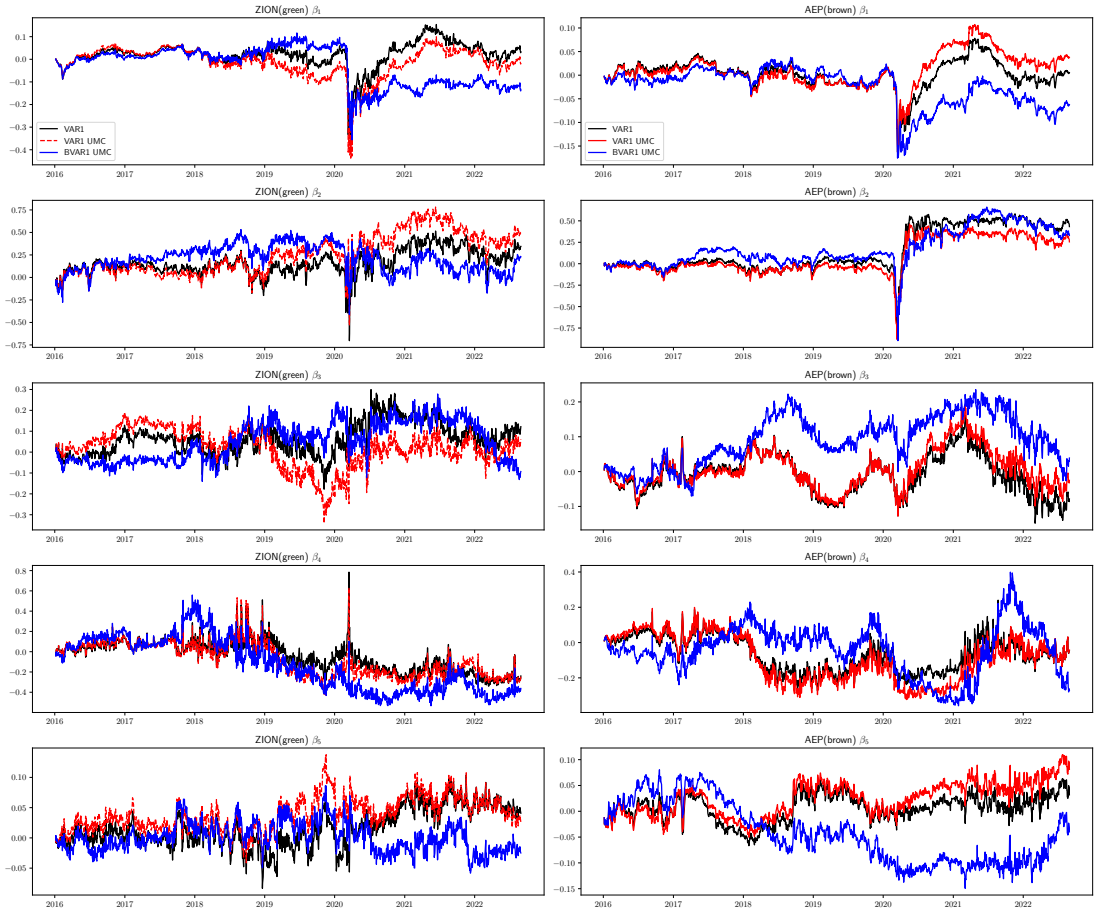

Figure 6: Accumulated difference of $\widehat{\beta} - \beta$ for green firm ZION and brown firm AEP
Plot concerns the $\sum_t \widehat{\beta}_t - \beta_t$ where the $\widehat{\beta}$ are computed by VAR1 (black curve), VAR1 UMC (dash red curve) and the BVAR UMC (blue curve).

6 Discussion

Our empirical results demonstrate several key findings regarding option pricing dynamics. The selected implied volatility (IV) surface model shows satisfactory performance for American options, achieving calibration accuracy comparable to the S&P 500 European options model in François et al. (2022) in terms of RMSE, though with more pronounced discrepancies observed for DOTM put options. In examining factor loading $\boldsymbol{\beta}$ predictions, we find that neither options written on green nor brown stocks benefit from the enhanced VAR-UMC specification, as it fails to outperform the standard VAR model across all tested scenarios. Notably, the simplest AR model consistently achieves the lowest RMSE and AMPE, suggesting that the more complex VAR structure provides no meaningful improvement in $\boldsymbol{\beta}$ projection accuracy across any dimension of our analysis. Further more, the IV surfaces predicted from $\boldsymbol{\beta}$ estimates reveals consistent patterns with the previous findings. The VAR1-UMC specification demonstrates weaker performance than the standard VAR1 model for green options while showing modest improvement for brown options - the UMC-enhanced models deliver better predictions on approximately 45% of trading days for brown firms. However, the overall results indicate that incorporating UMC factors fails to systematically enhance model performance relative to the baseline VAR specification. This is evidenced by generally higher RMSE values across most option categories, with only limited exceptions. Overall, the UMC model does not outperform the pure VAR model. At the same time, it is worth noticing that the DM statistics are much more significant for options on green than brown stocks. The Bayesian VAR extension yields similarly mixed results: while the BVAR-UMC generally underperforms VAR-UMC and standard VAR specifications, this outcome helps address concerns about dimensionality-related estimation instability. We also observed that the β_1 and β_5 of the brown option, after some point in time, start to benefit from the prior structure and then outperform the VAR UMC and VAR model. However, this is not observed from green options.

Some possible reasons could lead to such insignificant results. In the first place, our prediction analysis is carried out from January 6, 2016, to August 30, 2022, which spans the Covid-19 pandemic. Many researches shows significant transformation and the entangled effect caused by pandemic and climate risk (Karydas and Xepapadeas, 2022). Financial institutions started integrating climate risk into operations in the pre-pandemic period. Evidence shows investors already decreased brown stocks in their portfolio, hedging against climate risk while the COVID crisis abruptly broke in and shifted attention. Le Billon et

al. (2021) show that the economy is recovering from the pandemic in a way with intensive carbon emission from the perspective of the energy industry. Our calibration of β shows this unbiased effect as well. Not only for ZION and AEP, all 20 firms, regardless of green or brown, showed the same hike in β_1, β_2 at the beginning of 2020 as shown in Figure 3 when the pandemic just burst. Notice that β_1 measures the long-term ATM IV; the plot shows a clear level shift until late 2022, two years after the pandemic. Also, the MCCC shows very intensive zero values hit in the same period, as can be seen from Figure 1. Thus, it is very likely that the correlation between climate risk and volatility changed due to the shift in hedging demand from against climate risk to the potential price drop due to the pandemic. The VAR1 and VAR1 UMC prediction of β starts to diverge after the pandemic, especially for brown options, as shown in Figure 6. Therefore, before launching the research, finding a method to disentangle the climate risk from the pandemic effect is necessary.

Next, the MCCC index underlying the construction of the UMC may suffer from measurement inaccuracies. Ardia et al. (2023) provides MCCC indexes disaggregated by cluster and topic across multiple levels—for instance, separate indexes are developed for environmental effects, business impact, and societal debate. Serafeim and Yoon (2022) demonstrates that stock market prices respond primarily to financially material ESG (environmental, social, and governance) news. In the context of options, Ford et al. (2022) finds that traders are particularly attentive to environmental issues over other ESG dimensions. Therefore, constructing the UMC using the business impact and environmental components of the MCCC index may be more appropriate than relying on an aggregate one. Additionally, constructing the MCCC index omits control variables, raising concerns about potential endogeneity when the index is used in a VAR framework. For example, the UMC could be confounded with broader market volatility measures such as the VIX, which are often correlated with the options IV. Consequently, it is questionable whether a weak correlation with market-wide measures truly explains the observed insignificance.

7 Conclusion

In recent years, there has been a growing body of research on climate risk and its implications for equity returns. Ardia et al. (2023) empirically validate the hypothesis proposed by Pástor et al. (2021) that green stocks tend to outperform brown stocks when climate risk concerns rise unexpectedly. Building

on this literature, the present study conducts an ex-ante analysis using VAR and VARX models to examine whether such effects extend to equity IV. Specifically, we assess whether unexpected climate change concerns—proxied by the UMC index—enhance the prediction accuracy of the IV for brown options as compared to green options.

Our findings indicate no significant linear relation of unexpected climate change concerns and the IV of options on either green or brown stocks. Moreover, the observed effects from our model do not align with the GHG-based classification of green and brown assets. This conclusion holds not only for the prediction of IV surface factor loadings but also for the projected IV values derived from the surface. To ensure robustness, we incorporate a normal diffuse prior into the VAR and VARX models expecting to address concerns about increased dimensionality due to the inclusion of multiple lags and variables. Interestingly, a simple AR model outperforms the VAR in predicting both factor loadings and IV, suggesting that the VAR structure does not enhance forecasting power. We acknowledge a limitation of the analysis—namely, that the optimal number of lags was not selected based on formal information criteria, owing to dimensional constraints.

This study contribute to future research on exploring the link between climate change concerns and option volatility in the following ways. First, the assumption of linearity warrants reconsideration and alternative model specifications or transformations might be necessary. Second, the analysis period should ideally exclude the COVID-19 pandemic or, at minimum, employ methodologies capable of disentangling the distinct effects of the pandemic and climate change on volatility. Lastly, researchers must give careful attention to the selection of control variables to mitigate potential endogeneity concerns.

Bibliography

Ardia, D., Bluteau, K., Boudt, K., & Inghelbrecht, K. (2020). MCCC Index. Retrieved February 10, 2024, from https://sentometrics-research.com/download/mccc/

Ardia, D., Bluteau, K., Boudt, K., & Inghelbrecht, K. (2023). Climate change concerns and the performance of green vs. brown stocks. Management Science, 69(12), 7607–7632.

Azzone, M., & Baviera, R. (2022). Additive normal tempered stable processes for equity derivatives and power-law scaling. Quantitative Finance, 22(3), 501–518.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. Journal of Finance, 52(5), 2003–2049.

Bakshi, G. S., & Chen, Z. (1997). An alternative valuation model for contingent claims. Journal of Financial Economics, 44(1), 123–165.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. Review of Financial Studies, 9(1), 69–107.

Bertolotti, A., Basu, D., Akallal, K., & Deese, B. (2019). Climate risk in the US electric utility sector: A case study. Available at SSRN 3347746.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81(3), 637–654.

Bonato, M., Cepni, O., Gupta, R., & Pierdzioch, C. (2023). Climate risks and state-level stock market realized volatility. Journal of Financial Markets, 66, 100854.

Bouri, E., Iqbal, N., & Klein, T. (2022). Climate policy uncertainty and the price dynamics of green and brown energy stocks. Finance Research Letters, 47, 102740.

Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association, 65(332), 1509–1526.

Box, G. E., & Tiao, G. C. (2011). *Bayesian inference in statistical analysis*. John Wiley & Sons.

Bressan, G. M., & Romagnoli, S. (2021). Climate risks and weather derivatives: A copula-based pricing model. Journal of Financial Stability, 54, 100877.

Carriere, J. F. (1996). Valuation of the early-exercise price for options using simulations and nonparametric regression. Insurance: Mathematics and Economics, 19(1), 19–30.

Chalamandaris, G., & Tsekrekos, A. E. (2011). How important is the term structure in implied volatility surface modeling? Evidence from foreign exchange options. Journal of International Money and Finance, 30(4), 623–640.

Chen, R.-R., & Palmon, O. (2005). A non-parametric option pricing model: Theory and empirical evidence. Review of Quantitative Finance and Accounting, 24, 115–134.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. Journal of Financial Economics, 7(3), 229–263.

Dawson, R. J. (2015). Handling interdependencies in climate change risk assessment. Climate, 3(4), 1079–1096.

Duan, J.-C. (1999). Conditionally fat-tailed distributions and the volatility smile in options. Rotman School of Management, University of Toronto, Working Paper.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. Journal of Finance, 53(6), 2059–2106.

Eberlein, E., Keller, U., & Prause, K. (1998). New insights into smile, mispricing, and value at risk: The hyperbolic model. Journal of Business, 71(3), 371–405.

Ekström, E. (2004). Properties of american option prices. Stochastic Processes and their Applications, 114(2), 265–278.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging climate change news. Review of Financial Studies, 33(3), 1184–1216.

Fengler, M. R., Härdle, W. K., & Mammen, E. (2007). A semiparametric factor model for implied volatility surface dynamics. Journal of Financial Econometrics, 5(2), 189–218.

Ford, J. M., Gehricke, S. A., & Zhang, J. E. (2022). Option traders are concerned about climate risks: ESG ratings and short-term sentiment. Journal of Behavioral and Experimental Finance, 35, 100687.

François, P., Galarneau-Vincent, R., Gauthier, G., & Godin, F. (2022). Venturing into uncharted territory: An extensible implied volatility surface model. Journal of Futures Markets, 42(10), 1912–1940.

Gavriilidis, K. (2021). Measuring climate policy uncertainty. Available at SSRN 3847388.

Goncalves, S., & Guidolin, M. (2006). Predictable dynamics in the S&P 500 index options implied volatility surface. The Journal of Business, 79(3), 1591–1635.

Guo, K., Li, Y., Zhang, Y., Ji, Q., & Zhao, W. (2023). How are climate risk shocks connected to agricultural markets? Journal of Commodity Markets, 32, 100367.

Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies, 6(2), 327–343.

Ilhan, E., Sautner, Z., & Vilkov, G. (2021). Carbon tail risk. Review of Financial Studies, 34(3), 1540–1571.

Isah, K., Odebode, A., & Ogunjemilua, O. (2023). Does climate risk amplify oil market volatility? Energy Research Letters, 4(2).

Kadiyala, K. R., & Karlsson, S. (1997). Numerical methods for estimation and inference in Bayesian VAR-models. Journal of Applied Econometrics, 12(2), 99–132.

Karydas, C., & Xepapadeas, A. (2022). Climate change financial risks: Implications for asset pricing and interest rates. Journal of Financial Stability, 63, 101061.

Kruttli, M. S., Roth Tran, B., & Watugala, S. W. (2023). Pricing poseidon: Extreme weather uncertainty and firm return dynamics. Available at SSRN 3284517.

Le Billon, P., Lujala, P., Singh, D., Culbert, V., & Kristoffersen, B. (2021). Fossil fuels, climate change, and the covid-19 crisis: Pathways for a just and green post-pandemic recovery. Climate Policy, 21(10), 1347–1356.

Li, H., Bouri, E., Gupta, R., & Fang, L. (2023). Return volatility, correlation, and hedging of green and brown stocks: Is there a role for climate risk factors? Journal of Cleaner Production, 414, 137594.

Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions—Five years of experience. Journal of Business & Economic Statistics, 4(1), 25–38.

Long, S., Lucey, B., Kumar, S., Zhang, D., & Zhang, Z. (2022). Climate finance: What we know and what we should know? Journal of Climate Finance, 1, 100005.

Lv, W., & Li, B. (2023). Climate policy uncertainty and stock market volatility: Evidence from different sectors. Finance Research Letters, 51, 103506.

Malliaris, M., & Salchenberger, L. (1996). Using neural networks to forecast the S&P 100 implied volatility. Neurocomputing, 10(2), 183–195.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics, 3(1-2), 125–144.

Moon, H. R., & Perron, B. (2016). *Seemingly unrelated regressions*. Palgrave Macmillan UK.

OptionMetrics. (2024). IvyDB File and Data Reference Manual (6.0). OptionMetrics.

Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable investing in equilibrium. Journal of Financial Economics, 142(2), 550–571.

Rubinstein, M. (1985). Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from august 23, 1976 through august 31, 1978. Journal of Finance, 40(2), 455–480.

Schlenker, W., & Taylor, C. A. (2021). Market expectations of a warming climate. Journal of Financial Economics, 142(2), 627–640.

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In S. van der Walt & J. V. Millman (Eds.), *Proceedings of the 9th python in science conference* (pp. 57–61).

Serafeim, G., & Yoon, A. (2022). Which corporate esg news does the market react to? Financial Analysts Journal, 78(1), 59–78.

Tankov, P., & Tantet, A. (2019). Climate data for physical risk assessment in finance. Available at SSRN 3480156.

Venturini, A. (2022). Climate change, risk factors and stock returns: A review of the literature. International Review of Financial Analysis, 79, 101934.

Verschuur, J., Fernández-Pérez, A., Mühlhofer, E., Nirandjan, S., Borgomeo, E., Becher, O., Voskaki, A., Oughton, E. J., Stankovski, A., Greco, S. F., et al. (2024). Quantifying climate risks to infrastructure systems: A comparative review of developments across infrastructure sectors. PLoS Climate, 3(4), e0000331.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American statistical Association, 57(298), 348–368.

Zellner, W. (1996). *An introduction to bayesian inference in econometrics*. John Wiley & Sons.

Appendices

A A Bayesian VAR trial

The curse of dimensionality in the VAR model has been extensively studied due to the increase in lag and variables. For example, in our VAR(1) model, including five endogenous variables and one exogenous available lead to more than 30 parameters to be estimated. In order to see if the insignificant effect of the UMC index is due to the instability of large matrix estimation, we add a prior structure to the model for regularization purposes. Due to the limitation of computational power, we only test on two options.

A.0.1 Theory

The notations and setup mostly follow the settings proposed by Kadiyala and Karlsson (1997). To allow for a technical discussion of BVAR, we rewrite the VAR-X model (13) to be

$$\mathbf{y}_t = \mu + \mathbf{x}_t\mathbf{C} + \sum_{i=1}^{p} \mathbf{y_{t-i}}\mathbf{A}_i + \mathbf{u}_t \tag{24}$$

where $\mathbf{y}, \mathbf{x}, \mu$ are now row vectors. Equation (24) can be further converted to

$$\mathbf{y}_t = \mathbf{z}_t\mathbf{\Gamma} + \mathbf{u}_t \tag{25}$$

where $\mathbf{z} = \left\{1, \mathbf{x}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}\right\}$ is a row vector and $\mathbf{\Gamma} = \left(\mathbf{C}, \mathbf{A}_1, \cdots, \mathbf{A}_p\right)^\top$[2]. Then we stack the row vectors of $\mathbf{y}_t$, $\mathbf{z}_t$ and $\mathbf{u}_t$ for $t = 1, 2, \cdots, T$ as

$$\mathbf{Y} = \mathbf{Z}\mathbf{\Gamma} + \mathbf{U}. \tag{26}$$

Then for each column i of $\mathbf{Y}$, $\mathbf{\Gamma}$ and $\mathbf{U}$ we can write $\mathbf{y}_i = \mathbf{Z}\boldsymbol{\gamma}_i + \mathbf{u}_i$. Stacking all columns $\mathbf{y}_i$ results in

$$\mathbf{y} = (\mathbf{I}_m \otimes \mathbf{Z})\boldsymbol{\gamma} + \mathbf{u} \tag{27}$$

where $\mathbf{y} = \text{vect}(\mathbf{Y})$, $\boldsymbol{\gamma} = \text{vect}(\mathbf{\Gamma})$ and $\mathbf{u} = \text{vect}(\mathbf{u})$. We assume as usual a multivariate normal distribution for $\mathbf{u}$ which is

$$\mathbf{u} \sim N(\mathbf{0}, \Psi \otimes \mathbf{I}_T) \tag{28}$$

[2]The matrix $\mathbf{C}$ hereafter is cooperated with a row of $\boldsymbol{\mu}$ on the top.

which implies $\{\mathbf{u}_t\}$ are i.i.d $N(\mathbf{0}, \Psi)$ (i.e. independent observations overtime with a fixed variance matrix). Recall that a general Bayesian posterior distribution can be represented as

$$\underbrace{\pi(\Psi, \boldsymbol{\gamma} \mid y)}_{\text{posterior}} \propto \underbrace{f(y \mid \Psi, \boldsymbol{\gamma})}_{\text{likelihood}} \times \underbrace{\pi(\Psi, \boldsymbol{\gamma})}_{\text{prior}}$$

For likelihood, assumption (28) results in $\mathbf{y} \mid \boldsymbol{\gamma}, \Psi \sim \mathrm{N}\left((\mathbf{I}_m \otimes \mathbf{Z})\boldsymbol{\gamma}, \Psi \otimes \mathbf{I}_T\right)$. The multivariate normal density is common to be decomposed in to two part

$$f(\mathbf{y} \mid \gamma, \Psi) \propto \underbrace{|\Psi|^{-k/2} \exp\left\{-\frac{1}{2}(\gamma - \widehat{\gamma})' \left(\Psi^{-1} \otimes \mathbf{Z}'\mathbf{Z}\right)(\gamma - \widehat{\gamma})\right\}}_{\mathrm{N}\left(\widehat{\gamma}, \Psi \otimes \left[\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right]^{-1}\right)} \times \underbrace{|\Psi|^{-(T-k)/2} \times \exp\left\{-\frac{1}{2}\operatorname{tr}\left[\Psi^{-1}(\mathbf{Y} - \mathbf{Z}\widehat{\Gamma})^T(\mathbf{Y} - \mathbf{Z}\widehat{\Gamma})\right]\right\}}_{\mathrm{IW}\left((\mathbf{Y}-\mathbf{Z}\widehat{\Gamma})^{\mathrm{T}}(\mathbf{Y}-\mathbf{Z}\widehat{\Gamma}), \mathrm{T}-(mp+q)-1-m\right)} \tag{29}$$

where $\widehat{\boldsymbol{\gamma}} = \operatorname{vect}(\widehat{\boldsymbol{\Gamma}}), \widehat{\boldsymbol{\Gamma}} = (\mathbf{Z}^\top\mathbf{Z})^{-1}\mathbf{Z}^\top\mathbf{Y}$ is the OLS estimator of $\boldsymbol{\gamma}$. That means the multivariate density is proportional to the product of a normal density and inverse Wishart density. Researchers have also developed many different priors to choose. The common one is the Normal-Wishart natural conjugate prior (for normal likelihood) first proposed explicitly by Box and Tiao (2011), in which an inverse-Wishart prior is assumed for the covariance matrix. However, its limitation is widely discussed; for example, it is too rigid to control all variables' variance using only one shrinkage parameter. Here, we choose the normal-diffuse prior proposed by W. Zellner (1996) as follows. The normal-diffuse prior assumes an independent innovation variance (i.e., $\pi(\Psi, \boldsymbol{\gamma}) = p(\Psi)p(\boldsymbol{\gamma})$) which are

$$\boldsymbol{\gamma} \sim \mathrm{N}(\tilde{\gamma}, \tilde{\Sigma}), p(\Psi) \propto |\Psi|^{-(m+1)/2} \tag{30}$$

where $\sim$ indicates a prior setting of mean and variance. The advantage is that we usually do not have much information about the covariance matrix of innovation. Imposing an inverse-Whishart type of distribution seems too strong and arbitrary. Thus, we alleviate this issue by adopting an improper diffuse prior and reasonably assuming its independence with the means. However, the trade-off of imposing a diffuse prior is the absence of analytical posterior after joining together with the likelihood in (29). As stated by Kadiyala and Karlsson (1997), the (marginal) posterior for $\boldsymbol{\gamma}$ is a kernel constituted by the product of marginal normal prior and matricvariate t-distribution. Fortunately, the author provides us with two marginal distributions, which are

$$\boldsymbol{\gamma} \mid \Psi, \mathbf{y} \sim \mathrm{N}\left(\bar{\gamma}, \left(\tilde{\Sigma}^{-1} + \Psi^{-1} \otimes \mathbf{Z}'\mathbf{Z}\right)^{-1}\right) \tag{31}$$

$$\Psi^{-1} \mid \gamma, \mathbf{y} \sim W\left(\left[(\mathbf{Y} - \mathbf{Z}\widehat{\boldsymbol{\Gamma}})'(\mathbf{Y} - \mathbf{Z}\widehat{\boldsymbol{\Gamma}}) + (\boldsymbol{\Gamma} - \widehat{\boldsymbol{\Gamma}})'\mathbf{Z}'\mathbf{Z}(\boldsymbol{\Gamma} - \widehat{\boldsymbol{\Gamma}})\right]^{-1}, T\right) \tag{32}$$

where $\bar{\gamma} = \left(\Sigma^{-1} + \Psi^{-1} \otimes \mathbf{Z}'\mathbf{Z}\right)^{-1} \left[\Sigma^{-1}\tilde{\gamma} + \left(\Psi^{-1} \otimes \mathbf{Z}'\mathbf{Z}\right)\hat{\gamma}\right]$. Equation (31) and (32) thus allow us to utilize the Markov Chain Monte Carlo (MCMC) method to generate a sample of $\boldsymbol{\gamma}$ from the posterior distribution. We describe our Gibbs sampling algorithm in algorithm 1.

Our ultimate goal is to come up with an estimate for $\boldsymbol{\gamma}$ from the posterior distribution. In time series, when it comes to forecasting, it usually refers to the expectation conditioning on the latest information. Following Kadiyala and Karlsson (1997), the collection of (24) can be written as a first-order system $\mathbf{y}_t^* = \mathbf{y}_{t-1}^*\mathbf{A} + \mathbf{x}_{t-1}\mathbf{D} + \mathbf{u}_t^*$, where $\mathbf{y}_t^* = \{\mathbf{y}_t, \ldots, \mathbf{y}_{t-p+1}\}$, $\mathbf{D} = \{\mathbf{C}, \mathbf{0}, \ldots, \mathbf{0}\}$, $\mathbf{u}_t^* = \{\mathbf{u}_t, 0, \ldots, 0\}$ and

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{I} & 0 & \cdots & 0 \\ \mathbf{A}_2 & 0 & \mathbf{I} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & 0 \\ \mathbf{A}_{p-1} & 0 & 0 & \cdots & \mathbf{I} \\ \mathbf{A}_p & 0 & \cdots & \cdots & 0 \end{pmatrix} \tag{33}$$

Since we perform one-step-ahead forecast, we are trying to compute the following quantity

$$\widehat{\mathbf{y}}_{t+1} = \mathbb{E}\left(\mathbf{y}_{t+1} \,|\, \mathcal{I}_t\right) = \int \left(\mathbf{y}_t\mathbf{A} + \mathbf{x}_t\mathbf{D}\right) p\left(\boldsymbol{\gamma} \,|\, \mathcal{I}_t\right) \mathrm{d}\boldsymbol{\gamma} \tag{34}$$

then the sample mean is used to approximate the integral which is

$$\widehat{\mathbf{y}}_{t+1} \approx \frac{1}{N} \sum_{i=1}^{N} \left(\mathbf{y}_t\mathbf{A}(\boldsymbol{\gamma}_i) + \mathbf{x}_t\mathbf{D}(\boldsymbol{\gamma}_i)\right) \tag{35}$$

where $\mathbf{A}$ and $\mathbf{D}$ are reshaped matrix from each sample $\boldsymbol{\gamma}_i$. To complete, the prior mean $\tilde{\gamma}$ variance $\tilde{\Sigma}$ still needs to be provided in the first place. We then take the famous form of setting proposed by Litterman (1986). First, regarding the prior mean, Litterman (1986) observes the unit-roots appear in many economic variables (i.e., $y_t = y_{t-1} + \epsilon_t$), so he recommends setting the prior for parameters associating with its first lag to be unit one, and all rest to be zero. However, in our case, the variable does not have a clear economic meaning, and based on our empirical results in Table 5, the unit root is not seen in any of the $\Delta\beta$. Thus, we also impose a zero prior mean for those parameters. For prior variance, we follow Litterman (1986)'s setting. Recall the columns of (26) consisting of equation $\mathbf{y}_i = \mathbf{Z}\boldsymbol{\gamma}_i + \mathbf{u}_i$. We then set the covariance

matrix for $\boldsymbol{\gamma}_i$ as

$$\left(\sum\right)_{ij}^{\gamma_i} = \begin{cases} \pi_3 \sigma_i^2, & \text{if } i = 1, 2, \ldots, q+1 \\ \pi_1 / k, & \text{if } i = j \\ (\pi_2 \sigma_i^2)/(k\sigma_j^2). & \text{if } i \neq j \\ 0, & 0 \text{ / w} \end{cases} \tag{36}$$

where σ_i and σ_j are the residual standard error of a AR(p) process for variable i, k is the kth lag and π are the hyper parameters to be calibrated in the first place. π_1 is related to the prior variance of variables own lags, π_2 controls the prior variance of other variables and π_3 concerns the variance for exogenous (e.g. our UMC index) and deterministic variables (e.g. the $\boldsymbol{\mu}$). More precisely, we assume zero prior covariance between variables, and the variance decreases with lag length. Also, σ is taken to scale the variance of variables to different levels.

A.0.2 Application

We then apply the BVAR model on the greenest firm, ZION, and brownest firm, AEP. We analyze only the VAR(1) and VAR(1)-UMC models. First, we calibrate the hyper-parameters of π in (36) using a rolling window of 252 days and then predict the following year. By doing some sanity checks, the value suggested by Kadiyala and Karlsson (1997) and Litterman (1986) does not fit the data. For the sake of computation speed, we calibrate over a grid of (π_1, π_2) and set $\pi_3 = \pi_1$ to let the influence of exogenous variables be as significant as the variables' own lag. Considering the scale of different $\Delta\beta$, we determine the π by minimizing the mean square percentage error across all loading, which is

$$(\pi_1, \pi_2)^* = \text{argmin}_{(\pi_1, \pi_2)} \frac{1}{N} \sum_{j=1}^{5} \sum_{i=1}^{N} \left(\frac{\widehat{\Delta\beta}_{j,i} - \Delta\beta_{j,i}}{\Delta\beta_{j,i}} \right)^2 \tag{37}$$

We calibrate the VAR(2) and VAR(2)-UMC models separately on different ranges since a considerable variance imposed on the exogenous variable might lead to a more minor variance for the endogenous variable. Calibrating only over the VAR(2)-UMC model and assigning the same prior to VAR(2) is unfair and results in a super tight prior with inadequate flexibility. The best calibration results are shown in Table 11. Next, we run again the rolling window of 252 days and do a prediction one day ahead. The Gibbs sampling is detailed in Algorithm 1. On the first day, we draw 10,000 samples and discard the first 5000 draws to ensure the stationary state is reached. The starting points of $\boldsymbol{\Psi}$ are set to be the OLS estimate.

Algorithm 1 Gibbs sampling of equation (31) and (32)

for $t \leftarrow (p+1)$ **to** T **do**
  if $t = p + 1$ **then**
    $N \leftarrow 10000$ ▷ Number of samples
    $b \leftarrow 5000$ ▷ Burn-in, number of initial samples to discard
    $\mathbf{\Psi} \leftarrow \mathbf{\Psi}_{OLS}$ ▷ Staring point for MCMC, OLS estimation of $\mathbf{\Psi}$
  else
    $N \leftarrow 1000$
    $b \leftarrow 500$
    $\mathbf{\Psi} \leftarrow \overline{\mathbf{\Psi}}_{t-1}$ ▷ Staring point for MCMC
  $\{\boldsymbol{\gamma}\}, \{\mathbf{\Phi}\} \leftarrow \{\}, \{\}$ ▷ Vaults for simulations
  for $j \leftarrow 1$ **to** N **do**
    generate $\boldsymbol{\gamma}_j$ from $\boldsymbol{\gamma} \mid \mathbf{\Psi}$ by equation (31) ▷ Multivariate normal draw
    generate $\mathbf{\Psi}_j$ from $\mathbf{\Psi}^{-1} \mid \boldsymbol{\gamma}_j$ by equation (32) ▷ Inverse-Wishart draw then inverse
    $\mathbf{\Psi} \leftarrow \mathbf{\Psi}_j$
    append $\boldsymbol{\gamma}_j$ to $\{\boldsymbol{\gamma}\}$, $\mathbf{\Psi}_j$ to $\{\mathbf{\Phi}\}$
  $\overline{\mathbf{\Psi}}_t \leftarrow \sum_{i=1}^{N} \mathbf{\Psi}_i / N$ ▷ Estimate the starting point for simulation on day $t+1$
end

We then start from equation (31) and transit to (32) several times, and the simulated results are saved. For the rest of the days, since the samples used to construct the distribution only differ by one single $\boldsymbol{\beta}$ observation, the stationary state is expected to be close to the previous day. Thus, we start the chain from the previous day's expectation, $\overline{\boldsymbol{\Psi}}_t$, and use far fewer samples of 1000 to discard the first 500 as burn-in. Even though no formal test is carried out for convergence check, we did not detect severe convergence issues by looking at accumulated mean on some days. Our sanity check indicates the same case as found by Kadiyala and Karlsson (1997) that the convergence is not sensitive to the starting points of $\boldsymbol{\Psi}$.

Table 11: Out of sample prediction performance across difference models

NOTE: Table shows the hyper parameters used for BVAR in (33) by minimization in (37).

Tickers	Parameters	VAR(1)	VAR(1)-UMC
AEP	π_1	16.7	23.3
	π_2	27.8	4.0
ZION	π_1	21.1	16.7
	π_2	2.0	2.0

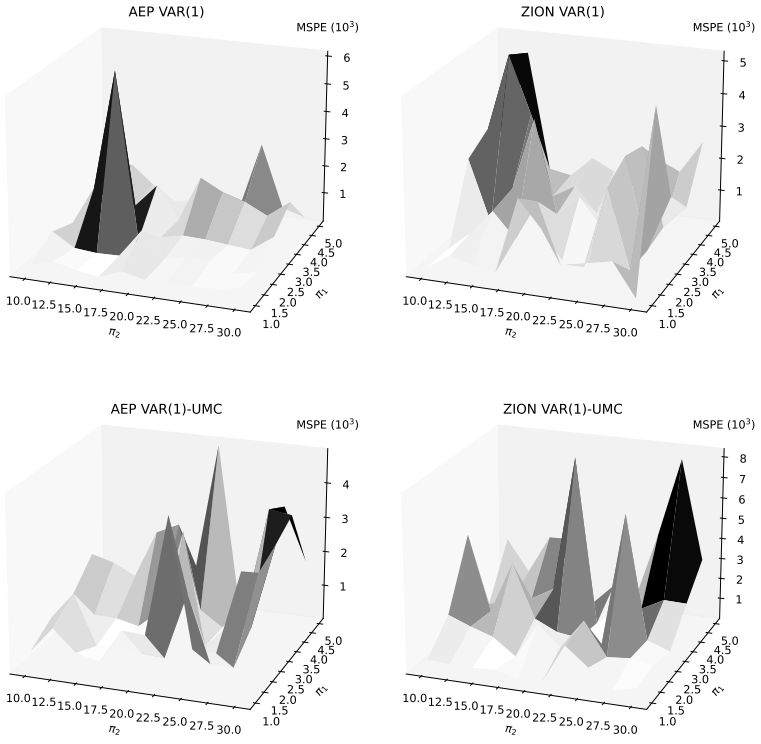

Figure 7: Hyper-parameter calibration results
Calibration results for VAR(1) and VAR(1)-UMC model targeting a minimized sum of square percentage error in (37).