THREE ESSAYS IN EMPIRICAL ASSET PRICING AND RETURN PREDICTABILITY

by

Yueliang Lu

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Business Administration

Charlotte

2023

Approved by:

____

Dr. Yufeng Han

____

Dr. Weidong Tian

____

Dr. I-Hsuan Ethan Chiang

____

Dr. Steven P. Clark

____

Dr. Christopher M. Kirby

____

Dr. Jiancheng Jiang

ABSTRACT

YUELIANG LU. Three Essays in Empirical Asset Pricing and Return Predictability.
(Under the direction of DR. YUFENG HAN & DR. WEIDONG TIAN)

This dissertation contains three essays on empirical asset pricing. The first essay presents the first evidence on how macro trends affect equity risk premium, going beyond the literature that rely on only the most recent values. We show that macro trends contribute statistically and economically to the out-of-sample aggregate market return predictability. Moreover, we present novel evidence that nonlinearity matters in market return predictability by combining macro trends with neural networks, yielding an out-of-sample R^2_{OS} statistic as high as 1.6%.

The second essay develops a theory of forward returns for an equity index. We obtain the forward returns using information from derivatives markets, including index option prices and gammas, VIX-futures, and prices of VIX-options. We document a pro-cyclical term structure of S&P 500 forward returns and a robust short-term reversal pattern. Moreover, by designing and implementing a market-timing strategy, we demonstrate that forward equity returns provide real-time trading signals with substantial economic value.

The third essay studies the causal effect of short-sale constraints on anomalies by examining an extensive set of 182 anomalies documented in the accounting, finance and economics literature. Our identification strategy relies on a persistent, robust and plausibly exogenous shock to short-selling supply induced by the dividend tax law change in the Job and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003. We find that anomalies become stronger following the dividend record months, driven by stronger overpricing as opposed to underpricing in the post-JGTRRA periods. While the shock magnifies returns to most anomaly types, the valuation anomalies seem unlikely to be driven by mispricing.

ACKNOWLEDGEMENTS

I am especially thankful to my committee chairs and advisors, Yufeng Han and Weidong Tian. Thank you so much for your continuous support and guidance throughout my entire PhD journey. I am also grateful to have Guofu Zhou (WashU Olin) as my informal advisor and coauthor. You three set a great example for me of what an inspirational scholar and researcher should be. I am looking forward to more collaborations in the future.

I would like to thank my dissertation committee members, I-Hsuan Ethan Chiang, Steven P. Clark (also my coauthor), Chris Kirby, and Jiangcheng Jiang, for your insightful comments and suggestions that greatly improve my work. To my coauthor Weike Xu, thank you for your support and wonderful teamwork. I am truly happy to have you as my friend.

I am also grateful to Joseph J. Cabral, Jun (Jeffery) Chen, Aparna Gupta, Dolly King, Gene Lai, Hong Liu, Alberto Martin-Utrera, David Mauer, Abena Owusu, Lisa Ponti, Majeed Simaan, Patrick Smith, Kelly Voster, and Yilei Zhang. To my Ph.D. cohorts: Hunter Fraizer, Helen Gardner, Shuangshuang Ji, Gunratan Lonare, Ziye Nie, Taichun Piao, Ke Shang, Chen Shen, and Zifeng Zeng. Last but not the least, to Deon Ritchie and Patrick Baker. Thank you Deon for giving me that cute coffee table; thank you Patrick for booking seminar rooms almost every week and preparing the reimbursement documents for me.

Finally, to my fur baby (a cute pug), Pepper: my everyday little sunshine.

DEDICATION

To my mother, Wen Liu.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: MACROECONOMIC TRENDS AND EQUITY RISK PREMIUM

1.1 Introduction

As emphasized by Fed Chair Powell, it is important to have more than just one month's worth of data to make informed decisions about monetary policy, and one must be cautious against overreaction to short-term data that may not provide a clear picture of the economic outlook. The implication is that in predicting the future economy, the past month data is unlikely to be sufficient, and the entire macro trends are likely to matter. To the extent that expected asset returns are function of the future states of the economy (see, e.g., Merton, 1973), the expected market risk premium must be a function of not only the most recent macroeconomic variables, but also their past trends. In the voluminous macroeconomic literature, researchers already use a time series approach to forecast GDP growth or inflation with their long-term lagged values (for example, Stock and Watson, 1993, 1999, 2003, 2004, 2006; Ang et al., 2006; Aruoba and Diebold, 2010; Bauer and Rudebusch, 2020; Yang, 2020). However, in the vast literature on market return predictability, there is no study yet that accounts for the role of macro trends or uses the data beyond the most recent ones.[1]

In this paper, we present the first evidence on how macro trends affect market risk premium, going beyond the literature that rely on only the most recent values. Specifically, we focus on how investors learn from economic fundamentals based on moving averages of observable macroeconomic variables to forecast the market excess return. Trend has received increasing attention and has been proved fruitful for macroeconomy and individual stocks.

[1]Nelson (1976); Fama and Schwert (1977); Rozeff (1984); Keim and Stambaugh (1986); Campbell (1987); Fama and French (1988a); Kothari and Shanken (1997); Pontiff and Schall (1998), among many others, find that a number of macro economic variables contain information on the market risk premium. Goyal and Welch (2008) examine a common set of 15 variables, and Rapach and Zhou (2022) provide a latest survey.

For example, Cieslak and Povala (2015); Han et al. (2016); Huang et al. (2020); De la O and Myers (2022), among others, extrapolate the trend in inflation, individual stock prices, firm fundamentals, and earnings growth, based on averages of past observations. However, a study on how macro trends affect the market return is missing in the literature. We fill the gap and show that macro trends are important and they contribute statistically and economically to the out-of-sample aggregate market return predictability.

We make several contributions. First, we show, for the first time, that macro trends matter to the aggregate stock market return predictability both statistically and economically, highlighting its significant unrecognized role in time-series return predictability. Compared with the 1-month forecast combination (simple pooling) that only uses the most recent data on a set of 14 macro variables, the trend-pooling methods, based on trends data, produce larger and more significant out-of-sample R^2_{OS} statistics. For instance, in the period of 2000:01−2020:12, 1-month simple pooling fails to outperform the historical average, resulting in an *insignificant* out-of-sample R^2_{OS} statistic of 0.41%. In contrast, the trend-pooling through 1- and 3-year moving averages outperform substantially, with R^2_{OS} statistics rising to significant levels of 0.56% and 0.69%, respectively. The empirical results are robust to using various dimension reduction techniques, including LASSO, PCR, and PLS.

Whether significant predictability of macro trends can yield sizable economic gains is another important question. We show that information of the macro trends indeed leads to sizable investment gains for a mean-variance investor from an asset allocation perspective. The annualized certainty equivalent return (CER) gains are 6.83%, 7.35%, 7.80%, and 8.76% at the annual horizons for 1-month simple pooling, 1-year trend-pooling, 1-year PLS-pooling, and 1-year LASSO-pooling, respectively, when the investor allocates investments between the market and risk-free rate. In general, investor portfolios based on 1-year macro trends consistently deliver higher average returns, lower standard deviation, and thus significantly larger out-of-sample Sharpe ratios. Our asset allocation results are robust to a proportional transaction cost of 0.50%.

Second, to the best of our knowledge, we contribute to the market risk premium literature by providing the first evidence that nonlinearity matters in the aggregate market return predictability. We apply neural networks into equity risk premium forecast, and allow for a large number of trend indicators and their complex nonlinear interactions. While neural networks have been applied into complex machine learning problems in finance and asset pricing, our paper employs it to predict the market directly.[2] We find that combining time-series trends with conventional neural networks is important. For example, a neural network with two hidden layers on average performs the best, yielding an out-of-sample R^2_{OS} statistic around 1.6%, an almost 80% increase relative to the linear model. We also find that increasing the number of hidden layers (or neurons) does not always translate into incremental gains, suggesting that "shallow" learning outperforms "deep" learning in the time-series market return prediction, which is consistent with Gu et al. (2020) finding in the cross-sectional stock return prediction.

Lastly, we contribute to the debate on market return predictability. The investigation of predictive power of macroeconomic variables on equity risk premium can go back to 1920 when Dow (1920) first explores the role of dividend ratios. Goyal and Welch (2008) call into question the out-of-sample predictability of the US market excess return, generating a substantial number of responses such as Campbell and Thompson (2008); Rapach et al. (2010); Neely et al. (2014); Rapach et al. (2016); Huang et al. (2015); Dong et al. (2022). A decade later, Goyal et al. (2021) re-examine the 29 variables from 26 papers published after Goyal and Welch (2008) and conclude that the predictive performance remains disappointing overall. The lack of consistent out-of-sample evidence in Goyal et al. (2021) indicates the need for improved forecasting methods to better establish the empirical reliability of market risk premium predictability. We show that the evidence is still in favor of

[2] Gu et al. (2020) provide a comparative analysis of machine learning methods in return prediction and find that neural networks perform best; Goyenko and Zhang (2020) extend their approach to include option characteristics. However, they both study cross-sectional return predictability using data of individual stocks, whereas we study the time-series return predictability with only market return data and the predictors.

the out-of-sample market excess return predictability, which is statistically and economically strong especially accounting for trends and nonlinearity.

Our paper uses a simple moving average method to capture the longer-term information in macroeconomics. An open question is why use a moving average. First, it has a long history. Relative to the extensive use of macro variables by academia, technical indicators based on moving averages of prices are commonly used by practitioners to track price trends and to make forecasts of price moves. We simply apply the same logic to macro variables. Second, moving average can reduce noise and smooth predictors as most macroeconomic variables are highly persistent. For instance, the T-bill rate is known to exhibit an autocorrelation as high as 0.99. The moving-average rule as a simple filter is less model dependent and thus more robust to the choice of underlying predictive variables. Third, moving average compares current levels to previous ones. In this regard, it is possible that the algorithm is able to capture under/over reaction to macroeconomic news, which is the key driver of return predictability (e.g., Hong and Stein, 2007, the gradual information flow). Lastly, moving average can be viewed as the simplest method that captures the sequential effect in time series. Because asset returns often exhibit sequential dependence, conventional machine learning models like the (feed-forward) neural networks may lose effectiveness with sequentially dependent input predictor variables (Cong et al., 2021). We show that one could also use the underlying lagged values to forecast, but using moving averages has economic interpretation as trends.[3] Using moving averages generally performs better than using lagged values directly in the long run.

1.2 Macro Trends

To learn the economic fundamentals and capture the trend in macroeconomics, we allow investors to use the past values of macroeconomic variables and summarize them using a

[3] A large body of literature use the term *trend inflation* to acknowledge the highly persistent expected inflation dynamics (e.g., Kozicki and Tinsley, 2001; Rudebusch and Wu, 2008; Bekaert et al., 2010; Cieslak and Povala, 2015; De la O and Myers, 2022). Among them, Cieslak and Povala (2015) show that a moving average of past core inflation forecasts the future inflation well, which reflects people's sluggish update of their inflation expectations.

moving average with a backward-looking window of l. Specifically, let $\{X_t^k\}_{k=1}^K$ denote K macroeconomic variables known at time t. Mathematically, for each variable, X_t^k, we define a macro trend signal

$$MA_{t,l}^k = \frac{X_t^k + X_{t-1}^k + \cdots + X_{t-l+1}^k}{l}, \quad 1 \le l \le L. \tag{1.1}$$

Obviously, when $l = 1$, the above equation reduces to the existing literature where only the most recent value, X_t^k, is used.

1.2.1 Linear trend-pooling

Given $K \times L$ moving-average-based predictors, we start with a linear *trend-pooling* method by running predictive regressions with one at a time, and use the estimated coefficients to compute the one-step ahead forecast of the market excess return,

$$\hat{r}_{t+1|t,l,k} = \hat{\alpha}_{t,l}^k + \hat{\beta}_{t,l}^k MA_{t,l}^k, \quad l \le L,\ k \le K. \tag{1.2}$$

We next take the arithmetic mean of $K \times L$ individual forecasts from Equation (1.2),

$$\hat{r}_{t+1|t} = \frac{1}{K \times L} \sum_{k=1}^{K} \sum_{l=1}^{L} \hat{r}_{t+1|t,l,k} \tag{1.3}$$

and use it as the final forecast.

It is well-known that a combination of forecasts often performs better than a single forecast in various domains since the seminal work by Bates and Granger (1969). When $L = 1$, the two-step pooling includes only the first step and reduces to the forecast combination method proposed by Rapach et al. (2010). For comparison, we name it *simple pooling*. We extend simple pooling to include macro trends into market return forecasting.

1.2.2 Nonlinear trend-pooling

We next apply neural networks into equity risk premium forecast, and allow for a large number of trend indicators and their complex nonlinear interactions. While neural networks have been applied into complex machine learning problems in finance and asset pricing, our paper employs it to predict the market return directly.

We focus on the conventional "feed-forward" neural network, which consist of an "input layer" of raw predictors, one or more "hidden layers" that interact and nonlinearly transform the predictors, and an "output layer" that aggregates hidden layers into an ultimate outcome prediction. Constructing a neural network requires the inputs of the number of hidden layers, the number of neurons of each layer, and which units are connected. However, selecting a network structure by cross-validation is in general a challenging task. Therefore, we follow Gu et al. (2020) to fix a variety of network architectures ex ante and estimate each one of them.

Specifically, we consider the neural network structures with up to five hidden layers. The simplest neural network has a single hidden layer of 2 neurons, denoted as NN1. Next, NN2 has two hidden layers with 4 and 2 neurons, respectively; NN3 has three hidden layers with 8, 4 and 2 neurons, respectively; NN4 has four hidden layers with 16, 8, 4 and 2 neurons, respectively; and NN5 has five hidden layers with 32, 16, 8, 4 and 2 neurons, respectively. All architectures are fully connected so that each unit receives an input from all units in the layer below. Overall, the consideration of multiple neural network structures helps to provide us with a good sense of the robustness of the out-of-sample forecasting results and the trade-offs of network depth in the equity risk premium forecasting.

We use the same activation function at all nodes, and choose a popular functional form in recent literature known as the rectified linear unit

$$\text{ReLU}(x) = \max(x, 0), \tag{1.4}$$

which encourages sparsity in the number of active neurons and allows for faster derivative evaluation. We use stochastic gradient descent (SGD) to train a neural network. For computational efficiency, we choose the adaptive moment estimation (Adam) by Kingma and Ba (2015).

We use multiple random seeds to initialize the neural network estimation and construct predictions by averaging the forecasts from the top five best networks in validation. Specifically, in each step of forecasting, we choose a fixed validation period and rank the forecasting performance by the out-of-sample R^2_{OS} statistics computed in the validation period. We then choose the five models with the highest R^2_{OS} and use the pooling of these five forecasts as the final forecast. In essence of forecast combination, this averaging step reduces prediction variance because the stochastic nature of the optimization can cause different seeds to produce different forecasts.

1.2.3 Dimension reduction

Apart from the linear forecast combination and the nonlinear neural networks, another straightforward method to incorporate a large set of information is a multiple predictive regression (or a "Kitchen sink" model),

$$r_{t+1} = \alpha_t + \sum_{k=1}^{K} \sum_{l=1}^{L} \beta_{t,l}^{k} MA_{t,l}^{k} + \varepsilon_t. \tag{1.5}$$

An obvious out-of-sample forecast by Equation (1.5) is given by

$$\hat{r}_{t+1|t} = \hat{\alpha}_t + \sum_{k=1}^{K} \sum_{l=1}^{L} \hat{\beta}_{t,l}^{k} MA_{t,l}^{k}, \tag{1.6}$$

where $\{\hat{\alpha}_t, \hat{\beta}_{t,l}^{k}\}$ are the OLS estimates in Equation (1.5) based on data up to time t.

However, the forecast based on Equation (1.5) is highly susceptible to in-sample overfitting, and thus leads to poor out-of-sample performance (Goyal and Welch, 2008; Rapach et al., 2010). Therefore, given $K \times L$ moving-average based predictors, we propose several more sophisticated dimension reduction techniques that can deal with a large number of potential predictors while guard against overfitting. We consider LASSO (Tibshirani, 1996), Elastic Net (Zou and Hastie, 2005), principal components regression, partial least squares (Kelly and Pruitt, 2013, 2015), and scaled PCR (Huang et al., 2021).

For shrinkage methods, like LASSO, we apply penalty to K macro variables of the same lag, and obtain a one-step ahead forecasts; we repeat the procedure for each lag, and aggregate the forecasts by a simple mean as the final forecast; conversely, for principal component methods, like PLS, we apply the algorithm to L moving averages of the same macro variable, and obtain a one-step ahead forecast; we repeat the procedure for each variable, and aggregate the forecasts by a simple mean as the final forecast.

1.2.4 Forecast evaluation—statistical accuracy

We assess equity risk premium forecasts in terms of statistical accuracy via mean square prediction error (MSPE). Denote the errors for the historical average benchmark and a competing forecast by

$$\hat{e}_{0,t} = r_t - \bar{r}^{HA}_{t|t-1}, \tag{1.7}$$

$$\hat{e}_{1,t} = r_t - \hat{r}_{t|t-1}, \tag{1.8}$$

respectively. The sample MSPE is given by

$$\widehat{\text{MSPE}}_j = \frac{1}{T}\sum_{t=1}^{T} \hat{e}^2_{j,t|t-1}, \quad j = 0,1, \tag{1.9}$$

where T is the number of out-of-sample observations.

To test for a difference in the population MSPEs, we use Clark and West (2007) procedure which can be conveniently implemented in a simple regression framework,

$$\underbrace{\hat{e}^2_{0,t|t-1} - \hat{e}^2_{1,t|t-1} + \left(\hat{r}^{HA}_{t|t-1} - \hat{r}_{t|t-1}\right)^2}_{f_{t|t-1}} = \mu + \varepsilon_t. \tag{1.10}$$

The t-statistic corresponding to the OLS estimate of μ in Equation (1.10) is used to test

$$\mathbb{H}_0 : \mu \leq 0 \quad \text{versus} \quad \mathbb{H}_1 : \mu > 0, \tag{1.11}$$

which is equivalent to

$$\mathbb{H}_0 : \text{MSPE}_0 \leq \text{MSPE}_1 \quad \text{versus} \quad \mathbb{H}_1 : \text{MSPE}_0 > \text{MSPE}_1. \tag{1.12}$$

The t-statistic is computed using a heteroskedasticity- and autocorrelation-consistent (HAC) standard error (Newey and West, 1987).

It is common to report the Campbell and Thompson (2008) R^2_{OS} statistic,

$$R^2_{OS} = 1 - \frac{\widehat{\text{MSPE}}_1}{\widehat{\text{MSPE}}_0}, \tag{1.13}$$

which gives the proportional reduction in the sample MSPE for the competing forecast with regard to the benchmark. Using Clark and West (2007) statistic to test the statistical significance of R^2_{OS} is tantamount to testing (in population)

$$\mathbb{H}_0 : R^2_{OS} \leq 0 \quad \text{versus} \quad \mathbb{H}_1 : R^2_{OS} > 0. \tag{1.14}$$

Because the predictable component in the monthly market excess return is necessarily limited, the R^2_{OS} statistic will be small. Nevertheless, Campbell and Thompson (2008) suggest that a monthly R^2_{OS} statistic as small as 0.5% can signal economic significance based on the market Sharpe ratio. In the next subsection, we assess the economic significance of market return forecasts by measuring their economic value to an investor.

1.2.5 Forecast evaluation—economic value

Consider a mean-variance investor who allocates across the market and risk-free Treasury bills each month. At the end of month t, the investor faces the objective function

$$\arg_{\hat{\omega}_{t+1}} \hat{\omega}_{t+1}\hat{r}_{t+1} - \frac{\gamma}{2}\hat{\omega}_{t+1}^2\hat{\sigma}_{t+1}^2, \tag{1.15}$$

where γ denotes the coefficient of relative risk aversion, $\{\hat{\omega}_{t+1}, 1-\hat{\omega}_{t+1}\}$ are allocation weights to the market portfolio and risk-free bills at month $t+1$, $\hat{r}_{t+1}$ is the market excess return forecast, and $\hat{\sigma}_{t+1}^2$ is the investor's forecast of the variance of the market excess return. The optimal mean-variance portfolio weight on the market can be computed as[4]

$$\hat{\omega}_{t+1}^{\star} = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2}\right). \tag{1.16}$$

We assume that the investor uses the sample variance computed over a 60-month rolling estimation window to forecast the variance in Equation (1.16).

We next compute three quantities (performance measures), based on the mean $\hat{\mu}_j$ and standard deviation $\hat{\sigma}_j$ of the out-of-sample realized returns by a forecasting method j. First, we measure the *out-of-sample Sharpe ratio* (SRatio)

$$SR\hat{a}tio_j = \frac{\hat{\mu}_j}{\hat{\sigma}_j}. \tag{1.17}$$

To test whether the Sharpe ratios of two strategies are statistically distinguishable, we follow DeMiguel et al. (2009) to compute the p-value of the difference.

Second, we compute the *certainty-equivalent return* (CEQ) of each strategy,

$$C\hat{E}Q_j = \hat{\mu}_j - \frac{\gamma}{2}\hat{\sigma}_j^2. \tag{1.18}$$

[4]We follow Campbell and Thompson (2008) to set $\gamma = 3$ and to constrain the portfolio weight on risky asset to lie within 0 and 1.5.

Relative to a benchmark, we also compute the CEQ difference, which is known as the utility gain in the forecasting literature (see, e.g., Rapach and Zhou, 2022).

Lastly, we compute the *performance fee* suggested in Fleming et al. (2001). It can be interpreted as the maximum fee that a quadratic-utility investor would be willing to pay to switch from the benchmark to the alternative. To estimate this fee, we find the value of Δ that solves

$$\sum_t \left[\left(R_{j,t} - \Delta\right) - \frac{\gamma}{2(1+\gamma)} \left(R_{j,t} - \Delta\right)^2 \right] = \sum_t \left[R_{i,t} - \frac{\gamma}{2(1+\gamma)} R_{i,t}^2 \right], \tag{1.19}$$

where $R_{j,t}$ and $R_{i,t}$ denote the out-of-sample realized returns by the competing forecast j and the benchmark forecast i, respectively. We report the estimate of Δ as annualized fees in basis points.

1.3 Data

We consider the same 14 macroeconomic variables from Goyal and Welch (2008) for which monthly data are available for 1926:12−2020:12.[5] The market excess return is the CRSP value-weighted market return minus the risk-free return (Treasury bill rate). The 14 macroeconomic variables are as follows.

1. *Dividend-price ratio, D/P*: Difference between the log of dividends paid on the S&P 500 index and the log of stock prices (S&P 500 index), where dividends are measured using a one-year moving sum.

2. *Dividend yield (log), D/Y*: Difference between the log of dividends and the log of lagged stock prices.

3. *Earnings-price ratio, E/P*: Difference between the log of earnings on the S&P 500 index and the log of stock prices, where earnings are measured using a one-year moving sum.

[5] The data can be obtained from Amit Goyal's website at `https://sites.google.com/view/agoyal145/?redirpath=/`

4. *Dividend-payout ratio (log), D/E*: Difference between the log of dividends and the log of earnings.

5. *Book-to-Market, B/M*: Ratio of book value to market value for the Dow Jones Industrial Average.

6. *Net equity expansion, $NTIS$*: Ratio of twelve-month moving sums of net issues by NYSE-listed stocks to total end-of-year market capitalization of NYSE stocks.

7. *Treasury bill rate, TBL*: Interest rate on a three-month Treasury bill (secondary market).

8. *Long-term yield, LTY*: Long-term government bond yield.

9. *Long-term return, LTR*: Return on long-term government bonds.

10. *Term spread, TMS*: Difference between the long-term yield and the Treasury bill rate.

11. *Default yield spread, DFY*: Difference between BAA- and AAA-rated corporate bond yields.

12. *Default return spread, DFR*: Difference between long-term corporate bond and long-term government bond returns.

13. *Inflation, $INFL$*: Calculated from the CPI (all urban consumers). We use X_{t-1} for inflation since inflation rate data are released in the following month (Goyal and Welch, 2008; Rapach et al., 2010).

14. *Stock variance, $SVAR$*: Sum of squared daily returns on the S&P 500 index.

We generate out-of-sample forecasts of the market excess return using a *recursive* window as in most of the literature. We consider one "long" and one "short" out-of-sample evaluation periods: i) 1965:01−2020:12 and ii) 2000:01−2020:12. The two corresponding initial estimation windows are originally used in Goyal and Welch (2008) and Rapach et al. (2010). We focus on the longer evaluation period as it covers more than half a century, which allows us to analyze market return predictability under a variety of economic conditions. As will be discussed in Section 1.4.4, we also consider alternative forecasting constructions, including rolling window estimation, different estimation periods, and quarterly data frequency. Those additional exercises help provide us with a good sense of the robustness of out-of-sample performance.

1.4 Out-of-Sample Results

This section resents the out-of-sample statistical results, economic values, along with a battery of robustness checks.

1.4.1 Statistical gains

Table 1.1 reports the R^2_{OS} statistics by linear trend-pooling methods. We choose the maximum moving average length L in Equation (1.3) to be 1 month, 6 months, 1 year, 2 years, and 3 years.[6] For example, the row with the heading of "6 mo" uses the moving averages of past 1, 2, 3, 4, 5 and 6 months. Together, we have $14 \times 6 = 84$ trend indicators (moving averages), and thus 84 individual market excess return forecasts. The final forecast is to pool those 84 individual ones using a simple average.

Panel A of Table 1.1 reports the results for the period 1965:01−2020:12. Column (2) reveals that, consistent with the finding in Rapach et al. (2010), the 1-month forecast combination (simple pooling) that only uses the most recent data on a set of 14 macro variables produces a significant R^2_{OS} statistic of 0.63%. In contrast, all trend-pooling methods with different macro trends substantially outperform with R^2_{OS} statistics ranging from 0.72% to 0.83%. The good out-of-sample forecasting performance of the conventional 1-month sim-

[6] According to Zhu and Zhou (2009), selecting some ex-ante value as the "optimal" lag length of the moving average might be been done only by trial and error, and might result in being suboptimal. Therefore, in this paper, we consider several fixed lag lengths for moving-average rules.

ple pooling does not preserve once we switch to the period 2000:01−2020:12. As shown in Panel B, the R^2_{OS} statistic for simple pooling is around 0.41% and insignificant. In contrast, the trend-pooling through 1- and 3-year moving averages outperform substantially, with R^2_{OS} statistics rising to significant levels of 0.56% and 0.69%, respectively.

Table 1.2 presents the R^2_{OS} statistics by combing macro trends with neural networks. Applying neural networks to the most recent data on the same set of 14 macro variables fails to outperform the historical average benchmark, as the R^2_{OS} statistics are mostly negative or positive but insignificant. On the contrary, capturing the macro trends recovers the out-of-sample predictability and generates substantial forecasting gains. In the period 1965:01−2020:12, the R^2_{OS} statistics peak at 1.31%, 1.62%, 1.90% and 1.46% for NN1, NN2, NN3 and NN4, respectively, and they are all significant at least at the 5% level. Turning to the evaluation period of 2000:01−2020:12, the most striking result is the extremely high R^2_{OS} statistics, larger than 4%, achieved by NN2 and NN4 when including the macro trends up to 1 year. The patterns documented in Table 1.2 also coincide with the finding in Gu et al. (2020) on individual stocks that shallow learning outperforms deeper learning. On average, we find that neural network performance peaks at two hidden layers then declines as more layers are added, when we consider a range of neural networks from very shallow (a single hidden layer) to deeper networks (up to five hidden layers). Overall, NN2 performs the best, followed by NN4, and NN5 performs the worst. One plausible reason is an artifact of the relatively small amount of data and tiny signal-to-noise ratio for our aggregate market return prediction setting. As argued by Gu et al. (2020), simple networks with only a few layers and nodes often perform best in small data sets. Training a very deep neural network is challenging because of the highly noncovex objective function and a large number of parameters.

On thing to note is that the poor performance of using 14 most recently macro variables is not necessarily a result of fewer input predictors. In the Internet Appendix, we apply neural networks to moving averages of different lags on the same set of 14 macro variables and compare their R^2_{OS} statistics. In other words, the input features are fixed for 14. We find that neural networks based on moving averages of different lags still significantly outperforms the historical average, though the results are not as strong as we observe in Table 1.2 when multiple macro trend signals are used all together. The results point to the economic value of incorporating both macro trends and complex nonlinear interactions, which are embedded in the neural networks but missed by other (linear) approaches. More discussions are provided in Section **??**.

Finally, Table 1.3 reports the R^2_{OS} statistics by dimension reduction. We consider PCR, PLS, Scaled PCR, LASSO, and ENet. As we mentioned in Section 1.2.3, for extracting principal components, we apply algorithm like PLS to moving averages of the same macro variable to generate one-step ahead forecast, repeat the procedure for $k = 1, \cdots, 14$, and aggregate the 14 individual forecasts with a simple mean. Thus, the row with the heading "1 mo" essentially becomes the simple pooling on 14 most recent macro variables. Compared with "1 mo", incorporating macro trends witnesses a substantial decrease in MSPE, leading to much larger and more significant R^2_{OS} statistics.

Taken together, Tables 1.1 $-$ 1.3 reveal that macro trends statistically and economically contribute to the out-of-sample aggregate market return predictability, substantially beyond using only the most recent data as in the literature. The market is more predictable than commonly believed once we incorporate trends and nonlinearity.

1.4.2 Economic gains

As described in Section 1.4.2, we measure the marginal economic benefit of the predictive ability of macro trends for a quadratic-utility investor who allocates between the market portfolio and risk-free Treasury bills. Figure 1.1 plots cumulative excess returns of one dollar for portfolios constructed based on market excess return forecast. In each panel, we compare the 1-month simple pooling with 1-year trend-pooling by either linear, PLS, ENet, or neural networks, respectively. Figure 1.1 reveals that portfolios that incorporate trend information generally exhibit superior performance compared to the simple pooling method that ignores it.

Table 1.4 reports several performance measures over the period 1965:01−2020:12. All results are annualized. We show that information of the macro trends indeed leads to sizable investment gains for a mean-variance investor from an asset allocation perspective. In general, investor portfolios based on 1-year macro trends consistently deliver higher average returns, lower standard deviation, and thus significantly larger out-of-sample Sharpe ratios, and larger CER gains. For instance, the annualized CER gains are 6.83%, 7.35%, 7.80%, and 8.76% for 1-month simple pooling, 1-year trend-pooling, 1-year PLS-pooling, and 1-year LASSO-pooling, respectively. We also observe consistently positive performance fees relative to simple pooling. For example, a quadratic-utility investor would be willing to pay an estimated 206 basis points annually to switch from the 1-month simple-pooling to the 1-year LASSO-pooling in order to acquire the longer-term trend information. Our results are robust to adjusting for transaction fees.

In summary, there are potentially large investment profits by capturing macro trends, emphasizing its important role in the market return predictability from the asset allocation perspective.

1.4.3 Variable importance

So far, results have shown that macro trends are important, and they contribute statistically and economically to the out-of-sample aggregate market return predictability. Given the $K \times L$ moving average trend indicators, an interesting question to ask is which macro variable matters, and likewise, which moving average lag matters? To understand the relative importance of each predictor, we use LASSO and ENet to assess "variable importance" to the predictability by counting how frequent each predictor is selected. Same as previous analysis, we choose the maximum backward-looking window of 1 year. Hence, we have 14 macro variables and 12 lags to rank.

The top two panels in Figure 1.2 present the relative importance of the 14 macro variables selected by either LASSO or ENet in predicting the market excess return, where the variable importance within the model is normalized to sum to one across all available predictors. Both LASSO and ENet agree that the top 3 most selected variables are *T-bill rate* (TBL), *Long term government bond return* (LTR), and *earnings-price ratio* (E/P). In contrast, the top 3 least selected ones are *net equity expansion* ($NTIS$), *stock variance* ($SVAR$), and *Inflation* ($INFL$). However, for the rest 7 variables, ENet tends to put more weight on *dividend yield (log)* (D/Y), *dividend-payout ratio (log)* (D/P), *term spread* (TMS), and *default spread return* (DFR), whereas LASSO tends to zero them out.

The middle two panels present the relative importance of the 12 moving average lags. For each lag l, we count the total number of predictors that have been selected for forecasting and use that as a proxy for "lag importance". Surprisingly, we find that the most important lag terms in market return predictability are 12-, 11-, and 10-month lags, which further indicates the importance of macro trends. By contrast, the 1-month (thus the most recent data), along with 2- and 3-month lags, turn out to be the least selected ones by both LASSO and ENet. Finally, we plot the heatmaps for the 14×12 moving averages in the bottom two panels in Figure 1.2.

The above findings echo with the Fed Chair Powell on the importance of having more than just one month's worth of data to make informed decisions, and on the importance of not overacting too much to short-term data that may not provide a clear picture of the economic outlook.

1.4.4 Robustness

In this subsection, we conduct a battery of additional robustness checks, including economic bounds, rolling window estimation, alternative estimation periods, and quarterly forecasts.

1.4.4.1 Economic bounds

In the time-series forecasts of the equity risk premium, Campbell and Thompson (2008); Pettenuzzo et al. (2014) are the pioneering examples that incorporate economic constraints into the forecasts. In this subsection, we consider imposing a simple economic lower and/or upper bounds on our trend-based forecasts. The lower bound follows the idea of Campbell and Thompson (2008) and requires the conditional mean of the equity risk premium to be non-negative. The upper bound requires the conditional Sharpe ratio to be smaller than a predetermined value. The zero lower bound is identical to the equity risk premium constraint, and the upper bound rules out that the price of risk becomes too high (Pettenuzzo et al., 2014).

Columns (3) − (5) in Table 1.1 report the R^2_{OS} statistics for economically constrained forecasts with the lower bound, the upper bound, and both. We use a value of one, suggested by Cochrane and Saa-Requejo (2000), to bound the Sharpe ratio. Across different rows, we find capturing macro trends still generate greater forecasting gains. Across different columns, we find that imposing economic constraints, especially the upper bound, yields better forecasting result for 1-month simple pooling, yet does not translate into greater forecasting gains in presence of macro trends. In other words, the predictive power of macro trends may not be subsumed by imposing economic constraints.

1.4.4.2 Rolling window forecast

For out-of-sample tests, it is equally important to have enough initial data to get a reliable estimate at the start of evaluation period and to have an evaluation period that is long enough to be representative. Researchers in this strand of literature, including Goyal and Welch (2008); Campbell and Thompson (2008); Rapach et al. (2010); Neely et al. (2014); Pettenuzzo et al. (2014), usually study a sequence of recursively generated out-of-sample equity risk premium forecasts.

Alternatively, we can adopt a rolling window and compute the corresponding out-of-sample R^2_{OS} statistics. A rolling window appears better able to accommodate changes in the parameters over time, despite at the cost of a shorter estimation sample and thus less precise parameter estimates (Pesaran and Timmermann, 2007). Column (2) in Table 1.5 reports the R^2_{OS} statistics for 1965:01−2020:12 using a 15-year rolling window. Different from the result in Table 1.1 with a recursive window, 1-month simple pooling that only uses the most recent data on a set of 14 macro variables can no longer outperform the historical average, resulting in an insignificant R^2_{OS} around 0.50%. By contrast, the trend-pooling methods continue to yield significantly positive R^2_{OS} statistics that are larger than 0.60%.[7]

1.4.4.3 Alternative estimation periods

In forecasting, the length of training and testing periods reflects different trade-offs between the desire to obtain statistical power and the desire to obtain results that remain relevant today. Apparently, a training sample starting from 1926 ensures the first objective—to obtain statistical power. We next focus on the second objective to obtain relevant results for today. For this purpose, we restrict the whole sample to 1999:12−2020:12 and evaluate the out-of-sample performance over the two more recent periods: (i) 2010:01−2020:12 and (ii) 2016:01−2020:12.

The results are presented in columns (3) and (4) in Table 1.5. As before, we find that the trend-pooling methods consistently outperforms simple pooling with much larger significant R^2_{OS} statistics. For example, in the period of 2010:01−2020:12, the R^2_{OS} statistic (around 1.19%) from 1-month simple pooling is marginally significant at the 10% level. In contrast, all trend-pooling methods achieve significant levels of R^2_{OS} statistics, ranging from 1.57% to 2.41%. A more striking result is in the second period, 2016:01−2020:12. The 1-month simple pooling leads to a *negative* R^2_{OS} around −0.06%, whereas the trend methods yield substantially positive R^2_{OS} statistics.

1.4.4.4 Quarterly forecast

We have shown that macro trends contain valuable information for predicting the monthly market excess return on an out-of-sample basis. In this subsection, we conduct the quarterly forecast with the same quarterly variables used in Rapach et al. (2010) available for

[7] We also consider rolling windows of 10 years and 20 years. The results are consistent.

$1947Q1-2020Q4$.[8] We consider the same two out-of-sample periods: $1965Q1-2020Q4$ and $2000Q1-2020Q4$.

Columns (5) and (6) in Table 1.5 report the R^2_{OS} statistics for quarterly return forecast. Consistent with the findings in Table 1.1 where monthly data are used, the trend-pooling strategies yield more significantly positive R^2_{OS} statistics. Specifically, in their original study, Rapach et al. (2010) document a reasonably significant R^2_{OS} of 3.04% at the 5% level, in the period $2000Q1-2005Q4$ (Table 1 in their paper). Here we find that the R^2_{OS} statistic drops to 0.99%, and becomes marginally significant at the 10% level, if we extend the evaluation period to 2020. The trend-pooling methods still significantly outperform the historical average.

1.5 Macro Trends and Longer-Term Information

In this section, we further examine the relevance of longer-term information for market return predictability based on macro trends. Instead of pooling moving averages of different lags (and different variables) together, we examine the moving average of the same set of 14 macro variables, based on each lag separately.

1.5.1 Moving averages and lagged values

Our paper uses a simple moving average method to capture the longer-term information in macroeconomics. A natural question is why not use lagged values directly? To answer this question, we consider the simple predictive regression such that

$$r_{t+1} = \alpha + \beta X_t^k + \varepsilon_{t+1} \tag{1.20}$$

where we replace X_t^k with either its lagged value, X_{t-l}^k, or its moving average, $MA_{t,l}^k$. We repeat the simple predictive regression for the same set of 14 macro variables and pool the individual forecasts based on either lagged values or moving averages, separately.

[8]The quarterly data include one more variable, *Investment-to-capital ratio, I/K*: ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the entire economy (Cochrane, 1991).

Table **??** compares the out-of-sample R^2_{OS} statistics from using lagged values up to 12 months with those from moving averages. We find that using moving averages generally performs better than using lagged values directly in the long run.[9] Moreover, using moving averages has economic interpretation as trends (see, e.g., Kozicki and Tinsley, 2001; Cieslak and Povala, 2015; Han et al., 2016).

1.5.2 Short-term or longer-term information?

The trend-pooling method in Equation (1.3) can be decomposed into two-step pooling. First, for each moving average lag, l, we run K simple predictive regressions with respect to $\{MA^k_{t,l}\}^K_{k=1}$, and pool the K forecasted values as

$$f(MA_l) = \frac{1}{K}\sum_{k=1}^{K}\left[\hat{\alpha}^k_{t,l} + \hat{\beta}^k_{t,l} MA^k_{t,l}\right]. \tag{1.21}$$

Equation (1.21) can be viewed as a simple pooling based on a single macro trend signal, MA_l, on a set of K macro variables.

We next repeat the above procedure for l from 1 to L, obtain L forecasts, $[f(MA_1), \cdots, f(MA_L)]$, and take another arithmetic mean to obtain the final forecast,

$$\hat{r}_{t+1|t} = \frac{1}{L}\sum_{l=1}^{L} f(MA_l). \tag{1.22}$$

Intuitively, the trend signal, MA_l, captures either short or longer-term macroeconomic information.

To test whether the longer-term signal contains more relevant information than the short term, we first directly compare the out-of-sample forecasting performance of $f(MA_l)$, from both statistical and economic aspects. Table 1.6 reports the R^2_{OS} statistics. Compared with using the most recent macro variables (MA1), we observe much larger R^2_{OS} statistics starting from MA4 to MA12. For instance, the R^2_{OS} for MA1 is 0.63% and rises to 1% for

[9] Similar results are obtained when we apply neural networks to lagged values of macro variables directly.

MA7, increasing by 58% in magnitude. Table 1.7 report the economic values by conducting asset allocation analysis. Relative to the most recent information, the longer-term macro trends on average deliver higher returns, smaller volatilities, higher Sharpe ratios, and larger CERs.

Overall, Tables 1.6 and 1.7 suggest that the longer-term macro trend contains valuable information in predicting market that can not be subsumed by the short-term (current) data. To formally test it, we next employ the encompassing test.

1.5.3 Encompassing test

Consider an optimal composite forecast of r_{t+1} as a convex combination of forecasts from two models, i and j,

$$\hat{r}^{\star}_{t+1} = (1-\lambda)\hat{r}_{i,t+1} + \lambda \hat{r}_{j,t+1}, \quad 0 \le \lambda \le 1. \tag{1.23}$$

$\lambda = 0$ suggests that model j does not contain any useful information, and thus is encompassed by model i. Likewise, $\lambda > 0$ indicates that model j is *not* encompassed by i.

To test the null hypothesis that model i encompasses j ($H_0 : \lambda = 0$), against the one-sided alternative hypothesis that the model i does not encompass j ($H_1 : \lambda > 0$), we follow Harvey et al. (1998) to compute the modified *HLN*-statistic over the out-of-sample evaluation period of T_0. Define $d_{t+1} = (\hat{e}_{i,t+1} - \hat{e}_{j,t+1})\hat{e}_{i,t+1}$, where $\hat{e}_{i,t+1} = r_{t+1} - \hat{r}_{i,t+1}$ and $\hat{e}_{j,t+1} = r_{t+1} - \hat{r}_{j,t+1}$. Let $\bar{d} = \frac{1}{T_0}\sum_k d_k$, and we compute

$$MHLN = \frac{T_0 - 1}{T_0}\left(\left[\hat{V}(\bar{d})\right]^{-\frac{1}{2}}\bar{d}\right) \sim t_{T_0-1}, \tag{1.24}$$

where $\hat{V}(\bar{d}) = \frac{1}{T_0}\hat{\phi}_0$ and $\hat{\phi}_0 = \frac{1}{T_0}\sum_k (d_k - \bar{d})^2$.

Table 1.8 reports the Harvey et al. (1998) *MHLN* statistic p-values applied to the forecasts in Equation (1.21), thus $[f(MA_1), \cdots, f(MA_{12})]$. Each entry corresponds to the null hypothesis that the row heading is encompassed by the column heading. The p-values in the right top are mostly larger than 0.8, suggesting that we can not reject the null hypothesis that short-term information, from MA1 (the most recent value) to MA4, is encompassed by longer-term, from MA5 to MA12; conversely, *p*-values in the left bottom are mostly smaller than 0.1, suggesting that we can reject the null hypothesis that longer-term trend information is encompassed by the short-term or the most recent data. In other words, in predicting the future market return, the past one or two month data is unlikely to be sufficient, and the entire macro trends are likely to matter.

1.5.4 Bias-Variance trade-off

The out-of-sample R^2_{OS} statistic essentially compares the MSPEs between the two forecasting approaches in Equation (1.13). Theil (1966) decomposes MSPE as follows,

$$MSPE = \left(\bar{\hat{e}}\right)^2 + Var(\hat{e}), \tag{1.25}$$

where $\hat{e}$ signifies the forecast error, $\left(\bar{\hat{e}}\right)^2$ is the squared forecast bias, and $Var(\hat{e})$ is the forecast variance.

Figure 1.3 is a scatterplot depicting the forecast variance and the squared forecast bias for the forecasts based on the historical average and simple pooling of different macro trend signals. The figure shows that the historical average is at the top right corner, suggesting both the largest forecast variance and the largest squared forecast bias. In contrast, pooling across 14 macro variables is concentrated at the (bottom) left corner. To avoid cluttering the diagram, we remove the historical average. As shown in the right panel of Figure 1.3, the simple pooling forecasts based on MA4 or longer-term moving averages yield much smaller forecast variance. Because the variance is much larger than the squared forecast bias, the reduction in variance dominates the performance and thus leads to a much larger R^2_{OS} statistics. Among them, MA10, MA11, and MA12 achieve both smaller variance and smaller squared bias, justifying the importance of longer-term macro trend signals in market excess return predictability. Overall, Figure 1.3 suggests that capturing macro trends helps to regulate more effectively the forecast variability and is thus more stable.

Apart from decomposing the R^2_{OS} statistics from a single macro trend signal, we can also apply it to the linear, nonlinear, and dimension reduction trend-pooling methods that are introduced and discussed in Section 2.3. To avoid cluttering the diagram, the left panel in Figure 1.4 plots the scatters of 1-month simple pooling, 1-year linear trend-pooling, 1-year LASSO-pooling, and 1-year ENet-pooling. Applying shrinkage significantly reduces the forecast variance at the cost of slightly larger forecast bias. Notwithstanding, the substantial reduction in forecast variance helps to offset the increase in squared forecast bias, thereby leading to much smaller MSPEs and much larger R^2_{OS} statistics. Likewise, the right panel plots the scatters of 1-month simple pooling, and 1-year trend-pooling by linear combination, PLS, PCR, and neutral networks (2-hidden layers). We find that extracting principal components or applying neural networks maintains a similar forecast bias square, but substantially reduces the forecasting variance. Remarkably, combing neural networks with macro trends delivers the highest forecasting precision (the lowest forecasting variance), due to the potential nonlinear interactions.

1.5.5 Performance over business cycles

To get a visual impression of the forecasting performance over the real economy (business cycles), we compute the square error difference relative to the historical average benchmark,

$$\text{square error difference} = (r_t - \bar{r}^{HA}_{t|t-1})^2 - (r_t - \hat{r}_{t|t-1})^2, \tag{1.26}$$

and plot the cumulative difference curve. Intuitively, when the curve increases, the competing forecast outperforms the historical average, while the opposite holds when the curve decreases. If the curve is higher (lower) at the end of the out-of-sample period than at the beginning, the forecasting approach (historical average) has a lower MSPE, leading to a positive (negative) R^2_{OS} statistic.

In total there are eight recessions over the out-of-sample period, 1965:01−2020:12, with business-cycle peaks (troughs) occurring at 1969:12 (1970:11), 1973:11 (1975:03), 1980:01 (1980:07), 1981:07 (1982:11), 1990:07 (1991:03), 2001:03 (2001:11), 2007:12 (2009:06), and 2020:02 (2020:04). Figure 1.5 plots the cumulative error difference curves for the simple pooling based on MA1, MA7, and MA12 of the same set of 14 macro variables, along with the vertical lines indicating NBER-dated business-cycle peaks and troughs. Consistent with the R^2_{OS} statistics documented in Panels A of Table 1.6, the figure shows that the MA7 trend signal achieves the largest cumulative square error difference at the end of 2020 (and thus the largest R^2_{OS}), followed by the MA12, whereas the MA1, which uses the most recent data only, yields the smallest cumulative error difference.

All three curves, however, become markedly negatively sloped starting from the mid-1990s and reach the trough around 2000. This is anticipated as the literature (see, e.g., Rapach et al., 2010; Neely et al., 2014) have demonstrated that the market return predictability is mostly concentrated during recessions. Too see it, we compute R^2_{OS} statistics separately for cyclical expansions and recessions. As shown in columns (3) and (4) in Table 1.6, the out-of-sample predictive ability is uniformly much stronger for recessions than for expansions. Within the recession periods, the longer-term macro trends produce much larger and more significant $R^2_{OS,REC}$ statistics than MA1. Moreover, we also observe that longer-term macro trends, starting from 5-month, also have predictive power during the expansions, with three of them significant at the 5% level. Utilizing macro trends helps to track more closely the important macroeconomic fluctuations and thus contributes to out-of-sample predictability in both recessions and expansions, substantially beyond using only the most recent data as in the literature.

To conclude the discussion, we also plot and compare the cumulative predictive errors among different trend-pooling methods introduced in Section 2.3. Figure 1.6 displays that, among different trend methods, ENet dominates PLS, which strictly dominates the linear combination. Furthermore, we find that the performance of neural networks heavily depends on the real economic conditions: it underperforms the historical average in expansions, mainly in the forecasting period of 1990s, but substantially outperforms the historical average in recessions because of the oil crisis in mid-1970s and the global financial crisis in 2008/09.

Taken together, in this section, we show that macro trends with longer-term information contain valuable information that can not be ignored in predicting equity risk premium. However, we also argue that when investors make real time decisions, it is important to use multiple macro trend signals together, from 1-month to L-month, rather than rely on one particular signal.

1.6 Conclusion

Echoing with the Fed Chair Powell's typical view that the Federal Reserve would not react too much to short-term fluctuations in the data, as they could be uninformative on the overall economic outlook, we provide the first evidence on how macro trends affect the equity risk premium, uncovering its significant and unrecognized role in market return predictability. To capture the trend in macroeconomics, we apply the technical rule used by practitioners—the moving average—to multiple macroeconomic variables to form various trend indicators. We find that macro trends statistically and economically contribute to the out-of-sample aggregate market return predictability, substantially beyond using only the most recent data as in the literature. Moreover, by applying neural networks, a powerful machine learning method, to various macro trends, we provide another novel evidence that nonlinear interaction matters in the aggregate market return predictability. Our study shows that the market is more predictable than commonly believed, once we incorporate trends and nonlinearity.

Table 1.1: R^2_{OS} statistics by trend-pooling
This table reports out-of-sample R^2 statistic, R^2_{OS} (in percent), from trend-pooling by using moving-average-based predictors up to *Lags* months together. The lower bound requires the forecasted equity premium to be non-negative. The upper bound requires the conditional Sharpe ratio to be no larger than one. “Lags up to" denote the maximum moving average lag starting from 1 month. Based on Clark and West (2007) test, *, **, and *** indicate significance for positive R^2_{OS} at the 10%, 5%, and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)
Lags up to	Unconstrained	Lower Bound	Upper Bound	Lower & Upper
Panel A: 1965:01−2020:12 out-of-sample period				
1 mo	0.63***	0.55***	0.70***	0.63***
6 mo	0.72***	0.66***	0.71***	0.66***
1 yr	0.83***	0.77***	0.78***	0.74***
2 yr	0.81***	0.76***	0.77***	0.74***
3 yr	0.74***	0.70***	0.72**	0.70***
Panel B: 2000:01−2020:12 out-of-sample period				
1 mo	0.41	0.31	0.58*	0.49**
6 mo	0.47*	0.38*	0.58*	0.49**
1 yr	0.56**	0.45**	0.70**	0.59**
2 yr	0.61*	0.44*	0.79**	0.61**
3 yr	0.69**	0.51**	0.83**	0.65**

Table 1.2: R^2_{OS} statistics by neutral networks
This table reports out-of-sample R^2 statistic, R^2_{OS} (in percent), by combing multiple macro trends with the neural networks. We use neural networks with 1 to 5 layers (NN1−NN5). “Lags up to" denote the maximum moving average lag starting from 1 month. Based on Clark and West (2007) test, *, **, and *** indicate significance for positive R^2_{OS} at the 10%, 5%, and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
Lags up to	NN1	NN2	NN3	NN4	NN5
Panel A: 1965:01−2020:12 out-of-sample period					
1 mo	-0.07	-0.19	0.40*	-2.09	-1.59
6 mo	0.59**	0.35	1.90***	0.57**	-0.32
1 yr	1.31***	1.62**	0.87***	1.46**	0.11**
2 yr	0.74**	0.89***	0.22**	0.66**	-0.02
3 yr	0.73**	0.69**	0.47**	0.09*	0.00*
Panel B: 2000:01−2020:12 out-of-sample period					
1 mo	0.31	0.90	0.52	-1.72	-3.49
6 mo	1.22*	0.87	3.90**	1.78*	-0.29
1 yr	2.69***	4.17**	1.45**	4.13**	1.39**
2 yr	2.14***	1.63***	1.00*	0.82*	0.65*
3 yr	1.61**	1.66**	1.77*	1.56**	1.27**

Table 1.3: R^2_{OS} statistics by dimension reduction
This table reports out-of-sample R^2 statistic, R^2_{OS} (in percent) from applying dimension reduction techniques to $K \times L$ moving-average-based predictors. "Lags up to" denote the maximum moving average lag starting from 1 month. Based on Clark and West (2007) test, *, **, and *** indicate significance for positive R^2_{OS} at the 10%, 5%, and 1% levels, respectively.

(1) Lags up to	(2) PCR	(3) PLS	(4) S-PCR	(5) LASSO	(6) ENet
Panel A: 1965:01−2020:12 out-of-sample period					
1 mo	0.63***	0.63***	0.63***	-0.18	-0.23
6 mo	0.76***	0.83***	0.83***	0.96**	0.71**
1 yr	0.95***	0.99***	0.95***	1.27***	1.37***
2 yr	0.95***	1.02***	1.02***	0.75**	0.84**
3 yr	0.79***	0.91***	0.96***	0.41*	0.56*
Panel B: 2000:01−2020:12 out-of-sample period					
1 mo	0.41	0.41	0.41	-0.26	-0.09
6 mo	0.47*	0.50*	0.43*	0.40	0.14
1 yr	0.59**	0.65**	0.55*	0.75*	0.91*
2 yr	0.61*	0.62*	0.60*	0.65*	1.14**
3 yr	0.68**	0.62*	0.65*	0.68**	1.02**

Table 1.4: Economic values relative to 1-month simple pooling, 1965:01−2020:12
This table reports various economic measures for a mean-variance investor with relative risk aversion coefficient of three who allocates *monthly* between equities and risk-free bills based on equity risk premium forecasts. We consider the simple forecasts that rely on only the 14 most recent lagged values, as well as 1-year trend-based methods, including trend-, PCR-, PLS-, sPCR-, LASSO, ENet-pooling, as well as neural networks. The performance measures include out-of-sample average return, standard deviation, Sharpe ratio (SRatio), certainty equivalent return (CEQ), and Fleming et al. (2001) performance fee (Fee). All results are annualized. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

(1)	(2) Avg. Ret (%)	(3) S.D. (%)	(4) SRatio	(5) CEQ (%)	(6) SRatio diff	(7) CEQ diff	(8) Fee (bps)
Panel A: 1-month simple pooling							
Linear	6.32	16.33	0.39	6.83			
Panel B: Trend-pooling combing macro trends from 1-month to 1-year							
Linear	6.82	16.29	0.42	7.35	0.03*	0.51	49.71
PCR	7.00	16.09	0.44	7.63	0.05**	0.79	70.59
PLS	7.07	15.87	0.45	7.80	0.06***	0.96	79.65
S-PCR	7.02	15.92	0.44	7.73	0.05***	0.89	74.57
LASSO	8.43	16.72	0.50	8.76	0.12*	1.92	206.05
ENET	8.19	16.13	0.51	8.80	0.12*	1.97	189.04
NN1	7.47	12.75	0.59	9.55	0.20**	2.72	154.07
NN2	8.00	14.23	0.56	9.50	0.17**	2.65	192.00
NN3	8.16	15.42	0.53	9.11	0.15*	2.34	194.32
NN4	8.09	15.55	0.52	8.99	0.14**	2.22	185.63

Table 1.5: R^2_{OS} statistics for robustness
This table reports out-of-sample R^2 statistics, R^2_{OS} (in percent) for alternative forecasting constructions. All statistics are for the period that starts at “Forecast Begin” and ends on December 2020. “Lags up to" denote the maximum moving average lag starting from 1 month. Based on Clark and West (2007) test, *, **, and *** indicate significance for positive R^2_{OS} at the 10%, 5%, and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
Data	Monthly	Monthly	Monthly	Quarterly	Quarterly
Sample begin	1926:12	1999:12	1999:12	1947Q1	1947Q1
Forecast begin	1965:12	2010:01	2016:01	1965Q1	2000Q1
Estimation	Rolling	Recursive	Recursive	Recursive	Recursive
Lags up to					
1 mo	0.50	1.19*	-0.06	2.44***	0.99*
6 mo	0.63*	1.57***	0.47	2.63***	1.33**
1 yr	0.82**	1.99***	0.88*	2.39***	1.49**
2 yr	0.72**	2.17**	0.98*	1.84***	1.33**
3 yr	0.71**	2.41**	1.06**	1.47**	1.24*

Table 1.6: R^2_{OS} statistics by a single macro trend signal, 1965:01−2020:12

This table reports out-of-sample R^2 statistic, R^2_{OS} (in percent), from forecast combination of 14 macro variables of the same moving average lag. We also report the R^2_{OS} for NBER-dated expansions and recessions, separately. "Lag" denote the moving average lag. Based on Clark and West (2007) test, *, **, and *** indicate significance for positive R^2_{OS} at the 10%, 5%, and 1% levels, respectively.

(1) Lag	(2) Overall	(3) Expansions	(4) Recessions
1 mo	0.63***	0.18	1.70***
2 mo	0.58**	0.26	1.36***
3 mo	0.60**	0.18	1.60***
4 mo	0.67***	0.27	1.62***
5 mo	0.87***	0.43*	1.95***
6 mo	0.94***	0.50**	1.98***
7 mo	1.00***	0.58**	1.99***
8 mo	0.96***	0.49**	2.10***
9 mo	0.95***	0.46*	2.11***
10 mo	0.88***	0.34*	2.17***
11 mo	0.86***	0.32*	2.15***
12 mo	0.91***	0.32*	2.32***

Table 1.7: Economic values of a single macro trend signal, 1965:01−2020:12
This table reports various economic measures for a mean-variance investor who allocates *monthly* between equities and risk-free bills. “Lag" denote the moving average lag. We compute out-of-sample average return, standard deviation, Sharpe ratio (SRatio), certainty equivalent return (CEQ), and Fleming et al. (2001) performance fee (Fee). All results are annualized. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag	Avg. Ret (%)	S.D. (%)	SRatio	CEQ (%)	SRatio diff	CEQ diff	Fee (bps)
			Panel A: 1-month simple pooling				
1 mo	6.32	16.33	0.39	6.83			
			Panel B: *L*-month simple pooling (a single macro trend)				
2 mo	6.15	16.44	0.37	6.61	-0.01	-0.22	-18.24
3 mo	6.22	16.03	0.39	6.87	0.00	0.04	-7.02
4 mo	6.47	16.35	0.40	6.98	0.01	0.15	14.94
5 mo	6.97	16.54	0.42	7.38	0.03**	0.55	62.41
6 mo	7.02	16.36	0.43	7.51	0.04**	0.68	68.85
7 mo	7.18	16.46	0.44	7.63	0.05**	0.79	83.92
8 mo	7.18	16.30	0.44	7.71	0.05**	0.87	86.12
9 mo	7.19	16.42	0.44	7.66	0.05*	0.82	85.44
10 mo	6.90	15.98	0.43	7.58	0.04	0.75	61.75
11 mo	6.95	16.07	0.43	7.59	0.05*	0.76	66.12
12 mo	6.98	15.83	0.44	7.73	0.05*	0.89	71.41

Table 1.8: Forecast encompassing test, *MHLN* statistic *p*-values, 1965:01−2020:12,
This table reports *p*-values for the Harvey et al. (1998) *MHLN* statistic. We compare the out-of-sample forecasts by pooling 14 macroeconomic variables of the same moving average lag, *l*, ranges from 1 to 12 months. The statistic corresponds to an upper-tail test of the null hypothesis that the forecast given in the row heading is encompassed by the forecast in the column heading.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	1 mo	2 mo	3 mo	4 mo	5 mo	6 mo	7 mo	8 mo	9 mo	10 mo	11 mo	12 mo
1 mo		0.28	0.33	0.60	0.95	0.97	0.97	0.94	0.88	0.84	0.81	0.85
2 mo	0.67		0.55	0.78	0.99	0.99	0.99	0.97	0.94	0.91	0.89	0.92
3 mo	0.60	0.40		0.86	0.99	1.00	1.00	0.99	0.95	0.93	0.91	0.93
4 mo	0.31	0.15	0.10		1.00	1.00	1.00	0.98	0.94	0.89	0.86	0.89
5 mo	0.02	0.01	0.00	0.00		0.81	0.88	0.74	0.63	0.43	0.38	0.48
6 mo	0.01	0.00	0.00	0.00	0.14		0.78	0.56	0.46	0.25	0.23	0.33
7 mo	0.01	0.00	0.00	0.00	0.08	0.17		0.27	0.25	0.10	0.08	0.18
8 mo	0.02	0.01	0.00	0.01	0.17	0.33	0.66		0.37	0.14	0.13	0.25
9 mo	0.04	0.02	0.02	0.02	0.24	0.40	0.64	0.55		0.16	0.16	0.30
10 mo	0.07	0.04	0.03	0.06	0.42	0.62	0.83	0.80	0.78		0.35	0.61
11 mo	0.09	0.04	0.04	0.07	0.46	0.65	0.85	0.80	0.76	0.59		0.81
12 mo	0.06	0.03	0.02	0.05	0.33	0.49	0.68	0.62	0.57	0.30	0.14	

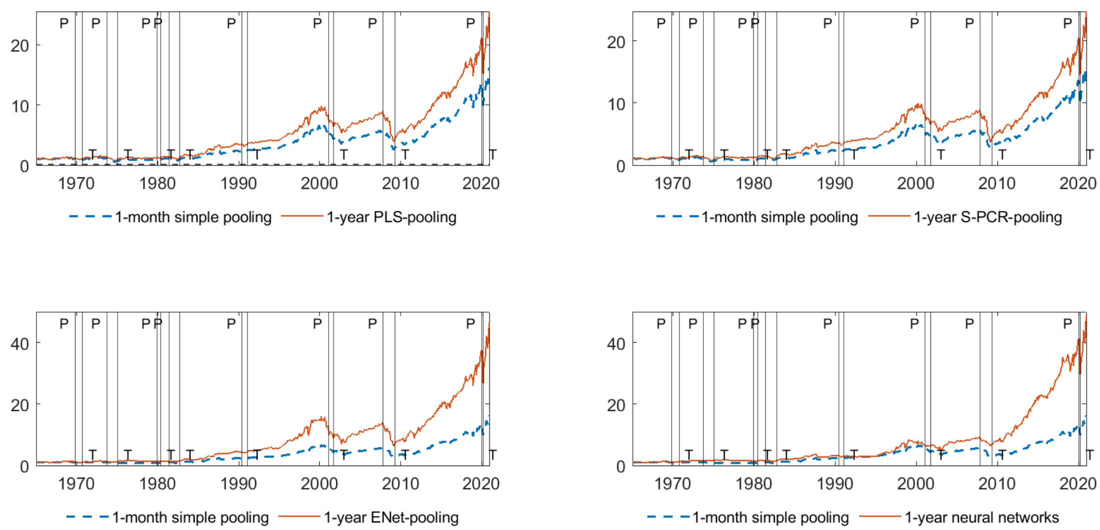
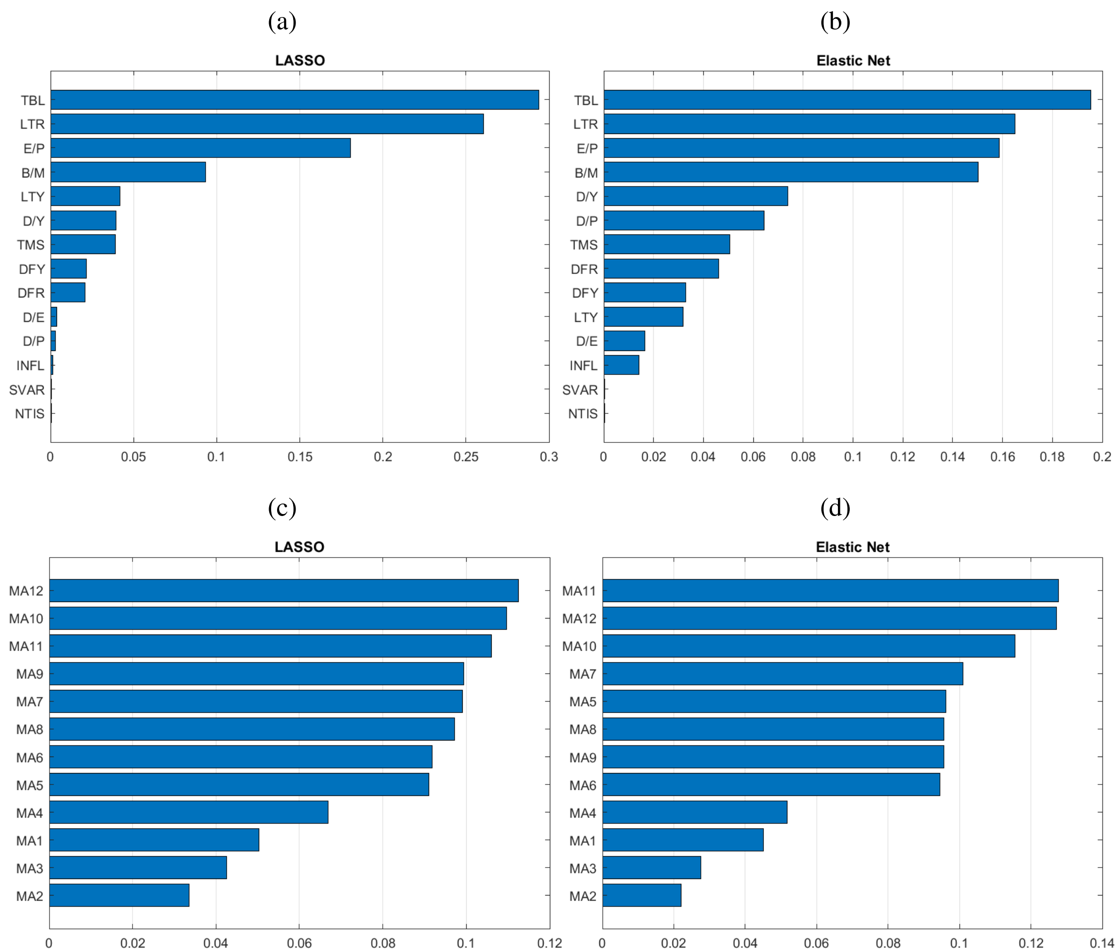


Figure 1.1: Cumulative excess returns for portfolios constructed based on macro trends, 1965:01−2020:12

Each panel depicts the cumulative excess return for a portfolio constructed based on forecasts from 1-month simple pooling or 1-year macro trends, including PLS, Scaled PCR, ENet, and neural networks. Vertical lines indicate NBER-dated business-cycle peaks (P) and troughs (T).

Figure 1.2: Selection frequency by LASSO and elastic net, 1965:01−2020:12

This figure demonstrates the selection frequency of 14×12 moving averages on a set of 14 variables in based on LASSO or ENet for the out-of-sample period 1965:01−2020:12.



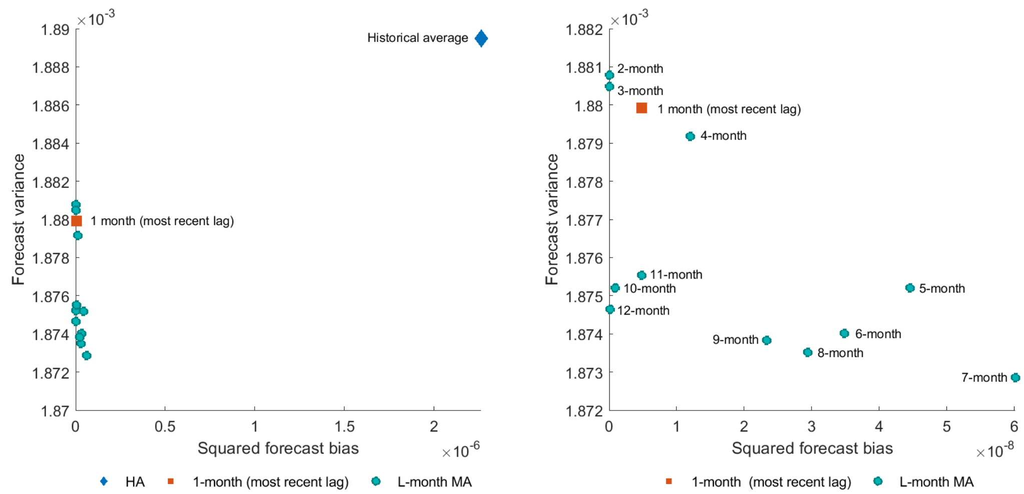


Figure 1.3: Scatterplot of forecast variances and squared forecast biases

This figure plots the bias-variance decomposition of the MSPEs based on historical average and forecast combination of 14 macroeconomic predictors of different moving average lags, in the period of 1965:01−2020:12

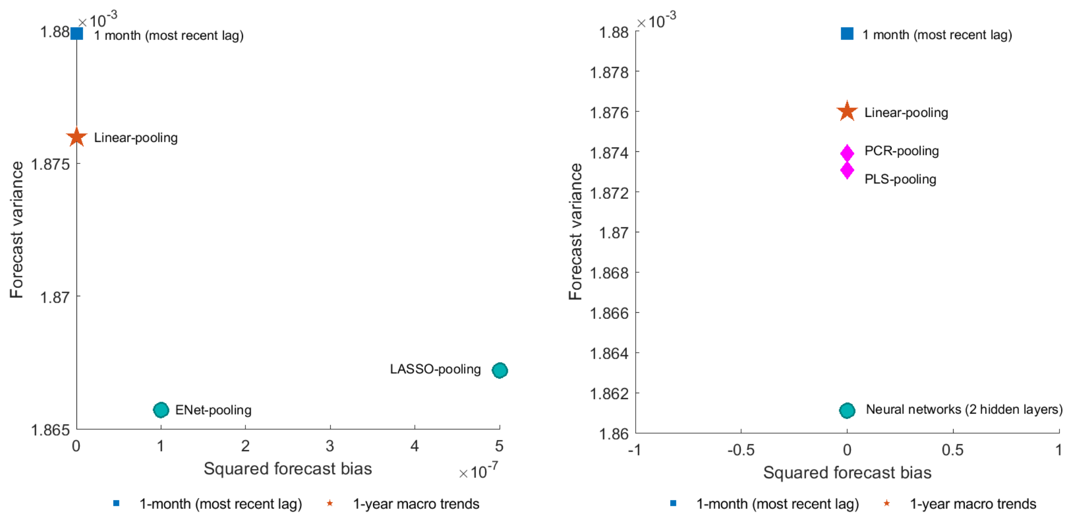


Figure 1.4: Bias-Variance decomposition for trend-based machine learning methods

This figure plots the bias-variance decomposition of the MSPEs based on 1-month forecast combination or 1-year macro trends, including ENet-pooling, PLS-pooling, and neural networks, in the out-of-sample period 1965:01−2020:12.

Figure 1.5: Cumulative square error difference from a single macro trend signal

This figure plots the cumulative square prediction error based on the forecast of 14 macro variables that are either 1-month (most recent), 7-month, or 12-month moving averages. The out-of-sample evaluation period is 1965:01−2020:12.

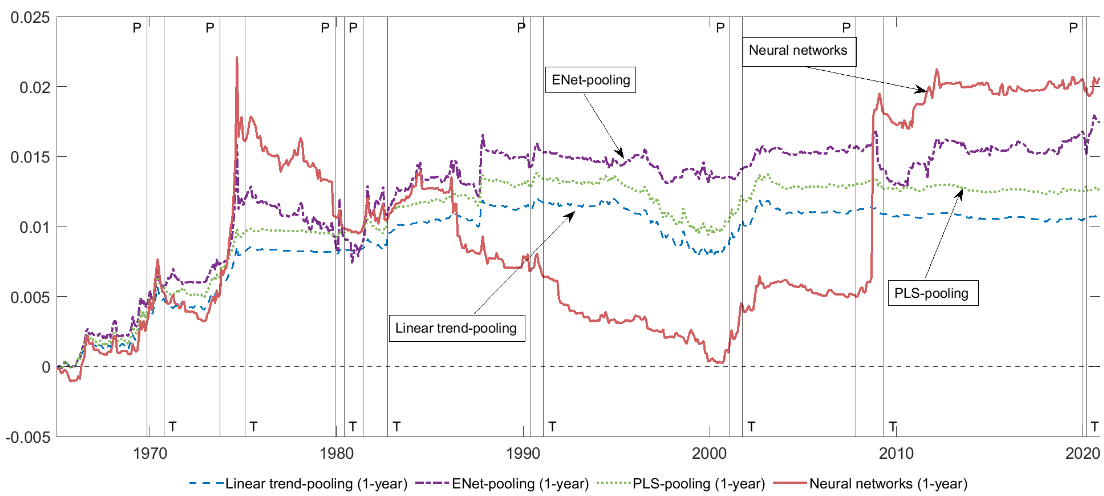


Figure 1.6: Cumulative square error difference of 1-year trend-pooling

This figure plots the cumulative square prediction error of the forecast based on 1-month forecast combination or 1-year macro trends, including ENet-pooling, PLS-pooling, and neural networks. The out-of-sample evaluation period is 1965:01−2020:12.

CHAPTER 2: EQUITY FORWARD RETURN FROM DERIVATIVES

2.1 Introduction

As argued by Miller (1999), "simply averaging the returns of the last few years, along the lines of the examples in the Markowitz paper (and later book), won't yield reliable estimates of the return *expected* in the future" (page 97). Since the derivatives market provides forward-looking information related to the expected return, previous studies have successfully derived the *expected spot return* from derivatives markets. For instance, Martin (2017); Chabi-Yo and Loudis (2020) for the aggregate market; Martin and Wagner (2019); Kadan and Tang (2020) for individual stocks; and Kremens and Martin (2019) for currencies.

In this paper, we introduce the notion of a *forward equity return* and establish its relationship to certain derivative securities. Specifically, for any positive number n, the forward return $\mathbb{E}_t[R_{t+n\to t+n+1}]$ is the expectation conditional on time t of the future return $R_{t+n\to t+n+1}$ of the underlying asset over a future time interval from $t+n$ to $t+n+1$. We develop a methodology to measure forward returns implied in information from the derivatives market, including prices of index options and VIX-derivatives. Furthermore, we can derive all higher moments of future returns using derivatives market information. Our approach even reveals new information about spot returns from the derivatives not previously studied in the literature, such as the conditional correlation between two spot returns and the joint distribution between two consecutive returns or two spot returns.[1] Importantly, our approach does not require us to impose any distributional assumptions on the aggregate equity market.

[1]Remarkable exceptions in the literature include Martin (2021), and Chabi-Yo (2019). We will explain these related studies and the distinct features of our paper.

Ultimately, we can construct an entire term structure of forward returns starting from the expected spot return when $n = 0$.[2] Knowing such a term structure of forward returns, we can investigate how a future return $R_{t+n \to t+n+1}$ *dynamically evolves* with market information available at time t. As an example, what is the autocorrelation coefficient between $R_{t+n \to t+n+1}$ and $R_{t+n+1 \to t+n+2}$? By answering this question, we can study the momentum or reversal patten of the equity market from a forward-looking derivative perspective.[3] Similarly, we can design dynamic trading strategies based on the conditional view of the aggregate market's future returns from derivatives.

We first express the equity index's forward return in terms of available VIX-derivatives (VIX-futures and VIX-options). Because of its high volume and vast liquidity, incorporating information from the VIX derivatives market is essential in constructing a complete picture of the equity market. In the expression we derive, VIX derivatives play an analogous role for forward returns as index options play for the expected spot return (see Martin, 2017). All quantities in this expression are observable in real-time. Thus, we can compute these forward returns in real-time as well. Since this expression depends on VIX-derivatives, we call it the *VIX-approach* for forward returns.

We present three applications of the VIX-approach. In the first application, we document a *pro-cyclical* term structure of forward returns: upward sloping in good times but downward sloping in bad times. Several studies have documented the average shape of the term structure of the equity risk premium, which is the expected spot return of dividend strips across different maturities. For instance, Binsbergen et al. (2012); Binsbergen and Koijen (2017) document that the equity term structure is downward sloping, on average.

[2]Our approach to modeling the term structure of equity forward returns, $\mathbb{E}_t[R_{t+n \to t+n+m}], m > n > 0$, instead of the sequence of expected spot returns $\mathbb{E}_t[R_{t \to t+n}], n > 0$, is conceptually similar to forward rate models compared to spot rate models in fixed income (see, for instance, Heath et al., 1992; Duffie and Singleton, 1999). However, there are substantial differences between equity returns and interest rates; and the forward equity return cannot be derived from the equity spot return. Specifically, spot returns are static, similar to the current yield curve in the fixed income market. In contrast, the term structure of forward returns provides a dynamic movement of the equity return, which resembles the movement of the yield curve.

[3]There has been extensive research about the realized autocorrelation using historical equity index data (see, for instance, Lo and MacKinlay, 1988, 1990; Fama and French, 1988b; Poterba and Summers, 1988; Moskowitz et al., 2012). These studies do not use derivatives.

Moreover, Gormsen (2021) finds that the equity term structure is downward sloping in good times, but upward sloping in bad times, and thus counter-cyclical. We document a new stylized fact about the shape of the term structure of forward equity returns. The key difference between the previous studies and ours is that, in our setting, we recover expected future returns on the aggregate market, whereas these previous studies focus on the spot return of a dividend-strip (Binsbergen et al., 2012; Binsbergen and Koijen, 2017), the futures of the dividend-strip (Bansal et al., 2021), or a short-maturity asset in excess of long-maturity asset (Gormsen, 2021). Moreover, we use VIX-derivatives, whereas dividend strips, index options, or asset pricing models are used in previous studies.

In the second application, we use the VIX-approach to compute conditional autocorrelation in real-time. Predicting the expected market return with past return observations has been a challenge for researchers and practitioners for decades. Can past returns forecast future returns? Are the returns of a given stock market index autocorrelated? How can we estimate the unconditional autocorrelation coefficients of market returns? Despite extensive research about the realized autocorrelation using historical data (see, for instance, Lo and MacKinlay, 1988, 1990; Fama and French, 1988b; Poterba and Summers, 1988; Moskowitz et al., 2012), the literature still offers no clear guidance as findings have varied depending on the horizon studied and on the sample frequency selected (Campbell, 2017; Baltussen et al., 2019). Moreover, what remains unclear is how to infer the *forward-looking autocorrelation* perceived by investors, as the true autocorrelation may diverge significantly from zero and fluctuate over time (LeRoy, 1973). Empirically, we document *significantly negative* autocorrelation on the S&P 500 index from index options and VIX-derivatives. For instance, the conditional autocorrelation between two consecutive monthly returns, $corr_t\left(R_{t\to t+1mo}, R_{t+1mo\to t+2mo}\right)$, is on average -20.90% with a t-stat of -18.10. On average, the market autocorrelation on the S&P 500 index is around -20% to -40%, suggesting a robust short-term reversal from the perspective of derivatives.

The third application illustrates the economic value of forward returns from derivatives.

Specifically, we construct a reversal signal to trade the market. This signal relies on the real-time autocorrelation identified from the derivatives market in the second application. We find that the reversal signal predicts market downturns well, particularly when the market declines significantly in the next month. Furthermore, we show that the corresponding market timing strategy is conservative and delivers higher Sharpe ratios compared to the buy-and-hold benchmark strategy. Moreover, the economic value of this timing strategy can be substantial during prolonged market downturns. For example, investors are willing to pay as much as 11% per annum to switch from the buy-and-hold strategy to the derivative-based market-timing strategy during the 2008/09 global financial crisis, January 2008−June 2009.

The VIX-approach relies on the assumption that $\theta_t \equiv corr_t^Q(R_{t\to t+1}, R^2_{t+1\to t+2}) = 0$, which states that the *risk-neutral conditional correlation coefficient* between $R_{t\to t+1}$ and $R^2_{t+1\to t+2}$ is zero. This risk-neutral zero correlation assumption seems rather technical and restrictive as our objective is to use available market information alone and avoid using any model assumption about the future return. Moreover, there are no available derivatives in the market yet to reveal this risk-neutral correlation coefficient directly. Therefore, we need to (1) justify the VIX-approach empirically for the equity index and (2) introduce econometric methods to estimate this risk-neutral correlation coefficient.

For this purpose, we derive an alternative expression for forward returns using another set of derivatives data - index option prices and their gammas. A disadvantage of this approach is that it cannot be applied in real-time. Nevertheless, by comparing this expression for forward returns to those derived using the VIX-approach, we can estimate past values of θ_t and analyze its time series properties. In the end, we document that the VIX-approach is reasonably good as the sample average of θ_t is fairly close to zero. Moreover, if we use the estimated θ_t, we derive a more robust expression of forward return. In the latter expression of forward return, we use all historical and current index option derivatives (option prices and gammas) and the VIX-derivatives. In this paper, we refer to this more general approach to estimating forward returns as the *VIX-Gamma approach*.

We implement the VIX-Gamma approach in the market-timing strategy and demonstrate that its economic value is even more significant than in the VIX-approach. For instance, we find that investors are willing to pay 31.153% per annum to switch from the buy-and-hold portfolio to the VIX-Gamma-based trading strategy, more than double the amount for the VIX-approach. Moreover, VIX-Gamma-based trading strategy produces a positive average return and a positive Sharpe ratio, even during the 2008/09 global financial crisis period.

Our study is related to Martin (2021), which reduces the conditional expected future return to the no-arbitrage price of a “forward-start option". Since the forward-start option is traded only in over-the-counter markets, he obtains quoted prices from a sophisticated investment bank for a small number of days. In contrast, we present two new expressions of the expected future return that rely on VIX-derivatives and index options—pricing data for both are publicly available. Notably, in our second application, the autocorrelation based on the exchange-traded derivatives is comparable (in magnitude) to that based on the over-the-counter derivatives in Martin (2021). Moreover, we demonstrate novel implications for the equity term structure, investment trading strategy, and risk-neutral density. Another related working paper is by Chabi-Yo (2019), who derives lower and upper bounds, varying from -28% to -3% with a mean value of -14%, on the market autocorrelation with index option prices. Other studies have documented how contingent claims can be used to elicit valuable forward-looking information about the market’s expected spot returns (Ross, 2015; Borovička et al., 2016; Jensen et al., 2019; Heston, 2021; Bakshi et al., 2022). But these authors do not study forward returns.

2.2 Theory

For a discrete time subscript t, S_t denotes the time-t price of the stock index. $R_{t\to t+1} = \frac{S_{t+1}}{S_t}$ is the gross market return over the time period from t to the next time $t+1$, and $R_{f,t\to t+1}$ is the gross risk-free return over the same time period. We denote the real-world probability measure by $\mathbb{P}$, and the information set at time t by $\mathscr{F}_t$. Let (M_t) be a pricing kernel process and $m_{t,t+1} = \frac{M_{t+1}}{M_t}$ be the stochastic discount factor (SDF) over the period from t to $t+1$. Consequently, the risk-neutral probability measure $\mathbb{Q}$ satisfies

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathscr{F}_{t+1}} = R_{f,t\to t+1} m_{t,t+1}.$$

For any $f(S_{t+1}) \in \mathscr{F}_{t+1}$ with suitable integrability condition, its conditional expectation under $\mathbb{P}$ is given by

$$\mathbb{E}_t^{\mathbb{P}}[f(S_{t+1})] = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}} f(S_{t+1})\right] = \frac{1}{R_{f,t\to t+1}} \mathbb{E}_t^{\mathbb{Q}}\left[\frac{f(S_{t+1})}{m_{t,t+1}}\right]. \tag{2.1}$$

Equation (2.1) states that a conditional expectation of $f(S_{t+1})$ under the real-world probability measure $\mathbb{P}$ is the no-arbitrage time-t price of a contingent claim with payoff $\frac{f(S_{t+1})}{m_{t,t+1}}$ at time $t+1$.[4] We use the notation $\mathbb{E}_t^{\mathbb{P}}[\cdot]$ to highlight the fact that those quantities are under the real-world probability measure. Henceforth, we drop the superscript and use $\mathbb{E}_t[\cdot]$ to denote conditional expectation under the $\mathbb{P}$-measure.

Consider a log-utility-based SDF such that,

$$m_{t,t+1} = \left(\frac{S_t}{S_{t+1}}\right).$$

By Equation (2.1), we obtain

$$\mathbb{E}_t[R^n_{t\to t+1}] = \frac{1}{R_{f,t\to t+1}} \mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{S_{t+1}}{S_t}\right)^{n+1}\right], \tag{2.2}$$

where the right-hand side of the above equation can be synthesized in terms of time-t prices of index call options $C_{t\to t+1}(K)$ that expire at $t+1$. Precisely,

$$\mathbb{E}_t[R^n_{t\to t+1}] = \frac{(n+1)n}{S_t^{n+1}} \int_0^\infty K^{n-1} C_{t\to t+1}(K) dK. \tag{2.3}$$

[4]Bakshi et al. (2022) refer to this known result as the *inverting the Girsanov theorem*.

This equation for the expected spot return (and higher moments) in terms of options is well studied in the literature (see, e.g. Bakshi et al., 2003; Bakshi and Madan, 2000; Carr and Madan, 1999; Carr et al., 1998; Martin, 2017; Martin and Wagner, 2019). We next move to the forward return (i.e., conditional expected future return).

2.2.1 VIX-approach

By Equation (2.1), the forward return is written as[5]

$$\mathbb{E}_t[R_{t+1\to t+2}] = \frac{1}{R_{f,t\to t+2}}\mathbb{E}_t^{\mathbb{Q}}\left[\frac{R_{t+1\to t+2}}{m_{t,t+2}}\right]. \tag{2.4}$$

Rewriting the right-hand side, we obtain

$$\begin{aligned}
\mathbb{E}_t[R_{t+1\to t+2}] &= \frac{1}{R_{f,t\to t+2}}\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{S_{t+2}}{S_{t+1}}\right)\times\left(\frac{S_{t+2}}{S_t}\right)\right] \\
&= \frac{1}{R_{f,t\to t+2}}\mathbb{E}_t^{\mathbb{Q}}\left[(R_{t+1\to t+2})^2\times R_{t\to t+1}\right], \\
&= \frac{1}{R_{f,t\to t+2}}\left\{\mathbb{E}_t^{\mathbb{Q}}\left[(R_{t+1\to t+2})^2\right]\times\mathbb{E}_t^{Q}[R_{t\to t+1}] + Cov_t^{Q}\left((R_{t+1\to t+2})^2, R_{t\to t+1}\right)\right\}.
\end{aligned}$$

Then,

$$\mathbb{E}_t[R_{t+1\to t+2}] = \frac{1}{R_{f,t+1\to t+2}}\mathbb{E}_t^{\mathbb{Q}}[(R_{t+1\to t+2})^2] + \frac{\theta_t}{R_{f,t\to t+2}}\sqrt{Var_t^{\mathbb{Q}}(R_{t+1\to t+2}^2)}\sqrt{Var_t^{\mathbb{Q}}(R_{t\to t+1})}, \tag{2.5}$$

where $\theta_t \equiv corr_t^{\mathbb{Q}}\left(R_{t\to t+1}, R_{t+1\to t+2}^2\right)$ is the correlation coefficient between the spot return, $R_{t\to t+1}$, and the future return square, $R_{t+1\to t+2}^2$, under the risk-neutral probability measure $\mathbb{Q}$. In Eq. (2.5), the risk-neutral conditional variance of $R_{t\to t+1}$ is computed from Eqs. (2.2) - (2.3) using index options. The other two terms are the risk-neutral conditional (upon at time t) moment, $\mathbb{E}_t^{\mathbb{Q}}[(R_{t+1\to t+2})^2]$, and the risk-neutral conditional variance of $R_{t+1\to t+2}^2$, which are discussed in the following result.

Proposition 2.2.1 *Suppose that interest rates are deterministic. For simplicity, let* $R =$

[5] For simplicity we only derive the result for $\mathbb{E}_t[R_{t+1\to t+2}]$. The expression of $\mathbb{E}_t[R_{t+n\to t+n+s}]$ is similar and given in the Appendix.

$R_{t+1\to t+2}$, $R_f = R_{f,t+1\to t+2}$. Let $F_t = FVIX_{t,t+1\to t+2}$ be the futures price of the VIX index, and σ_t be the implied volatility of at-the-money options on the VIX index. Then $\mathbb{E}_t^{\mathbb{Q}}[R^n]$ can be solved recursively for $n = 2,3,4,\cdots$ via the following approximation formulas,

$$F_t^2(1+\sigma_t^2) \sim \left(\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{R}{R_f}\right)^2\right] - 1\right), \tag{2.6}$$

$$F_t^2(1+T_1\sigma_t^2) \sim \frac{1}{T_2}\left(-\frac{2}{3}\mathbb{E}_t^{Q}\left[(\frac{R}{R_f})^3\right] + 3\mathbb{E}_t^{Q}\left[(\frac{R}{R_f})^2\right] - \frac{7}{3}\right), \tag{2.7}$$

$$\frac{1}{2}F_t^2\left(1+\sigma_t^2\right) \sim \sum_{i=1}^{n}(-1)^i\frac{1}{i}\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{R}{R_f}-1\right)^i\right], n \geq 3. \tag{2.8}$$

Proof: See Appendix □

Proposition 2.2.1 states that all risk-neutral higher moments of $R_{t+1\to t+2}$ can be computed from the publicly available VIX index, VIX futures, and VIX options.[6] For instance, the risk-neutral conditional moment of $R_{t+1\to t+2}$ is given by

$$\mathbb{E}_t^{\mathbb{Q}}[(R_{t+1\to t+2})^2] \sim R_{f,t+1\to t+2}^2\left(1+F_t^2(1+\sigma_t^2)\right). \tag{2.9}$$

This formula can be understood as follows. The CBOE's VIX index measures the risk-neutral entropy,

$$VIX_{t\to t+1}^2 = \frac{2}{T}L_t^{\mathbb{Q}}\left(\frac{R_{t\to t+1}}{R_{f,t\to t+1}}\right), \tag{2.10}$$

where $L_t^{\mathbb{Q}}(X) \equiv \log\left[\mathbb{E}_t^{\mathbb{Q}}(X)\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\log(X)\right]$.

Among quadratic polynomials, $\mathbb{E}_t^{\mathbb{Q}}[(R_{t+1\to t+2})^2]$ is the best one (up to a constant) to approximate the risk-neutral conditional entropy of $R_{t+1\to t+2}$, i,e., the second moment of a future VIX. By the equation, $\mathbb{E}^{Q}[X^2] = \mathbb{E}^{Q}[X]^2 + Var^{Q}(X)$, the second moment of a future VIX (represented by a random variable X) is the sum of a square of a *VIX futures* price, $\mathbb{E}^{Q}[X]$, and a conditional risk-neutral variance of a future VIX. Since the latter can be proxied by the square of the implied volatility of *VIX options*, we use VIX-derivates

[6] Although Proposition 2.2.1 is given as an approximation, we show that the approximation error is very small for the empirical application in the Appendix.

to compute the risk-neutral conditional second moment of a future return. Similarly, by high-degree polynomial best approximation, we obtain other equations (recursive formula) in Proposition 2.2.1. Then we can calculate $\mathbb{E}_t^{\mathbb{Q}}[(R_{t+1\to t+2})^3], \mathbb{E}_t^{\mathbb{Q}}[(R_{t+1\to t+2})^4]$, and so on. Accordingly, we calculate $Var_t^Q(R_{t+1\to t+2}^2)$ using VIX-derivatives.

Remark 2.2.1 *Following Martin (2013), the risk-neutral conditional cumulant-generating function $K(\lambda)$ of the relative future return $\frac{R}{R_f}$ is $K(\lambda) = \log\left(\mathbb{E}_t^{\mathbb{Q}}\left[e^{\lambda R/R_f}\right]\right)$. Notice that $\mathbb{E}_t^{\mathbb{Q}}\left[\frac{R}{R_f}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{E}_{t+T}^{Q}\left(\frac{R}{R_f}\right)\right] = 1$, then*

$$K(\lambda) = \log\left(1 + \lambda + \sum_{n=2}^{\infty} \frac{1}{n!}\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{R}{R_f}\right)^n\right]\lambda^n\right).$$

By Proposition 2.2.1, the function $K(\lambda)$ can be computed from VIX-derivatives.

Now, in computing the forward return in Equation (2.5), the last ingredient is θ_t. Theoretically, θ_t can be obtained from the risk-neutral bivariate distribution of $(R_{t\to t+1}, R_{t+1\to t+2})$. Thus basket or correlation options can be used to recover the value of this parameter.[7] Nevertheless, there are no available real-time basket or correlation options yet in the financial market to derive the value of θ_t.

In this paper, we suggest two methods to estimate the value of θ_t. The first method is as follows. While θ_t is unknown, we know that the risk-neutral correlation coefficient, $corr_t^{\mathbb{Q}}\left(R_{t\to t+1}^2, R_{t+1\to t+2}\right) = 0$, and in general, $corr_t^{\mathbb{Q}}\left(g(R_{t\to t+1}), R_{t+1\to t+2}\right) = 0$, for any function $g(\cdot)$. This implies that, *to a certain extent*, $R_{t+1\to t+2}$ is independent from $R_{t\to t+1}$. Therefore, it is reasonable to expect that θ_t is close to zero.[8] Indeed, in Section 2.4 below, we introduce an alternative expression of forward returns in terms of index options and

[7]For this particular case, Martin (2021) reduces it to be a forward-start option valuation. In general, if $\mathscr{F}_{t+T}$ is generated by $S_1, \cdots, S_{t+T}$, then all conditional information at time t should be recovered by the time t value of some options such as basket options with state variables $S_{t+1}, \cdots, S_{t+T}$. The theory is developed in Tian (2014, 2019) by using the universal approximation theorem from neural networks. Carr and Laurence (2011) also develop a theory in terms of basket options based on random transformation.

[8]For instance, if Stein's lemma holds for the risk-neutral bivariate distribution of $(R_{t\to t+1}, R_{t+1\to t+2})$, then $\theta_t = 0$. In general, however, the bivariate distribution of (X, Y) has a rich structure, yielding non zero correlation coefficient between X and Y^2, but zero-correlation between X^2 and Y. We will discuss this issue in Appendix C.

option greeks. As will be shown in Section 2.4, we can use historical option data to estimate θ_t (the second method), and the averages are fairly close to zero. For these reasons, we first assume that $\theta_t = 0$; thus, the forward return is

$$\mathbb{E}_t[R_{t+1\to t+2}] = \frac{1}{R_{f,t+1\to t+2}}\mathbb{E}_t^Q[R^2_{t+1\to t+2}] \sim R_{f,t+1\to t+2}\left(1+F_t^2(1+\sigma_t^2)\right). \tag{2.11}$$

Since VIX-derivatives data alone are sufficient to compute the forward return in the last equation, we call this approach the *VIX-approach* for estimating the forward return.

Remark 2.2.2 *Similar to Equation (2.3), in which the conditional moments of spot return can be implied by index options, we can also express the higher moments of future return with VIX-derivatives. Precisely, by assuming* $corr_t^Q\left(R_{t\to t+1}, (R_{t+1\to t+2})^k\right) = 0, k \geq 2$, *we have*

$$\mathbb{E}_t\left[(R_{t+1\to t+2})^k\right] = \frac{1}{R_{f,t+1\to t+2}}\mathbb{E}_t^Q\left[(R_{t+1\to t+2})^{k+1}\right]. \tag{2.12}$$

Thus far, we assume a log-utility-based specification of the pricing kernel process. It can easily be extended to a power specification for a representative CRRA-type agent with a coefficient of constant relative risk aversion γ,

$$m_{t,t+1} = \left(\frac{S_t}{S_{t+1}}\right)^{\gamma}, \quad \gamma \geq 1.$$

2.3 Empirical Results and Applications

In this section, we use the theoretical results of Section 2.2 in three novel applications. We start with the first application by applying the VIX-approach to derive the term structure of forward one-month returns. Next, we compute the expected spot return (and higher orders) from index options. Combined with the moments of future return from VIX-derivatives, we recover the market autocorrelation on a real-time basis. Finally, we show how the real-time market autocorrelation can be used to construct a market timing strategy that outperforms the buy and hold benchmark. A common theme of these applications is that the recovered future return and autocorrelation contain valuable forward-looking information not captured by historical measures.

2.3.1 Data

We collect data for S&P 500 index options and VIX options from OptionMetrics, and obtain VIX index and VIX futures data from the CBOE. On each trading day, we follow Hu and Jacobs (2020) to use linear interpolation to compute daily VIX futures prices, $\text{FVIX}_{t,t+T_1\to t+T_1+T_2}$, with constant maturities for T_1 = 1, 2, 3, 4, 6, and 9 months.[9] Since both VIX index and VIX futures measure the forward-looking implied index volatility over 30 days, T_2 always represents one month.

Panels A and B in Table 2.1 report summary statistics for VIX index and VIX futures prices. Typically, the VIX futures market is in contango. That is, on average, VIX futures prices are higher than the VIX index, reflecting the volatility risk premium paid by holders of long volatility positions. Panel C reports summary statistics for implied volatility of at-the-money VIX call and put options. After applying standard filters and merging data from different databases, we end up with a sample of daily observations from February 24, 2006 to December 31, 2019. All results are annualized.

2.3.2 Term structure of forward returns

The expected excess return on the market, or expected equity risk premium, is one of the central quantities of interest in finance and macroeconomics (Martin, 2017; Rapach and Zhou, 2022). In this subsection, we study the shapes of the term structure of forward returns and equity forward risk premiums, as the first application.

[9] CFE may list futures for up to nine near-term serial months, as well as five months on the February quarterly cycle associated with the March quarterly cycle for options on S&P 500 (Mencia and Sentana, 2013). We thus choose the maximum constant maturity to be nine months. VIX futures expiration calendar can be found at `https://www.macroption.com/vix-expiration-calendar/`.

To be specific, let

$$f_{t,T} = \mathbb{E}_t\left[R_{t+T\to t+T+1}\right], \forall t, \forall T = 0, 1, \cdots, \tag{2.13}$$

where $f_t \equiv f_{t,0}$ is the conditional expected spot return $\mathbb{E}_t\left[R_{t\to t+1}\right]$. The one-period forward return $f_{t,T}$ forms a term structure of future returns analogous to the term structure of forward rates: Conditional on the time-t, $f_{t,T} = \mathbb{E}_t\left[\mathbb{E}_{t+T}[R_{t+T\to t+T+1}]\right] = \mathbb{E}_t[f_{t+T}]$, which is similar to the equation that the implied forward rate equals the expected spot rate in the fixed-income market.

However, the relationship between equity spot and forward returns is fundamentally different than the relationship between bond spot and forward returns. To see this, suppose $\hat{r}_{t\to t+n}$ is a default-free continuously compounded spot interest rate in effect from time t until the future time $t+n$.[10] In other words, $R_{f,t\to t+n} = \exp(n\hat{r}_{t\to t+n}) = 1/P(t,t+n)$, where $P(t,t+n)$ is the time t price of a default-free zero coupon bond paying 1 at time $t+n$. By the (continuously compounded) yield curve at time t, we mean the mapping $n \to \hat{r}_{t\to t+n}$. At time t, the return that will be realized from holding the zero-coupon bond until maturity at time $t+n$ is *known*. On the other hand, we can also construct a term structure of equity spot returns, say, $n \to \mathbb{E}_t[R_{t\to t+n}]$. However, the realized return $R_{t\to t+n}$ on the equity index is *unknown* at time t, as the expected spot return is distinct from the realized return due to the perpetual nature of equity claims.

Now let $\hat{f}_{t,t+n\to t+n+m}$ be the implied (continuously compounded) forward rate, at time t, from time $t+n$ to $t+n+m$. The well-known relationship between forward and spot rates,

$$\hat{f}_{t,t+n\to t+n+m} = \frac{(m+n)\hat{r}_{t\to t+n+m} - n\hat{r}_{t\to t+n}}{m}, \tag{2.14}$$

is a straightforward consequence of the no-arbitrage principle. However, the forward equity return, $\mathbb{E}_t[R_{t+n\to t+n+m}]$, cannot be derived similarly in terms of expected spot returns. In

[10] A similar argument also holds for any discrete compounding convention.

fact, we can determine the time-t fair rate, K, of a forward contract maturing at time $t+n$, to exchange the realized return $R_{t+n\to t+n+m}$ at time $t+n+m$. By the no-arbitrage asset pricing principle,

$$K = \mathbb{E}_t^Q[R_{t+n\to t+n+m}] = \mathbb{E}_t^Q\left[\mathbb{E}_{t+n}^Q[R_{t+n\to t+n+m}]\right] = \mathbb{E}_t^Q[R_{f,t+n\to t+n+m}]. \tag{2.15}$$

Hence, the “implied" forward return on the equity index is the risk-free forward return.

The difference between $\mathbb{E}_t[R_{t+n\to t+n+m}]$ and $\mathbb{E}_t^Q[R_{t+n\to t+n+m}]$ is called the time-t conditional expected forward risk premium. Compared to the equity index's forward return, the risk-free return in each short time period is small and relatively stable. Therefore, the term structure of $f_{t,T}$ is essentially comparable (in shape) to the term structure of *expected forward risk premiums*, $f_{t,T} - \mathbb{E}_t^Q[R_{f,t+T\to t+T+1}]$, which reduces to $f_{t,T} - R_{f,t+T\to t+T+1}$, provided interest rates are deterministic. Thus, we obtain the term structure of forward one-period returns and expected forward risk premiums by the VIX-approach in Equation (2.11) and Proposition 2.2.1.

Table 2.2 reports the summary statistics for the T-forward one-month returns. We choose T to be 1, 2, 3, 4, 6, and 9 months in $f_{t,T}$. Panel A considers the full sample period from February 24, 2006 to December 31, 2019. The term structure of forward one-month equity returns (and equity risk premiums) is mainly upward-slopping in normal times. On average, the forward one-month return and expected risk premium *increase* with respect to maturity T, except for a slightly downward/flat feature when $T = 6$ months.

We next examine the shape of the term structure $f_{t,T}$, for $T = 1,2,3,4,6,9$, when the market is in bad (good) times. Precisely, we use the NBER recessions period, January 1, 2008−June 30, 2009 to represent bad times, and the post-NBER recessions, July 1, 2009−December 31, 2019, to represent good times. As shown in Panel B of Table 2.2, the term structure of forward one-month returns (and equity forward risk premiums) is *downward* sloping on average in bad times. In contrast, Panel C reveals an *upward*-sloping term structure of one-month future returns (and expected risk premiums), on average, in good times.

Collectively, Figure 2.1 illustrates the term structure of future one-month returns on average, in “good", “bad", or “overall" times, respectively. The slope of the term structure is pro-cyclical. Furthermore, Figure 2.2 displays the term structure of future one-month returns for all time t during the NBER recessions. More specifically, we divide the sample period into four shorter ones: January 2008−October 2008; November 2008−January 2009; February 2009−April 2009; and May 2009−June 2009. We observe that the downward-sloping term structure is significantly steep between October 2008 and April 2009 (the most severe period of the 2008/09 global financial crisis). By contrast, Figure 2.3 shows an upward-sloping term structure of forward one-month returns (expected risk premiums) most of the time between 2009 and 2019.

It is interesting to compare our results on the equity forward term structure with recent studies of the equity term structure in the literature (see, for instance, Binsbergen et al., 2012; Binsbergen and Koijen, 2017; Gormsen, 2021). By a term structure of equity risk premia, these previous studies refer to the relationship between a one-period spot return premia of maturing asset with the maturity.[11] Specifically, a zero-coupon equity or dividend strip is a claim with only one dividend payment at a future time, analogous to a zero-coupon bond. Let $P_{n,t}$ be the price at time t of a claim (dividend strip) to the dividend at time $t+n$. Then, the time-t price of the underlying index is $S_t = \sum_{n=1}^{\infty} P_{n,t}$. Let $R_{n,t}$ be the one-period spot return of the dividend strip maturing $t+n$ from period t to $t+1$. That is,

$$R_{n,t+1} = \frac{P_{n-1,t+1}}{P_{n,t}}. \tag{2.16}$$

[11] In Chabi-Yo and Loudis (2020), a term structure is a lower bound of hold-to-maturity expected spot returns at various horizons. The authors show that the term structure of the (lower bound) of spot returns is downward-sloping during turbulent times but upward-sloping during normal times.

The spot return of the underlying index is

$$R_{t\to t+1} = \sum_{n=1}^{\infty} \omega_{n,t} R_{n,t+1}, \omega_{n,t} = \frac{P_{n,t}}{S_t}. \tag{2.17}$$

Then, the term structure of the dividend return, $n \to R_{n,t+1}$, illuminates the contribution of the dividend return $R_{n,t+1}$ to the spot return $R_{t\to t+1}$. [12] For example, an upward-sloping term structure of the dividend return shows a higher contribution of the dividend strip maturing $t+n$ to the underlying aggregate market index return $R_{t\to t+1}$ when the maturity n is higher, and vice versa. In contrast, we focus on the term structure of the aggregate equity market's forward returns, $T \to R_{t+T\to t+T+1}$, a relationship between T and the forward return starting from period $t+T$ to $t+T+1$.

In this regard, we document a new stylized fact that the term structures of forward one-period returns and expected forward risk premiums implied by derivatives markets are *pro-cyclical*. The pro-cyclicality can be explained as follows. By Proposition 2.2.1, the conditional expected future one-month returns in T months are essentially determined by the futures prices of VIX over the same period. Therefore, a pro-cyclical term structure of equity risk premia is consistent with Hu and Jacobs (2020) which documents that VIX futures prices tend to have an upward sloping term structure during normal times and tend to become inverted or hump-shaped in times of market turbulence.

2.3.3 Market autocorrelation

Building on the theoretical results of Section 2.2, we turn now to the question of predicting the expected market return with past return observations. Specifically, we are interested in computing the conditional market autocorrelation, $corr_t(R_{t\to t+1}, R_{t+1\to t+2})$, under the real-world probability measure. This conditional market autocorrelation reveals how two future returns change from the perspective of time t in two consecutive periods. Our goal

[12]Gormsen (2021); Bansal et al. (2021) study the term structure of the dividend future return, or $\theta^{n,m} = \mathbb{E}_t[R_{n,t} - \mathbb{E}_{m,t}]$ for long maturity n and short maturity $m < n$. Similarly, Binsbergen et al. (2012); Binsbergen and Koijen (2017) consider the difference between short-term assets with all dividend payments until T, say three years, and long-term assets with all remaining future dividends.

is to compute these conditional market autocorrelations from derivatives data.

For this purpose, we compute

$$Cov_t\left(R_{t\to t+T}, R_{t+T\to t+T+1}\right) = \mathbb{E}_t\left[R_{t\to t+T+1}\right] - \mathbb{E}_t\left[R_{t\to t+T}\right] \times \mathbb{E}_t\left[R_{t+T\to t+T+1}\right]. \quad (2.18)$$

By Equations (2.3) and (2.11), the expected spot return can be recovered from equity index option prices, and the expected future return can be obtained from the VIX-derivatives. Similarly, we can estimate $Var_t(R_{t\to t+T}), Var_t(R_{t+T\to t+T+1})$, and then the autocorrelation coefficient $corr_t\left(R_{t\to t+1}, R_{t+1\to t+2}\right)$, under the real-world probability measure.

Table 2.3 (Panel A) reports the market autocorrelation on the S&P 500 index. We use the average value between the implied volatility of at-the-money put and call VIX options as a proxy for σ_t. Following the VIX futures structure, we consider T_1 to be 1, 2, 3, 4, 6 and 9 months, and T_2 to be fixed for 1 month. Across columns, we observe significantly *negative* coefficients, suggesting a persistent short-term reversal. For instance, when T_1 is 1 month, $corr_t(R_{t\to t+1mo}, R_{t+1mo\to t+2mo})$ is on average -20.90% with a *t*-stat of 18.10. On average, the market autocorrelation on S&P 500 index is around -20% to -40%. Notably, the numbers in Table 2.3 are comparable to Chabi-Yo (2019) and Martin (2021). Using index options data, Chabi-Yo (2019) estimates that the upper bounds of autocorrelation vary from -28% to -3%; with price quotes of forward-start options from a major investment bank, Martin (2021) estimates that the autocorrelation of the S&P 500 index is between -20% and -40% for a small number of days. Table 2.3 suggests that our no-arbitrage framework, in Proposition 2.2.1, is consistent with the pricing of over-the-counter derivatives in the market.

Figure 2.4 displays the time-series of $corr_t(R_{t\to t+1mo}, R_{t+1mo\to t+2mo})$. It is well known that the month-month autocorrelation coefficient from the historical data is close to zero ((Lo and MacKinlay, 1988, 1990), and Table 2.3, Panel B). From the perspective of the derivatives market, however, the autocorrelation coefficients can be either negative or positive, though they are negative most of the time.

As a comparison, we also compute the *month-to-month autocorrelation* between two consecutive calendar months using historical stock return, including January/February, February/March, $\cdots$, and December/January. Figure 2.5 plots the consecutive month-to-month autocorrelation over various periods with the data prior to January 1927 obtained from Robert Shiller's website. In contrast to virtually zero month-month autocorrelation coefficient in (Lo and MacKinlay, 1988, 1990), the autocorrelation between two consecutive months can be significantly nonzero. It can be either positive or negative, depending on the sample of the data. For example, the autocorrelation of March/April is around 10% over 1871-2019, but -20% over a recent time period 1990-2019.

We next compute the consecutive month-to-month autocorrelation from the derivative market as explained in Section 2.2. For consistency, we restrict the sample period to $2006-2019$, calculate $corr_t(R_{t\to t+1mo}, R_{t+1mo\to t+2mo})$ on the first day of each month and take the simple average within each of the 12 calendar months of the year. For example, for March/April, we compute the correlation coefficient with VIX-derivatives data on March 1 in each year and then take a simple average. Our results are displayed in Figure 2.6, in which the solid red line displays the month-to-month autocorrelations from the derivative market. By contrast, the blue dot line represents the month-month autocorrelation from the historical stock return (as in Figure 2.5). Both methods yield a similar pattern of consecutive month-month autocorrelation, but the derivative approach results more negatively in magnitude. In fact, we demonstrate negative autocorrelation between any two consecutive months from the derivatives market. As an illustration, both yield similar autocorrelation coefficients of -35% between February and March, and -10% between December and January. Between May and June, the derivative approach implies an autocorrelation as large as -40%, while the historical returns suggest a value of -20%. Moreover, the monthly return displays a stronger reversal in specific periods than others (for instance, from February to March, May to June, July to August, and December to January) by both approaches. However, Figure 2.5 also displays a moderate reversal between two consecutive calendar months in specific periods. For instance, the corresponding autocorrelations are significantly negative from February to March, March to April, May to June, November to December, and December to January. In summary, the derivative market reveals a robust short-term reversal in the stock market from a forward-looking perspective.

Finally, using the VIX approach, we can derive the conditional correlation coefficient between two spot returns. For instance, in the following conditional correlation between $R_{t\to t+1}$ and $R_{t\to t+2}$,

$$corr_t(R_{t\to t+1}, R_{t\to t+2}) = \frac{\mathbb{E}_t[R_{t\to t+1}R_{t\to t+2}] - \mathbb{E}_t[R_{t\to t+1}]\mathbb{E}_t[R_{t\to t+2}]}{\sqrt{Var_t(R_{t\to t+1})}\sqrt{Var_t(R_{t\to t+2})}}, \tag{2.19}$$

every term except $\mathbb{E}_t[R_{t\to t+1}R_{t\to t+2}]$ is obtained from index options. Similar to Equation (2.5), we have

$$\begin{aligned}\mathbb{E}_t[R_{t\to t+1}R_{t\to t+2}] &= \frac{1}{R_{f,t\to t+2}}\mathbb{E}_t^{\mathbb{Q}}\left[(R_{t+1\to t+2})^2 \times (R_{t\to t+1})^3\right], \\ &= \frac{1}{R_{f,t\to t+2}}\left\{\mathbb{E}_t^{\mathbb{Q}}\left[(R_{t+1\to t+2})^2\right] \times \mathbb{E}_t^{Q}\left[(R_{t\to t+1})^3\right] + Cov_t^{Q}\left((R_{t+1\to t+2})^2, (R_{t\to t+1})^3\right)\right\} \\ &\sim \frac{1}{R_{f,t\to t+2}}\mathbb{E}_t^{\mathbb{Q}}\left[(R_{t+1\to t+2})^2\right] \times \mathbb{E}_t^{Q}\left[(R_{t\to t+1})^3\right],\end{aligned}$$

assuming $corr_t^{Q}\left((R_{t+1\to t+2})^2, (R_{t\to t+1})^3\right) = 0$.[13] Proposition 2.2.1 and Equation (2.3) can be used to derive $\mathbb{E}_t[R_{t\to t+1}R_{t\to t+2}]$ with index options and VIX derivatives.

Panels C and D in Table 2.3 report this new correlation coefficient calculated using either derivatives or historical stock returns. As shown, using historical data, the autocorrelation between two spot returns is significantly positive. For example, on average, the correlation coefficient between the one-month spot return and the two-month spot return is 0.746.

[13] By a similar method in Section 2.4, we can also justify this assumption empirically. Without the VIX approach, we need prices of some generalized forward-start options or spread options which are at present traded only in over-the-counter markets.

However, based on derivatives market information, the one-month spot return and two-month spot return are more significantly positively correlated (0.83). Furthermore, the market autocorrelation between two spot returns, for $T_1 = 1,2,3,4,6,9$ months, using derivatives market information is more significant than the correlation coefficients derived from the historical stock returns data. Given the robust short-term reversal from the derivatives (Panel A), we document that the derivatives market information reveals a very substantial and higher correlated movement between two spot returns.

2.3.4 Market timing

To demonstrate the investment value of the VIX-approach, we next propose a market timing strategy as the third application. We start by constructing a reversal trading signal based on the forward-looking autocorrelation from the derivatives. We then discuss several evaluation criteria and report the out-of-sample performance of the marketing timing strategy.

2.3.4.1 A reversal trading signal

Motivated by the short-term reversal documented in Table 2.3, we construct a reversal signal based on both realized cumulative excess returns and the conditional derivative-based autocorrelation. Specifically, we define the reversal signal at time t as, the realized cumulative excess return over the past K months, $r_{t-K\to t}$, and the corresponding conditional autocrrelation $corr_{t-K}(r_{t-K\to t}, r_{t\to t+1})$ by Q-approach computed at time $t-K$ as follows.

$$\tilde{S}_{t,K}\left[r_{t-K\to t}, corr_{t-K}(r_{t-K\to t}, r_{t\to t+1})\right] = \begin{cases} 1, & \text{if} \quad r_{t-K\to t} > 0 \quad \& \quad corr_{t-K}(r_{t-K\to t}, r_{t\to t+1}) > 0, \\ 1, & \text{if} \quad r_{t-K\to t} < 0 \quad \& \quad corr_{t-K}(r_{t-K\to t}, r_{t\to t+1}) < 0, \\ 0, & \text{otherwise,} \end{cases} \tag{2.20}$$

where $r_{t-K\to t} = R_{t-K\to t} - R_{f,t-K\to t}$ is the realized cumulative excess return over the past K months, $corr_{t-K}(r_{t-K\to t}, r_{t\to t+1})$ is calculated from the derivatives, and $K = 1$, 2, 3, 4, 6, and 9 months. In total, we have six market reversal signals at time t.[14]

Following the market reversal signal in 2.20, we trade the market by implementing a zero-cost strategy at the beginning of the subsequent month. As an illustration, if we use the one-month reversal signal as a trading signal at time t and implement the corresponding market timing strategy, the realized return in the subsequent month is

$$\eta\left[\tilde{S}_{t,1}\right] = \begin{cases} r_{t\to t+1}, & \text{if} \quad \tilde{S}_{t,1} = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{2.21}$$

We call the strategy based on the 1-month reversal signal the *single timing strategy*. It is also possible to use all six reversal signals together, which we call *combination timing strategy*. For instance, employing the combination strategy would entail being long the market if $\sum_K \tilde{S}_{t,K}$ is greater than some threshold, ξ, an integer ranging from 2 to 5. Following the combination timing strategy, the realized return in the next month is

$$\eta\left[\tilde{S}_{t,K}, \forall K; \xi\right] = \begin{cases} r_{t\to t+1}, & \text{if} \quad \sum_K \tilde{S}_{t,K} \geq \xi, \\ 0, & \text{otherwise.} \end{cases} \tag{2.22}$$

In other words, we should be long the market if and only if at least ξ reversal signals defined in Equation (2.20) indicate long signals.

2.3.4.2 Performance evaluation

To evaluate the market timing strategy's performance, we compute four performance measures based on the mean $\hat{\mu}_j$ and standard deviation $\hat{\sigma}_j$ of the out-of-sample realized returns of strategy j.

First, we measure the *out-of-sample Sharpe ratio* (SRatio) and the *certainty-equivalent*

[14] Notice that computing the autocovariance is sufficient in constructing the reversal signal. We do not necessarily need the autocorrelation in this case.

return (CEQ) of each strategy,

$$\hat{s}_j = \frac{\hat{\mu}_j}{\hat{\sigma}_j} \tag{2.23}$$

and

$$C\hat{E}Q_j = \hat{\mu}_j - \frac{\gamma}{2}\hat{\sigma}_j^2, \tag{2.24}$$

where γ is chosen to be 1, consistent with the log-utility specification of the stochastic discount factor in Section 2.2.

Next, we compute DeMiguel et al. (2009) *return-loss* with respect to a benchmark. We choose the buy and hold strategy as the benchmark as in Gao et al. (2018). Precisely, if $\{\hat{\mu}_b, \hat{\sigma}_b\}$ are the monthly out-of-sample mean and volatility of the excess returns from the buy-and-hold strategy, the return-loss from strategy j is

$$\text{return-loss}_j = \left(\frac{\hat{\mu}_b}{\hat{\sigma}_b}\right) \times \hat{\sigma}_j - \hat{\mu}_j. \tag{2.25}$$

In other words, the return-loss is the additional return required in order for the performance of strategy j to be consistent with the performance of the benchmark. A negative value indicates that strategy j outperforms the benchmark as measured by the Sharpe ratio.

Lastly, we calculate the *performance fee* suggested in Fleming et al. (2001), defined as the maximum fee that a quadratic-utility investor would be willing to pay to switch from the benchmark to the timing strategy. This fee is estimated as the value of Δ that solves

$$\sum_t \left[\left(R_{j,t} - \Delta\right) - \frac{\gamma}{2(1+\gamma)} \left(R_{j,t} - \Delta\right)^2 \right] = \sum_t \left[R_{b,t} - \frac{\gamma}{2(1+\gamma)} R_{b,t}^2 \right], \tag{2.26}$$

where $R_{j,t}$ and $R_{b,t}$ denote the out-of-sample realized returns for timing strategy j and the benchmark, respectively. We report the estimates of Δ in units of basis points per annum.

2.3.4.3 Out-of-sample performance

Panel A of Table 2.4 reports the performance measured on (annualized) returns generated from the market timing strategies over the full sample period. The market timing strategy delivers good realized returns, but does not necessarily lead to the highest realized return on average. The average realized return is about 4.433% *per annum* by the single timing strategy, whereas the average return from buy and hold is about 5.489%. This is reasonable given the market's upward trend from 2006 to 2019, regardless of the financial crisis around 2008 or a market downturn in 2018.

More important questions to ask are: (i) whether the market timing strategy delivers a higher Sharpe ratio; and (ii) whether it "predicts" bad market times.

For the first question, all timing strategies produce minor standard deviations than the benchmark, suggesting that the market timing strategy is more conservative than the benchmark. For instance, the standard deviation is 9.616% per annum for the single timing strategy; but 14.789% for buy and hold, which is almost twice large. As a result, the single timing strategy produces a Sharpe ratio of 0.461, whereas the buy and hold only achieves 0.371. We also see that the last combination timing strategy delivers a higher Sharpe ratio of 0.467.

For the second question, we evaluate the out-of-sample performance during the NBER recessions in Panel B of Table 2.4. Not surprisingly, the buy and hold strategy suffers a dramatic loss, yielding a negative average return of −32.304% per annum, along with a standard deviation as high as 25.565%, during January 2008−June 2009. Consequently, the Sharpe ratio of the benchmark is around −1.264. In contrast, the single timing strategy, $\eta\left[\tilde{S}_{t,1}\right]$, achieves an average return of −7.027% per annum, along with a much smaller standard deviation of 14.420%. Although the Sharpe ratio from the single timing strategy is also negative, around −0.487, it exhibits a significant economic value relative to the benchmark, as suggested by the return-loss and the performance fee. The −11.194% return-loss value of the single timing strategy suggests that investors are willing to pay as high as 11% per annum to switch from the buy and hold to the market timing strategy. Likewise, the quadratic-utility investor would be willing to pay an estimated 2630 basis points annually to switch from the benchmark portfolio to the single timing strategy. Remarkably and consistently, during the market crisis, all single and combination timing strategies yield higher average returns, smaller standard deviations, higher Sharpe ratios, larger CEQs, negative return-loss measures, and positive performance fees than the buy and hold strategy.

Furthermore, we plot the realized returns generated from the single market timing strategy and the buy and hold strategy during the NBER recessions from January 2008 to June 2009 in Figure 2.7. This recession period overlaps the 2008/09 global financial crisis. We observe that the market timing strategy based on a one-month reversal signal avoids significant market crashes in January, June, September, October of 2008, and January of 2009. To summarize, we show that the robust short-term reversal identified by the derivative market does reveal valuable information on future market downturns, and the associated economic value can be substantial.

2.4 VIX-Gamma approach

So far, we have assumed that $\theta_t = corr_t^{\mathbb{Q}}\left(R_{t\to t+1}, R^2_{t+1\to t+2}\right) = 0$. One appealing feature of the VIX-approach is that it provides an effective way to compute forward return in real-time.

However, we need to justify this assumption for the equity index, as this zero risk-neutral correlation coefficient assumption fails for a general risk-neutral bivariate distribution in a no-arbitrage pricing model.[15] Consequently, we turn our attention now to the problem of estimation and prediction of θ_t using available derivatives data.

In this section, we first provide an alternative expression of forward return in terms of option prices and gammas. Although this formula cannot be used in real-time to compute forward return (see explanations below), we can combine this expression and Equation (2.5) to construct a predictor for the parameter θ_t using available historical index options and VIX-derivatives. Then, we use this estimation of θ_t at time t to compute the forward return. We call this methodology, the VIX-Gamma approach, and apply the VIX-Gamma

[15] We provide a simple example in **??** to illustrate that the risk-neutral correlation coefficient can be any nonzero number between -1 and 1 in a simple two-period no-arbitrage asset pricing model.

approach to the market timing strategy on a real-time basis.

2.4.1 Forward return from index option market

In this subsection, we provide an alternative expression of forward return from the options market.

Proposition 2.4.1 *Suppose that interest rates are deterministic. Then,*

$$\mathbb{E}_t[R_{t+1\to t+2}] = 2S_t \int_0^\infty \frac{C_t''(S_t,L)}{L^3} \left(\underbrace{\int_0^\infty C_{t+1}(L,K)\,dK}_{\textit{inside-integral, } I_{t+1}(L)\textit{, known at } t} \right) dL, \tag{2.27}$$

where

- $S_t =$ *underlying index price observed at time t;*
- $C_t''(S,L) =$ *the call option gamma at time t with the underlying S_t and strike price L;*
- $C_{t+1}(L,K) =$ *the call option price at time $t+1$ with the underlying L and strike price K.*

Proof: See Appendix □

Compared with Eq. (2.11) for the equity index, Equation (2.27) presents an alternative formula of forward return for a general underlying variable R_t. Specifically, to compute a forward return at time t, there are two sets of option data required in Equation (2.27) in addition to the asset price S_t. First, the call option gamma, $C_t''(S,L)$, with the underlying S_t, strike price L, and maturity $t+1$, is needed. The option gamma is available in real-time. Second, the time $t+1$ price of call option price with time to maturity $t+2$, *when the underlying price is L at time* $t+1$, from time t perspective. Notice that this price $C_{t+1}(L,K)$ is known at time t, but it is not real-time, since the underlying index only achieves one particular number at time $t+1$. Therefore, we need to explain why the option gamma and $C_{t+1}(L,K)$ are involved in this equation (the details are given in the Appendix).

First of all, the number $C_{t+1}(L,K)$ involved in Equation (2.27) is well-defined at time t. As a simple illustrative example, the Black Scholes option formula, assuming the underlying asset has a normal distribution with a constant volatility parameter σ, presents a known option price at a future time as follows:

$$C_{t+1}(L,K) = LN(d_1) - \frac{K}{R_{f,t+1\to t+2}} N(d_1 - \sigma\sqrt{\Delta t}), d_1 = \frac{log(L/K) + (r_t + \frac{1}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}}.$$

When we write $C_{t+1}(S_{t+1},K)$ at time t, the reason we do not know the option price precisely is that we do not know the realized value S_{t+1}. However, the no-arbitrage asset pricing theory guarantees the **relationship** between $C_{t+1}(S_{t+1},K)$ and S_{t+1}. In particular, when S_{t+1} achieves a number L, the price $C_{t+1}(L,K)$ is known. This argument holds in general, regardless of the distribution of S_{t+1}. The reason is simple. Given the specification of the stochastic discount factor, we know at time t a precise relationship between the underling index price S_{t+1} and the option price $C_{t+1}(S_{t+1},K)$ for any conditional distribution of S_{t+1}. Hence, $C_{t+1}(L,K)$ is well-defined at time t.

Second, even though $C_{t+1}(L,K)$ is known at time t, its expression could be complicated. Under the power-specification of the stochastic discount factor, we obtain

$$\mathbb{E}_t[R_{t+1\to t+2}] = \frac{2}{R_{f,t\to t+1}S_t} \int_0^\infty \mathbb{E}_t^{\mathbb{Q}} \left[\frac{C_{t+1}(S_{t+1},K)}{S_{t+1}} \right] dK, \tag{2.28}$$

which resembles a similar insight in Equation (2.3) for a future return (a proof is given in the Appendix). If the specification of the stochastic discount factor contains a volatility component, as discussed in Bakshi et al. (2021); Babaoğlu et al. (2018); Christoffersen et al. (2013), the above expression of the forward return would be different, since $C_{t+1}(L,K)$ would also depend on the volatility at time t. But, such a specification involving a volatility component is beyond the scope of the present paper.

Third, why do we need option gammas in Equation (2.27) for a future return, whereas only option prices are required to compute expected spot return $\mathbb{E}_t[R_{t\to t+1}]$ in Equation

(2.3)? This difference seems substantial since $R_{t+1\to t+2}$ is just the index return in a future time period $[t+1,t+2]$. We notice that the term inside the integral is time t price (ignoring the interest rate) of a future payoff $\frac{C_{t+1}(S_{t+1},K)}{S_{t+1}}$ at time $t+1$; thus, it is essentially a forward-start option price with payoff $\frac{\max(S_{t+2}-K,0)}{S_{t+1}}$ at time $t+2$. We write $\Pi(L|S_t)$ be the conditional distribution of S_{t+1}. Namely,

$$\Pi(L|S_t) = \int_0^L q(z|S_t)dz$$

where $q(z|S_t)$ is the conditional density function under the risk-neutral probability measure. It is known in the options literature that

$$d\Pi(L|S_t) = R_{f,t\to t+1}\frac{\partial^2 C_t(S_t,L)}{\partial L^2}dL$$

Therefore, we can represent the forward-start option at time t as

$$\mathbb{E}_t^{\mathbb{Q}}\left[\frac{C_{t+1}(S_{t+1},K)}{S_{t+1}}\right] = \int_0^\infty \frac{C_{t+1}(L,K)}{L}R_{f,t\to t+1}\frac{\partial^2 C_t(S_t,L)}{\partial L^2}dL.$$

Finally, Equation (2.27) follows from the following relationship between option gamma and strike-gamma as follows.

$$L^2\frac{\partial^2 C_t(S_t,L)}{\partial L^2} = S_t^2\frac{\partial^2 C_t(S_t,L)}{\partial S_t^2},$$

in which we can use the option gamma to replace the strike gamma up to a constant.

Remark 2.4.1 *Similarly, we can derive* $\mathbb{E}_t\left[R^k_{t+1\to t+2}\right], k\geq 2$ *as follows.*

$$\mathbb{E}_t\left[R^k_{t+1\to t+2}\right] = (k+1)kS_t\int_0^\infty \frac{C_t''(S_t,L)}{L^{k+2}}\left(\underbrace{\int_0^\infty K^{k-1}C_{t+1}(L,K)\,dK}\right)dL. \qquad (2.29)$$

Therefore, we can obtain the conditional distribution of a future return $R_{t+1\to t+2}$ *in terms* *of index option prices and index gamma.*

It is worth mentioning that Proposition 2.4.1 cannot be used directly since $C_{t+1}(L,K)$ cannot be calculated *precisely*, at time t, since we do not specify the conditional distribution of S_{t+1}. For this reason, there is no way to find such data at time t to derive the forward return in real-time.

This limitation of Proposition 2.4.1 prevents us from deriving the forward return from the option market, compared with the real-time VIX-approach. In the next section, we explain how to combine both Proposition 2.4.1 and Eq. (2.11) to introduce an improved VIX-Gamma approach for the forward return.

2.4.2 VIX-Gamma approach

Suppose our objective is to estimate the number θ_t at time t. One procedure is as follows.

First, at time $t-1$, we calculate all risk-neutral return quantities on the right-hand side of Equation (2.5) by Proposition 2.2.1, except for θ_{t-1}. Second, assuming the index price S_t is realized at time t. By the homogeneous property of the option price, $C_t(L,K) = C_t(S_t, S_tK/L)\frac{L}{S_t}$, we are able to calculate $C_t(L,K)$ at time t for any L. Therefore, at time $t-1$, we compute the left-hand side, $\mathbb{E}_{t-1}[R_{t\to t+1}]$, of Equation (2.5) directly by Proposition 2.4.1. By equating the left-hand and right-hand sides calculations, we calculate the value of θ_{t-1}. Figure 2.8 provides a visual illustration on how to understand Equation (2.27).

Finally, at time t, we compute the forward return by,

$$\mathbb{E}_t[R_{t+1\to t+2}] = \frac{1}{R_{f,t+1\to t+2}}\mathbb{E}_t^Q[R_{t+1\to t+2}] + \frac{\hat{\theta}_t}{R_{f,t\to t+2}}\sqrt{Var_t^Q(R_{t\to t+1})}\sqrt{Var_t^Q(R_{t+1\to t+2}^2)}, \tag{2.30}$$

where $\hat{\theta}_t = \theta_{t-1}$,

In the above procedure, we use the risk-neutral correlation coefficient θ_{t-1} as a predictor of θ_t. The reason is straightforward. From an econometrics perspective, we can investigate the time-series property of $\theta_s, s < t$, at time t. Then, we can use statistics of this time-series $\{\theta_s, s < t\}$ to predict θ_t by an econometrics study. For example, if this time series is stationary, θ_{t-1} is a good indicator of θ_t. More generally, if this time series is ergodic, the sample average of $\theta_s, s < t$ is a good indicator of θ_t. A VIX-Gamma approach is to use a predictor $\hat{\theta}_t$ from the statistics of time-series of all past θ's in Equation (2.30).

Table 5 reports some statistics of the time series of θ_t that is calculated with available derivative data. We find out the sample average of θ_t is relatively stable across different periods and close to zero. For instance, the average value of θ_t in different periods varies between -0.092 to 0.049 except for the calendar year 2010−2011. Moreover, the moving average of θ_t from 6 months to 5 years belongs to [-0.046, 0.013]. And for the entire period, the moving average of θ_t is about -0.06. Therefore, we have justified the VIX-approach by assuming a zero value of θ_t for the equity index.

The difference between the VIX-Gamma approach, Equations (2.30), and the VIX-approach is a non-zero predictor $\hat{\theta}_t$, which relies on all index (prices and gammas) up to (and include) time t, and VIX derivatives data prior to time t. In the end, we make use of all available VIX derivative data. all option price and option gamma data until to time t, to calculate the forward return $\mathbb{E}_t[R_{t+1\to t+2}]$. Since the VIX-Gamma approach relies on all historical and current derivatives data, it provides a *forward-backward* perspective of future returns. In contrast, the VIX approach is *forward-looking* because current VX-derivatives data are required.

2.4.3 Market timing by VIX-Gamma

In this subsection, we implement the same market timing strategy with the VIX-Gamma approach. Same as before, we compute the market autocovariance (autocorrelation) on the market and construct the market timing strategy. For brevity, we simply choose the single timing strategy where we rely on the 1-month reversal signal only. We use the θ_t that are predicted recursively as explained in the last subsection.

Table 2.5 reports several key out-of-sample performance measures of VIX-Gamma approach and the buy and hold benchmark. Same as in Table 2.4, we consider both full sample period and the NBER recession subperiod during January 2008−June 2009. We observe that, in the full sample, the VIX-Gamma approach outperforms the buy and hold with a higher average return, smaller standard deviation. Remarkably, the Sharpe ratio increases by 35%, from 0.371 by buy and hold to 0.500 by VIX-Gamma. Moreover, the Sharpe ratio difference of 0.129 is statistically significant. Likewise, the CEQ difference is also significantly positive at the 5% level. The negative return-loss and positive performance fee both suggest that the market timing strategy improves significantly by considering much more options data in VIX-Gamma approach.

A more striking result is given in Panel B of Table 2.5, during the NBER recessions period. In comparison to the negative mean returns and Sharpe ratios from buy and hold strategy, and those from the VIX-approach based timing strategy in Panel B of Table 2.4, here we observe a positive average return, a positive Sharpe ratio, and a positive CEQ. Compared with the return-loss of -11.194% and the performance fee of 2,630 basis points for the VIX-approach based single timing strategy in Panel B of Table 2.4, the two quantities jump to -31.153% and 3,948 basis points, separately, once we switch to the VIX-Gamma approach. On the whole, those results highlight the improved forecasting gains associated with the VIX-Gamma approach, and justify the investment value of studying the conditional expected returns from the derivatives.

The difference between our market timing strategy and the benchmark strategy is that we long the market only when the signal shows a positive market excess return in the following month. In contrast, the benchmark strategy is long the market persistently. In other words, our market timing strategy is to stay away from the stock market if the signal from the derivative market suggests a future market downturn. Therefore, the relative performance of the market timing strategy mainly depends on whether the reversal signal identified indeed reveals valuable information about the market return in the following month.

As shown in Figure 2.9, the reversal signal from the VIX-Gamma approach more accurately *predicts* the market downturn than the VIX-approach in Figure 2.7. Remarkably, we find that the trading signals from the VIX-Gamma approach successfully predicted all market crashes during the 2008/09 global financial crisis, except for the most severe one in October 2008. Remarkably, it also captures the upside potentials, for instance, in April and May of 2009, which seems missing in the VIX-approach.

Finally, we plot the realized returns during in full sample period in Figure 2.10. To highlight the predictive power of the timing strategy, we shadow the area below zero. The reversal signal from VIX-Gamma does seem to predict the market downturn, particularly when the market crashed in 2008−2009, 2014, 2015, and 2018−2019.

2.5 Conclusion

In this paper, we express the equity index's forward return by market available derivatives data–index option prices and gammas, VIX-futures, and VIX-option prices. Since this expression depends on all historical derivative and current derivative data, this expression yields a term structure of forward returns from a forward-backward perspective without relying on any model assumption about the equity index.

We present three applications of this expression of forward return from derivatives, including the pro-cyclic term structure of forward returns, robust autocorrelation analysis and short-term reversal pattern, and a profitable dynamic market-timing strategy. The forward return reveals future market drawdowns and captures upward market movements, yielding substantial economic value. Overall, we demonstrate the significance of derivatives market information in estimating expected returns in the future (Miller, 1999, page 100).

Table 2.1: Summary statistics of VIX and VIX-derivatives

This table provides the summary statistics for VIX, VIX futures, and implied volatility of VIX options. The sample period is from March 26, 2004 (February 24, 2006) to December 31, 2019 for VIX futures (options).

		Mean	Std dev	p25	Median	p75	Skew	Kurt
Panel A: VIX index								
		18.869	9.013	13.150	16.210	21.490	2.556	11.963
Panel B: VIX futures prices								
Maturity (in months)								
1		19.417	8.157	14.300	16.883	22.100	2.308	9.896
2		20.144	7.328	15.350	17.800	23.000	1.946	7.809
3		20.767	6.596	16.286	18.558	23.702	1.610	6.072
4		21.072	6.260	16.682	18.925	24.107	1.428	5.097
6		21.578	5.835	17.355	19.504	24.669	1.179	4.024
9		22.039	5.874	17.993	20.130	25.579	0.518	3.880
Panel C: Implied volatility of VIX options								
Maturity (in months)								
1	Put	0.891	0.168	0.789	0.871	0.974	1.268	8.293
	Call	0.893	0.164	0.788	0.875	0.978	1.139	7.033
2	Put	0.790	0.111	0.716	0.795	0.858	0.453	6.279
	Call	0.789	0.110	0.714	0.794	0.857	0.444	5.125
3	Put	0.717	0.089	0.660	0.724	0.776	0.159	4.776
	Call	0.716	0.089	0.657	0.723	0.774	0.245	5.208
4	Put	0.668	0.078	0.616	0.673	0.720	0.070	3.692
	Call	0.667	0.078	0.613	0.673	0.719	0.269	5.711
6	Put	0.630	0.071	0.579	0.635	0.678	0.067	3.396
	Call	0.628	0.072	0.576	0.634	0.676	0.153	4.176
9	Put	0.617	0.073	0.568	0.622	0.667	0.003	3.247
	Call	0.615	0.074	0.565	0.622	0.665	-0.019	3.161

Table 2.2: Expected future one-month return from the VIX-derivatives

This table provide the summary statistics for the expected future one-month return from the VIX-derivatives. The maturities are 1, 2, 3, 4, 6 and 9 months. We report mean, median, standard deviation, skewness and kurtosis. Panel A, B, and C consider three different sample periods: (i) full sample: February 24, 2006−December 31, 2019; (ii) Bad times (NBER recessions): January 1, 2008−June 30, 2009; and (iii) Good times (post-NBER recessions): July 1, 2009−December 31, 2019. All results are annualized and expressed in percentage.

	Avg. Ret (%)	Std dev (%)	p25	p50	p75	Skew	Kurt
Panel A: Sample period: February 24, 2006−December 31, 2019							
Maturity (in months)							
1	5.891	5.202	2.813	4.557	6.887	3.810	22.653
2	6.254	4.419	3.386	5.160	7.392	3.077	16.245
3	6.602	3.802	3.993	5.680	7.845	2.530	12.536
4	6.853	3.553	4.350	5.951	8.102	2.155	9.718
6	6.818	3.371	4.386	5.780	8.204	1.895	7.380
9	7.197	3.891	5.093	6.412	9.116	0.837	4.739
Panel B: Bad times (NBER recessions): January 1, 2008−June 30, 2009							
Maturity (in months)							
1	14.952	10.142	7.527	9.864	20.475	1.359	4.103
2	14.071	7.860	7.959	10.163	19.375	1.131	3.365
3	13.178	6.205	8.286	10.262	18.021	1.075	3.290
4	12.915	5.440	8.528	10.385	17.486	0.868	2.660
6	12.329	4.801	8.435	9.718	17.007	0.712	2.086
9	11.859	4.436	8.549	9.702	16.067	0.997	4.787
Panel C: Good times (post-NBER recessions): July 1, 2009−December 31, 2019							
Maturity (in months)							
1	4.332	2.477	2.569	3.683	5.233	2.056	8.809
2	4.891	2.399	3.066	4.357	5.753	1.574	5.815
3	5.442	2.356	3.582	4.968	6.331	1.330	4.548
4	5.781	2.360	3.939	5.294	6.714	1.289	4.395
6	6.409	2.451	4.499	5.724	7.366	1.174	3.754
9	7.446	2.931	5.360	6.439	9.011	1.023	3.683

Table 2.3: Market autocorrelation on S&P 500 index from the derivatives

Panels A and C report the statistics of the conditional market autocorrelation, $corr_t(R_{t\to t+T_1}, R_{t+T_1\to t+T_1+T_2})$ and the conditional correlation between two spot returns, $corr_t(R_{t\to t+T_1}, R_{t\to t+T_1+T_2})$, on S&P 500 index from the derivatives, respectively. We also compute them from historical stock returns in Panels B and D, respectively. The sample period is from February 24, 2006 to December 31, 2019. *** indicates significance at the 1% level.

T_1	1 month	2 months	3 months	4 months	6 months	9 months
T_2	1 month	1 month	1 month	1 month	1 month	1 month
Panel A: $corr_t(R_{t\to t+T_1}, R_{t+T_1\to t+T_1+T_2})$ from derivatives						
Mean	-0.209***	-0.279***	-0.362***	-0.313***	-0.268***	-0.257***
p25	-0.290	-0.364	-0.462	-0.383	-0.318	-0.346
p50	-0.196	-0.238	-0.346	-0.270	-0.251	-0.255
p75	-0.087	-0.151	-0.229	-0.194	-0.199	-0.204
Skew	-0.944	-0.645	0.252	-1.364	0.018	1.048
Kurt	5.467	7.696	8.711	5.814	8.116	8.160
Panel B: $corr_t(R_{t\to t+T_1}, R_{t+T_1\to t+T_1+T_2})$ by realized historical returns						
$\hat{\rho}$	0.093	0.044	0.079	0.118	0.035	0.029
Panel C: $corr_t(R_{t\to t+T_1}, R_{t\to t+T_1+T_2})$ from derivatives						
Mean	0.829***	0.944***	0.978***	0.982***	0.976***	0.971***
p25	0.734	0.912	1.000	1.000	1.000	1.000
p50	0.850	0.997	1.000	1.000	1.000	1.000
p75	0.963	1.000	1.000	1.000	1.000	1.000
Skew	-0.765	-8.093	-9.046	-4.847	-4.448	-3.833
Kurt	3.151	114.762	111.636	31.996	24.396	18.117
Panel D: $corr_t(R_{t\to t+T_1}, R_{t\to t+T_1+T_2})$ by realized historical returns						
$\hat{\rho}$	0.746***	0.841***	0.889***	0.919***	0.944***	0.961***

Table 2.4: Market timing

This table reports the investment value of timing the previous cumulative market excess return and conditional market autocorrelation. We consider six reversal signals, $\tilde{S}_{t,K}$, for $K = 1$, 2, 3, 4, 6, and 9 months.

The single timing strategy, $\eta\left[\tilde{S}_{t,1}\right]$ takes a long position in the market when the one-month reversal signal, $\tilde{S}_{t,1}\left[r_{t-1\to t}, corr_{t-1}(r_{t-1\to t}, r_{t\to t+1})\right]$ equals one, and invests in the risk-free asset otherwise. The combination timing strategy, $\eta\left[\tilde{S}_{t,K}, \forall K; \xi\right]$ utilizes all six reversal signals, and takes a long position in the market only if at least ξ out of six reversal signals take values of ones. We consider ξ to be 1, 2, 3, 4, and 5.

Panel A and B consider two different out-of-sample periods: 1) full sample period; 2) NBER recession period: January 2008−June 2009. The average value, standard deviation, and return-loss are expressed in percentage, and the performance fee is in basis points. All results are annualized.

	Avg ex-Ret (%)	Std dev (%)	SRatio	CEQ	SRatio Diff	CEQ Diff	Ret-Loss (%)	Fee (bps)
Panel A: Full sample period								
Buy and hold	5.489	14.789	0.371	0.044				
$\eta\left[\tilde{S}_{t,1}\right]$	4.433	9.616	0.461	0.040	0.090	-0.004	-0.863	-74.433
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 2\right]$	2.389	11.175	0.214	0.018	-0.157	-0.026	1.759	-286.856
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 3\right]$	1.997	9.986	0.200	0.015	-0.171	-0.029	1.709	-319.872
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 4\right]$	1.733	9.343	0.185	0.013	-0.186	-0.031	1.735	-343.248
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 5\right]$	3.347	7.166	0.467	0.031	0.096	-0.013	-0.687	-172.763
Panel B: NBER recessions: January 2008−June 2009								
Buy and hold	-32.304	25.565	-1.264	-0.356				
$\eta\left[\tilde{S}_{t,1}\right]$	-7.027	14.420	-0.487	-0.081	0.776	0.275	-11.194	2630.850
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 2\right]$	-23.918	20.746	-1.153	-0.261	0.111	0.095	-2.297	890.633
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 3\right]$	-11.842	18.780	-0.631	-0.136	0.633	0.220	-11.888	2116.325
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 4\right]$	-11.842	18.780	-0.631	-0.136	0.633	0.220	-11.888	2116.325
$\eta\left[\tilde{S}_{t,K}, \forall K; \xi = 5\right]$	-0.456	11.728	-0.039	-0.011	1.225	0.344	-14.364	3304.994

Table 2.5: Market timing by VIX-Gamma

This table reports the single timing strategy based on the 1-month reversal signal identified from the VIX-Gamma approach. Panel A and B consider two different out-of-sample periods: 1) full sample period; 2) NBER recession period: January 2008−June 2009. The average value, standard deviation, and return-loss are expressed in percentage, and the performance fee is in basis points. All results are annualized. The statistical significance of the Sharpe ratio difference (SRatio Diff) and certainty equivalent return difference (CEQ Diff) are evaluated based on *p*-values using the Jobson and Korkie (1981) methodology described in Section 2 of DeMiguel et al. (2009). ** and * indicate significance at the 5% and 10% levels, respectively.

	Avg ex-Ret (%)	Std dev (%)	SRatio	CEQ	SRatio Diff	CEQ Diff	Ret-Loss (%)	Fee (bps)
Panel A: Full sample period								
Buy and hold	5.489	14.789	0.371	0.044				
VIX-Gamma	5.659	11.322	0.500	0.050	0.129*	0.006**	-1.456	39.453
Panel B: NBER recessions: January 2008−June 2009								
Buy and hold	-32.304	25.565	-1.264	-0.356				
VIX-Gamma	6.538	19.480	0.336	0.046	1.599	0.402	-31.153	3948.062

Table 2.6: Implied risk-neutral correlation

This table reports the average value of the risk-neural correlation, $\theta_t = corr_t^{\mathbb{Q}}\left(R_{t\to t+1}, R_{t+1\to t+2}^2\right)$, that are "calibrated" from the VIX-Gamma approach. We compute the average values either by moving-averages (MAs) or by the calendar years.

By MAs:	6-month	1-year	3-year	5-year	Overall
	-0.046	-0.029	0.006	0.013	-0.060
By years:	2006−2009	2010−2011	2012−2014	2015−2017	2018−2019
	-0.059	-0.281	0.049	-0.092	0.040

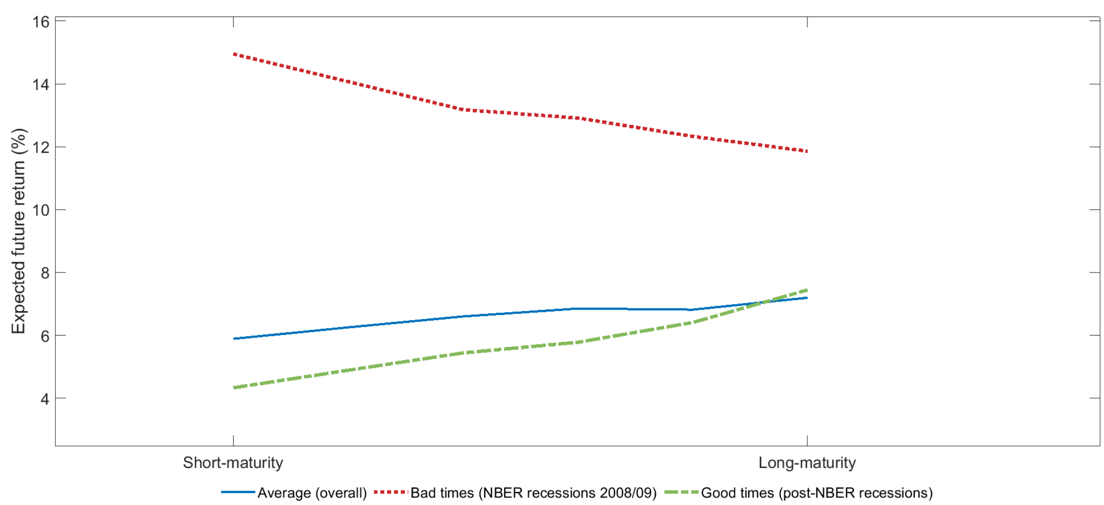


Figure 2.1: The term structure of expected future one-month return

This figure plots the term structure of the expected future one-month returns by VIX-derivatives. The figure shows the unconditional average return (solid line), the average return in bad times from January 2008 to June 2009 during the NBER recessions (dashed line), and the average return in good times during the post NBER recessions (dash-dotted line).

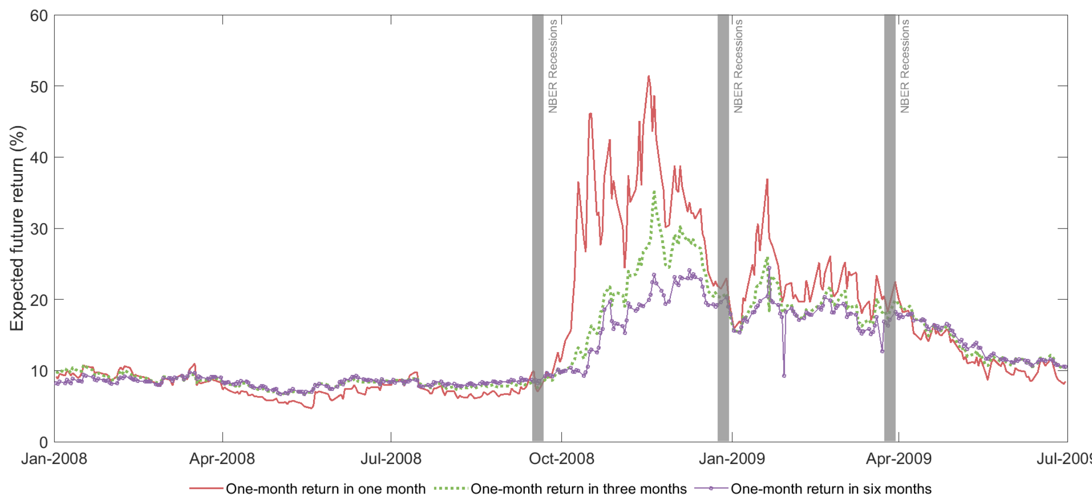


Figure 2.2: Expected future one-month returns during the NBER recessions

This figure plots the expected one-month returns in one, three, and six months by VIX-derivatives during the NBER recessions from January 1, 2008 to June 30, 2009. All results are annualized and expressed in percentage.

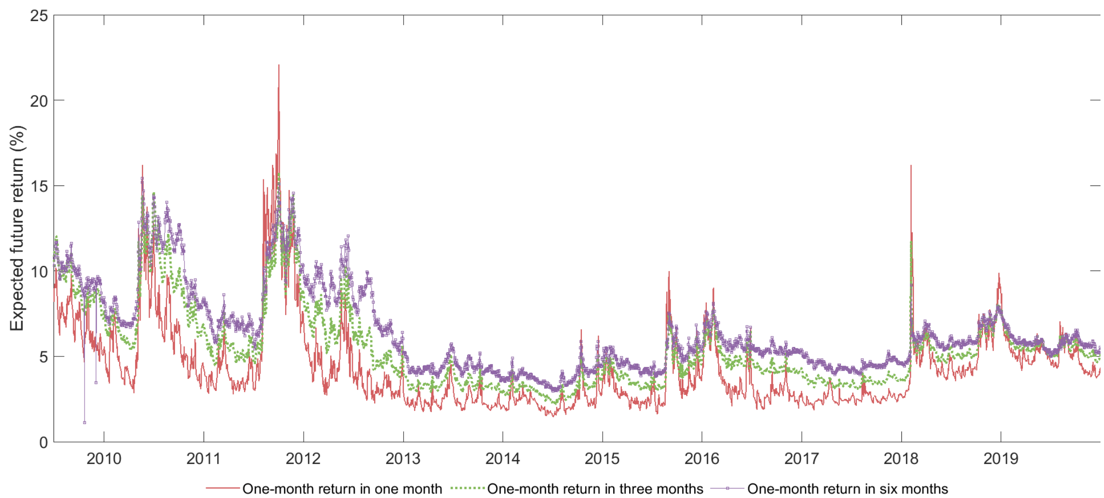


Figure 2.3: Expected future one-month returns post the NBER recessions

This figure plots the expected one-month returns in one, three, and six months by VIX-derivatives during the *post* NBER recession period from July 1, 2009 to December 31, 2019. All results are annualized and expressed in percentage.

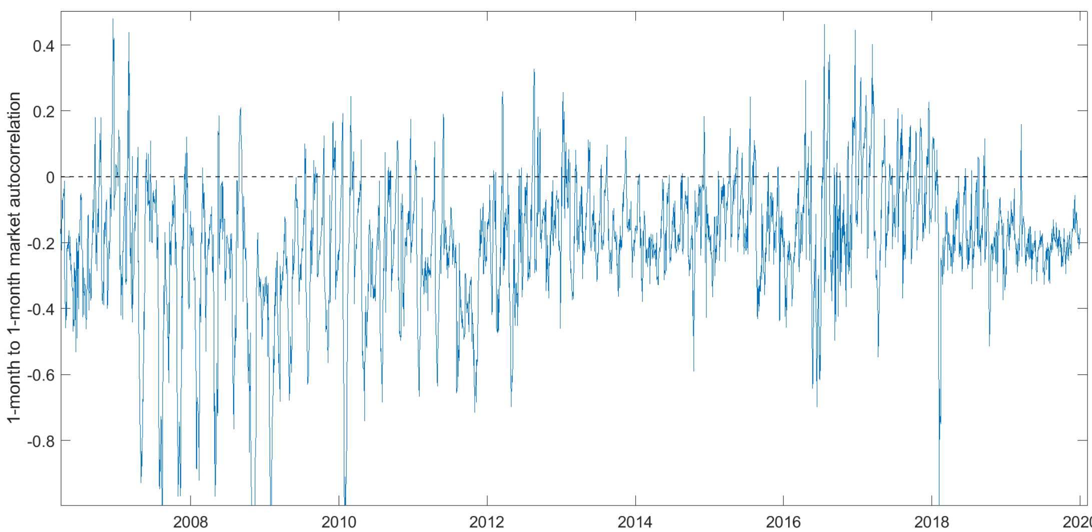
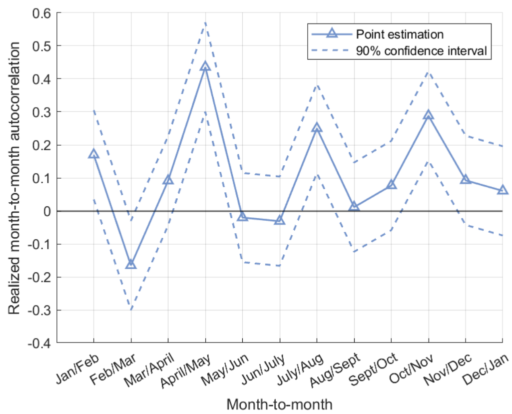
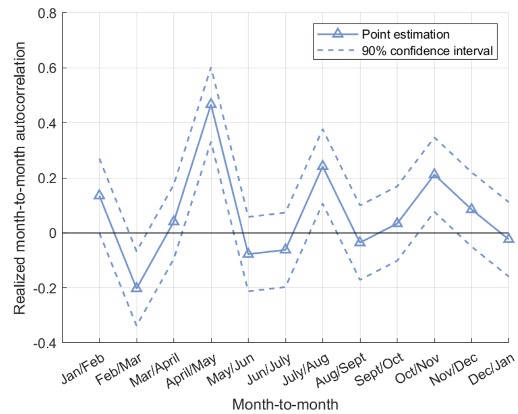


Figure 2.4: Market autocorrelation on S&P 500 index from VIX-derivatives

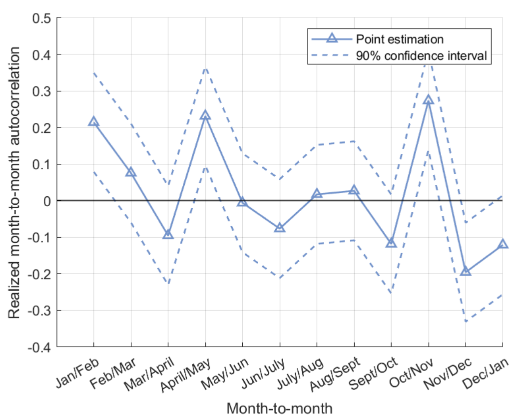
This figure plots the real-time forward-looking 1-month to 1-month market autocorrelation, $corr_t(R_{t\to t+1mo}, R_{t+1mo\to t+2mo})$, on S&P 500 index recovered from VIX-derivatives.



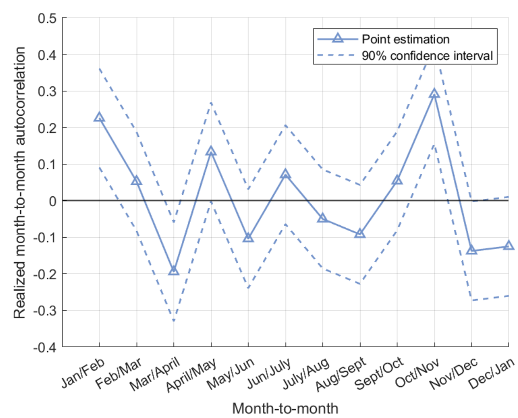
(a) Sample period 1871 - 2019



(b) Sample period 1920 - 2019



(c) Sample period 1970 - 2019



(d) Sample period 1990 - 2019

Figure 2.5: Realized market autocorrelation between adjacent calendar months

This figure plots the realized month-to-month autocorrelation of the S&P 500 monthly returns between two consecutive months. The area between the dotted line represents the 90% confidence interval for the sample autocorrelation by assuming the standard error equals one over the square root of the sample size. We consider four time periods: (a) 1871 – 2019, (b) 1920 – 2019, (c) 1970 – 2019, (d) 1990 – 2019. The data prior to January 1927 are obtained from Robert Shiller's website.

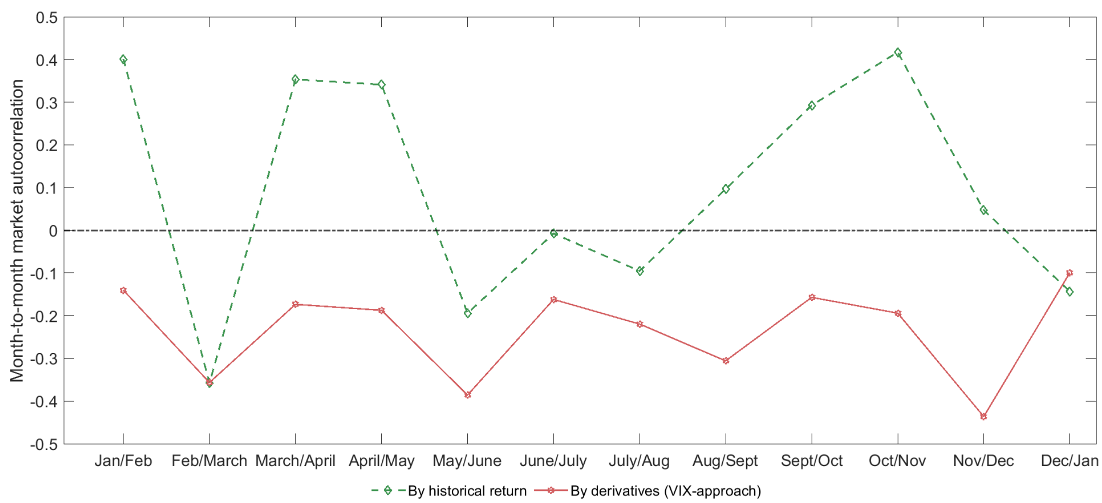


Figure 2.6: Month-to-month market autocorrelation by derivatives and historical data

This figure plots the month-to-month autocorrelation on S&P 500 index between two consecutive months. By historical return, we compute the sample autocorrelation using historical monthly return data; by VIX-approach, we compute $corr_t(R_{t\to t+1mo}, R_{t+1mo\to t+2mo})$ by derivative data on the first day of each month, and then take the average within January, February, ..., and December. The sample period is from 2006 to 2019.

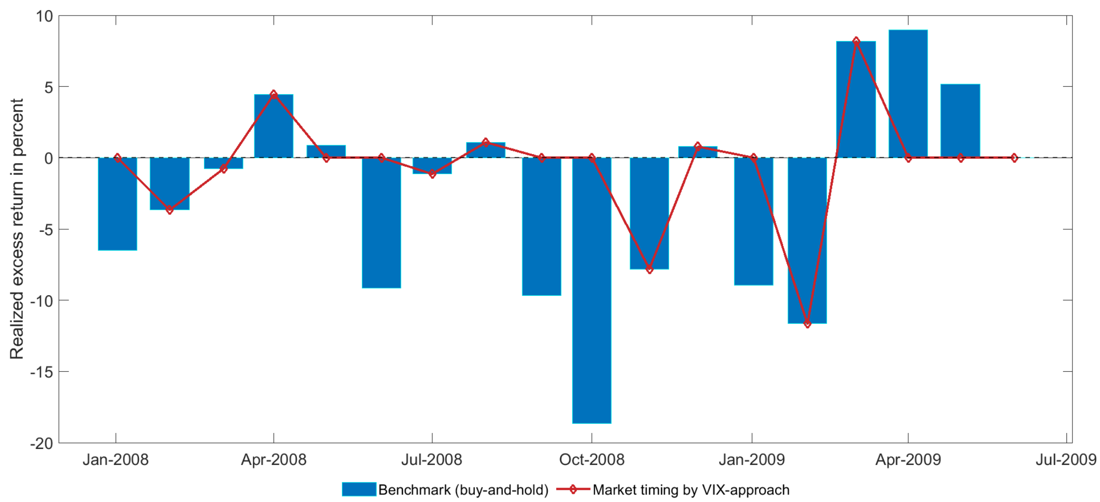
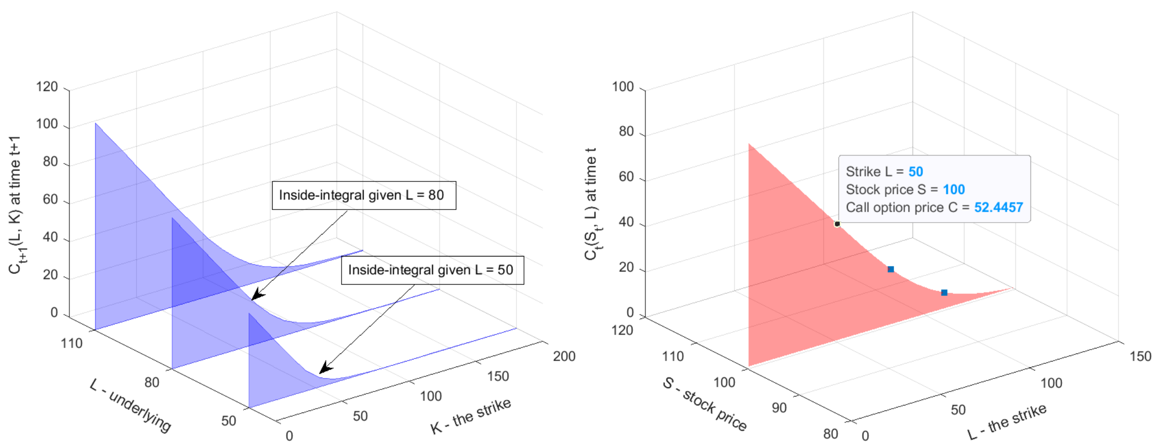


Figure 2.7: Market timing during NBER recessions

This figure plots the realized out-of-sample excess returns generated from either buy-and-hold strategy (benchmark) or the market timing strategy over the NBER recessions from January 2008 to June 2009. The market timing strategy, $\eta\left[\tilde{S}_{t,1}\right]$ takes a long position in the market when the one-month reversal signal, $\tilde{S}_{t,1}\left[r_{t-1\to t}, corr_{t-1}\left(r_{t-1\to t}, r_{t\to t+1}\right)\right]$ equals one, and invests in the risk-free asset otherwise.



(a) Call option at time $t+1$ (b) Call option at time t

Figure 2.8: Calculating the two-integral from call option prices

This figure illustrates the calculation of the two-integral in Proposition 2.4.1. At time t in the right-side panel, we plot the call option price, $C_t(S_t, L)$ for a sequence of strike prices, $L \geq 0$, assuming $S_t = 100$, $r_f = 5\%$, $\sigma = 25\%$, and $T = 1$ year. At time $t+1$ in the left-side panel, we plot the call option prices, $C_{t+1}(L, K)$, given each L "observed" at time t in the right-side panel as the new underlying prices, and for a sequence of strikes, $K \geq 0$.

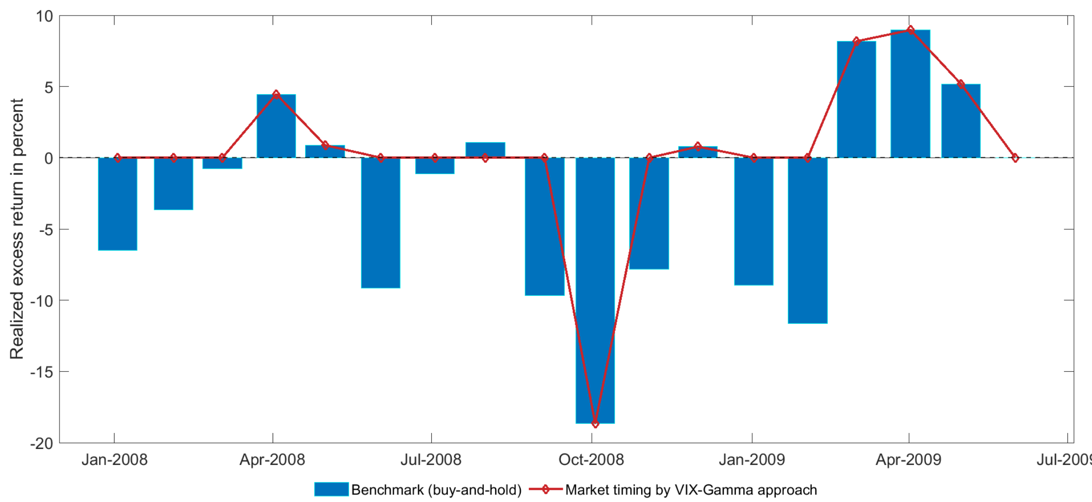


Figure 2.9: Market timing by VIX-Gamma approach during NBER recessions

This figure plots the realized out-of-sample excess returns generated from either buy-and-hold strategy (benchmark) or the market timing strategy over the NBER recessions from January 2008 to June 2009. The market timing strategy, $\eta\left[\tilde{S}_{t,1}\right]$ takes a long position in the market when the one-month reversal signal, $\tilde{S}_{t,1}\left[r_{t-1\to t}, corr_{t-1}\left(r_{t-1\to t}, r_{t\to t+1}\right)\right]$ equals one, and invests in the risk-free asset otherwise.

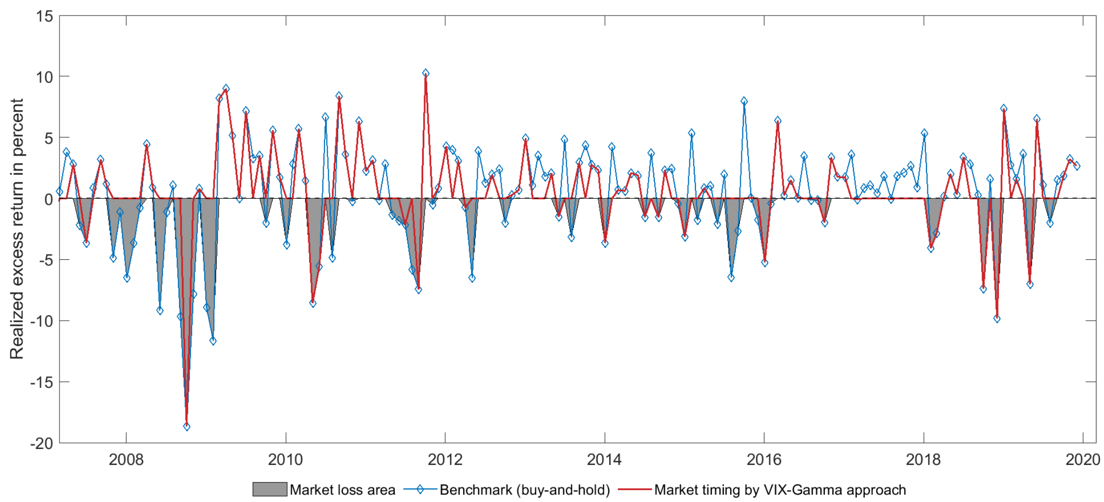


Figure 2.10: Market timing by VIX-Gamma approach

This figure plots the realized out-of-sample excess returns generated from either buy and hold strategy (benchmark) or market timing strategy over the out-of-sample evaluation period from 2006 to 2019. The market timing strategy, $\eta\left[\tilde{S}_{t,1}\right]$ takes a long position in the market when the one-month reversal signal, $\tilde{S}_{t,1}\left[r_{t-1\to t}, corr_{t-1}(r_{t-1\to t}, r_{t\to t+1})\right]$ equals one, and invests in the risk-free asset otherwise.

CHAPTER 3: MISPRICING AND ANOMALIES: AN EXOGENOUS SHOCK TO SHORT SELLING FROM JGTRRA

3.1 Introduction

There is a large asset pricing literature that documents that many firm characteristics can predict future stock returns (see e.g., Haugen and Baker, 1996; Bali et al., 2016; Harvey et al., 2016; Hou et al., 2020), yielding a number of anomalies that standard asset pricing models cannot explain. A fundamental question is what causes anomalies. Despite many studies on anomalies, researchers still disagree on the source of return predictability. The literature offers two major explanations.[1] First, return predictability could be a result of compensation for rational risks (e.g., Fama and French, 1992, 1998). Second, it could reflect mispricing due to limits to arbitrage (e.g., Shleifer and Vishny, 1997; Barberis and Thaler, 2003; Engelberg et al., 2018). In either case, anomalies exist for good economic reasons. However, empirically, it is difficult to distinguish between the two explanations as both explanations can exist (Lewellen, 2010; Lam and Wei, 2011).

This paper investigates the economic causes of anomalies in a comprehensive way. First, we analyze the issue for all the 182 significant ones out of 355 anomalies identified in the accounting, economics and finance literature. Second, we utilize the Job and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003 as a plausibly exogenous shock to short selling supply and examine its causal effect on anomalies. After the JGTRRA, equity lenders are reluctant to lend shares around the dividend record dates because "substitute dividends" that they would receive from short sellers are taxed at ordinary income rates

[1] A third explanation is that anomalies could result from data mining (Harvey et al., 2016). However, as shown by Chen (2021), it is unlikely to attribute all the anomalies to *p*-hacking. Additionally, Bowles et al. (2017) use a database with the precise release dates of accounting information to show that anomaly returns are real but quickly exploited after information announcements.

while qualified dividends are taxed at 15 percent, thus creating a negative shock to the short selling supply (Thornock, 2013).[2] In this paper, we investigate the *causal* effect of the differential dividend taxation-induced short selling shocks on mispricing and anomalies. Since JGTRRA is still in effect, and if we find that it impacts anomalies, it will serve as a persistently present arbitrage barrier to support the mispricing interpretation.

To investigate how JGTRRA shocks affect limits to arbitrage, we construct a cross-sectional aggregate net overpriced score (*NOPS*) from 182 anomalies. Stocks with the highest values of *NOPS* are the most “overpriced", whereas those with the lowest values are the most “underpriced". We focus on two main hypotheses: 1) Mispricing is stronger in the dividend record months compared to the other months after the JGTRRA of 2003, and as a result, anomalies are stronger in the subsequent months; 2) The effect mainly comes from the overpriced stocks.

To assess the causal effect of short selling on mispricing and anomalies, we use a stock-level difference-in-differences (DID) panel regression framework. Specifically, we regress future one-month stock returns on *NOPS*, a dividend record month dummy (*DivR*) and *JGTRRA* dummy, the interaction terms between *NOPS* and each of two dummy variables, and finally a three-way interaction term between *NOPS*, *DivR*, and *JGTRRA*. The coefficient on the three-way interaction term measures the DID effect, namely, the difference between after and before the enactment of JGTRRA of 2003 of the differences in the predictive power between the dividend record months and non-dividend record months. In other words, it captures the differential responses of anomalies to JGTRRA between following the dividend record months and following the other months. We show that the coefficients on the three-way interaction term, $NOPS \times DivR \times JGTRRA$, are significantly negative (the same sign as *NOPS*) at the 1% level for various fixed effects and clustering methods. These results indicate that after JGTRRA, anomalies become stronger following

[2] According to Thornock (2013), lending fees on average spike by 24% over the median rate, and loan quantities for tax-sensitive lenders decrease by 18% over the median quantities before the dividend record dates.

the dividend record months than following the other months, because stocks become more mispriced in the dividend record months when it is harder for arbitrageurs to do short sells.

To test the hypothesis that the effects of shocks to short selling come mostly from the overpriced stocks, we extend our DID analysis by adding *High NOPS*, *Low NOPS* and their corresponding interaction terms. If our conjecture is true, then we expect that the coefficient on *High NOPS*×*DivR*×*JGTRRA* would be significantly negative, while the coefficient on *Low NOPS*×*DivR*×*JGTRRA* would be insignificant. Indeed, we find that the coefficient on *High NOPS*×*DivR*×*JGTRRA* is -0.744 with a t-stat of -2.33, whereas the coefficient on *Low NOPS*×*DivR*×*JGTRRA* is positive but insignificant. In summary, our results are consistent with the mispricing explanation for anomalies. This tax-driven exogenous shock to short selling prevents arbitrageurs from exploiting overpricing and thereby amplifying anomalies.

We further demonstrate that our results are unlikely driven by risk or data-mining. We find that our results are robust to controlling for various dynamic risk factors including the market portfolio and five macroeconomic risk factors in Chen et al. (1986). Moreover, we conduct various placebo tests to address the data-mining concern. We change the timing of JGTRRA to various periods and re-estimate our DID regressions. We find that these fictitious scenarios do not have the same significant impact on anomalies between dividend record months and the other months. Next, we randomly create pseudo dividend record months without changing the timing of JGTRRA, and find that the coefficient of the three-way interaction term is always statistically insignificant.

We conduct a battery of additional robustness checks. First, we show that our results hold in a portfolio-level DID framework. Economically, in response to JGTRRA, the increase in the anomaly return is on average 1.677% higher after the dividend record months than after the other months, indicating a stronger response of anomalies after the dividend record months. Second, we find that the effect of shocks to short selling on mispricing is more pronounced in periods with higher investor sentiment and stronger for stocks with younger age, smaller size, higher idiosyncratic volatility, and lower size-adjusted institutional ownership. Lastly, we show that our results are robust to alternative sample periods, regression specifications, and mispricing measures.

Equipped with the large number of anomalies, we separate them into four groups: event, market, fundamentals and valuation following Engelberg et al. (2018), and construct the mispricing measure (*NOPS*) for each of the anomaly groups. We then conduct the DID analysis for each group and find that the impact of JGTRRA on anomalies is significant for event, market, and fundamentals anomalies but is insignificant for the valuation anomalies.[3] These results indicate that valuation anomalies are unlikely due to mispricing. Our evidence casts some doubts on the mispricing explanation for the value premium (see, e.g., Porta et al., 1997; Ali et al., 2003) and is consistent with various risk-based explanations in the literature.[4]

It is worth noting that the DID framework used in this paper only requires a rather weak exogeneity condition that the decision to issue dividends does not affect the evolution of anomalies over time, which is likely to be true as firms rarely change their dividend policies. In particular, it does not require that dividend stocks and non-dividend stocks are indistinguishable. In other words, dividend stocks can differ from non-dividend stocks in systematic ways as long as the differences do not depend on some time-varying unobservable that affect the anomalies because of the double differences approach. Nevertheless, we further test the robustness of our results using only firms that issue dividends so as to have a matched sample. In addition, Chetty and Saez (2005) report that a number of firms initiate dividends immediately after the enactment of the law, which in itself should not affect the evolution of anomalies over time since it is a one-time change. We nevertheless further

[3]We find similar results after separating anomalies into momentum, profitability, investment, intangible, trading frictions, or value & growth according to Hou et al. (2020). The effect is highly significant in all but value & growth anomalies.

[4]The risk-based explanations for the value premium include financial distress risk (Fama and French, 1995), investment risk (Berk et al., 1999), investment irreversibility and countercyclical price of risk (Zhang, 2005), consumption risk (Lettau and Ludvigson, 2001), aggregate cash flow risk (Campbell and Vuolteenaho, 2004). More recently, Gerakos and Linnainmaa (2018) show that value premium is mainly driven by variations in size. Ball et al. (2020) document that earnings yield explains value premium.

exclude all firms that initiate dividend payouts after the JGTRRA of 2003 to mitigate the potential concern that firms may initiate dividends in response to the act. An additional benefit of this exclusion is that we now have the same firms as treated (dividend record months) and control (other months) before and after JGTRRA. In other words, we have a setting that resembles the controlled experiment often available in physics or other natural science disciplines, but rarely available to finance or economics. Our results are robust in both samples.

Our study is related to Chu et al. (2020), which presently is the only other study to investigate the causal effect of short-selling constraints on anomalies. Chu et al. (2020) make a first attempt by using Regulation SHO (Reg SHO) as a quasi-exogenous shock to the short-sale constraint to study its impact on the 11 anomalies in Stambaugh et al. (2012) and argue that those anomalies reflect mispricing. While their results are indicative, it remains unclear whether or not mispricing is a pervasive phenomenon beyond those 11 ones. We utilize a new exogenous shock, which tightens the short-sale constraint instead of loosening it in Reg SHO and still remains effective today, to study its causal effect on all the anomalies documented in the literature. We provide strong evidence that mispricing drives anomalies in general but also find new evidence that if we separate the anomalies into different types, valuation anomalies do not seem to be driven by mispricing. The two papers complement each other by shedding light on the source of anomalies.

3.2 JGTRRA dividend tax cut and shocks to short selling

The Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003 is a tax law passed by the United States Congress on May 23, 2003. This law reduces the maximum federal tax rate on qualified dividends from 38.6% to 15%. This tax cut remains effective for taxpayers whose income does not exceed the thresholds set for the highest income tax.

JGTRRA provides a new opportunity for examining the causal effect of dividend taxation on financial markets. First, this tax cut was largely a surprise to the market prior to 2003 as it moved from an initial proposal to a signed law in under five months. Consequently, researchers generally consider it as an exogenous event. Second, JGTRRA is free of other major changes to the tax law that might confound the empirical analysis of its effects.

The JGTRRA dividend tax cut substantially affects short selling. Thornock (2013) first documents the effect of dividend taxation on short selling around the dividend record dates using proprietary short lending data between 2005 and 2007. He argues that dividend taxation can affect short selling through "loan effect", which stems from the different tax treatments for qualified and unqualified dividends. If a short seller borrows a stock over the dividend record date, then she repays the amount of dividend to the lender because the buyer in the short sale is the legal shareholder of record. This repayment is referred to as the "substitute dividend", which is taxed at the ordinary income rate rather than the rate of qualified dividends.

The following numerical example explains the above tax effect. An investor in the 35% marginal tax bracket owns 100,000 shares of a stock that has an annual dividend payment of $1.00. After JGTRRA of 2003, this dividend of $100,000 could be taxed at 15% and therefore the investor would pay $15,000 in taxes. However, if the investor lends the shares, she would pay $35,000 in taxes. This tax differential of $20,000 is economically large. Consequently, tax-sensitive equity lenders would increase their fees and decrease their lending quantities around the dividend record dates. Another adverse effect of dividend taxation on short selling is associated with dividends received deduction (DRD) from corporate income. The DRD allows for a 70% deduction on dividends received from other corporations. However, substitute dividends are not qualified for the DRD.

Dixon et al. (2021) also observe a significant tightening of the equity lending market around dividend record days. Blocher et al. (2013) find that prices of hard-to-borrow stocks surge around ex-dividend dates due to a decline in short selling supply driven by dividend taxation. However, none of these studies directly test a causal relation between dividend taxation and short selling. We document a causal relation between the short interest ratio and JGTRRA in the dividend record months in the later section.

3.3 Data and research design

In this section, we discuss the data, the construction of the aggregated mispricing measure, and the difference-in-differences panel regression framework used to detect the impact of shocks to short selling on mispricing and the strength of anomalies.

3.3.1 Data

We collect dividend information, prices, and monthly returns from the Center for Research in Security Prices (CRSP) between July 1965 and December 2019. One concern is that from 1954 to 1984, a dividend income exemption was introduced that initially started at $50, and a 4% tax credit for dividends above the exemption. After 1985, dividends were fully taxed under ordinary income rates, without any exemption, until the JGTRRA of 2003.[5] To that end, we restrict our main analysis using the sample from July 1985 to December 2019. We also restrict our sample to ordinary taxable cash dividends (CRSP distribution code = 1232) of $0.01 or greater that are paid by ordinary common shares listed on the NYSE, AMEX, or NASDAQ exchanges. We exclude stocks with prior month prices below $5 per share.

We define $DivR_{i,t}$ as a dummy variable that equals one if stock i reports a dividend record date in month t and zero otherwise. Panel A Table 3.1 provides the descriptive statistics of $DivR_{i,t}$ for our sample. In total, we obtain 1,588,481 firm-month observations with a mean *DivR* of 14.20%. We obtain firm information from CRSP/Compustat Merged annual and quarterly files, IBES, Thompson Reuter's 13F database, and OptionMetrics to construct anomaly variables.

3.3.2 Net overpriced score

We use a comprehensive set of anomalies to construct the mispricing measure. Our initial anomaly pool consists of 355 individual anomaly variables. These variables are

[5] For more details regarding the history of dividend tax rates in the U.S., please refer to https://www.dividend.com/taxes/a-brief-history-of-dividend-tax-rates/.

primarily drawn from Harvey et al. (2016); McLean and Pontiff (2016); Green et al. (2017); Engelberg et al. (2018); Hou et al. (2020); Chen and Zimmermann (2021).[6] We follow Green et al. (2017) to exclude variables that are insignificant in predicting future returns and end up with 182 significant ones.[7]

Inspired by Stambaugh et al. (2015) and Engelberg et al. (2018), we construct a cross-sectional aggregated mispricing measure, net overpriced score, (*NOPS*). Stocks with the highest values of *NOPS* are the most "overpriced", whereas those with the lowest values are the most "underpriced". We construct the mispricing measure as follows. Each month, we sort stocks into decile portfolios based on each anomaly characteristic. We use the extreme deciles to define the long or short side for each anomaly. Next, for each firm and month, we sum the number of short-side and long-side anomalies that the firm belongs to. Doing so produces *NShort* and *NLong*. Finally, the cross-sectional mispricing measure, *NOPS* is defined as $NShort - NLong$. Panel B Table 3.1 provides the descriptive statistics of *NOPS*. On average, a stock is in 12.54 short portfolios and 14.84 long portfolios. *NOPS* has a mean value of -2.29, a standard deviation of 10.12, a maximum value of 63, and a minimum value of -61.

3.3.3 Difference-in-differences regressions

Ideally, the best approach to identify a *causal* effect is a controlled (random) experiment that is often done in physics, biology, and other natural science disciplines. In finance and economics, the best available situation most of the time is a quasi-experiment such as JGTRRA or Reg SHO. In this case, the validity of the approach crucially depends on the identification assumption. An instrumental variable approach such as 2-step Least Square is often used if an exogenous variable that only affects the treated can be identified. Alter-

[6] Chen and Zimmermann (2021) cover all independent anomalies in Hou et al. (2020); 98% of the portfolios in McLean and Pontiff (2016); 90% of the characteristics from Green et al. (2017); and 90% of the firm-level predictors in Harvey et al. (2016). We thank Andrew Chen for making their data available. We also obtain additional variables from Han et al. (2016); **?**); Avramov et al. (2021), and among others.

[7] We also drop nine dividend-related anomalies such as dividend initiation, dividend omission, and dividend yield, and among others.

natively, if conditional exogeneity assumption (or selection on observables) holds, a simple difference approach or propensity score matching approach would suffice. However, if instead a weaker exogeneity condition holds such as exogeneity of selection to changes in outcomes, the appropriate approach is the difference-in-differences (DID) approach that compares the difference from before and after the quasi-experiment for a treated group to the same difference for a control group. In particular, DID approach does not require the treated and the control groups to be matched or random. Instead, it only requires that the selection does not change over time or if it changes, it will not affect the changes in the outcome. Since a firm's dividend policy rarely changes, it seems that the DID approach is appropriate to study the effect of JGTRRA.[8]

In this paper, we investigate the effect of the differential tax-driven shock to short selling on mispricing and anomalies in a stock-level DID panel regression framework. Specifically, we estimate the following regression equation,

$$\begin{aligned} ret_{i,t} = \alpha_0 + \alpha_t + \alpha_i + b_1 NOPS_{i,t-1} + b_2 DivR_{i,t-1} + b_3 NOPS_{i,t-1} \times DivR_{i,t-1} + b_4 NOPS_{i,t-1} \times JGTRRA_{t-1} \\ + b_5 DivR_{i,t-1} \times JGTRRA_{t-1} + b_6 NOPS_{i,t-1} \times DivR_{i,t-1} \times JGTRRA_{t-1} + \varepsilon_{i,t}, \end{aligned} \quad (3.1)$$

where $ret_{i,t}$ is the percentage return of stock i in month t; $DivR_{i,t-1}$ is a dummy variable indicating stocks that report dividend record dates in the previous months; $JGTRRA_{t-1}$ is a dummy variable which equals one if month $t-1$ is after May 2003 (after the JGTRRA of 2003); $JGTRRA_{t-1}$ itself is subsumed by the time fixed effect, and thus is dropped from the regression; α_t is the time fixed effect that captures the common factor and/or market-wide or economy-wide trends that drive the stock returns in both dividend record months and other months; and α_i is the firm fixed effect to mitigate the potential omitted variable bias.

The three-way interaction term, *NOPS*×*DivR*×*JGTRRA* captures the moderating effect of the JGTRRA of 2003 and the dividend record months on the predictive power of *NOPS*. The DID coefficient b_6 is the coefficient of interest, capturing the difference between the

[8] JGTRRA also allows for a situation resembling a controlled experiment, which will be discussed in detail in Section **??**.

dividend record months and non-dividend record months in their respective changes of *NOPS*' predictive power between after and before the enactment of JGTRRA of 2003. In other words, it captures the differential response of anomalies to JGTRRA between the dividend record months and the non-dividend record months. If our hypothesis is true, we would expect b_6 to be significantly negative because of stronger mispricing in the dividend record months.

We also investigate the causal effect of short selling on anomalies in a portfolio-level DID panel regression framework following Chu et al. (2020). To construct the decile portfolios, for stocks that report dividend record dates in the previous months, we sort them into ten deciles based on NOPS and then compute the monthly portfolio returns. We repeat the procedure for stocks that do not report dividend record dates in the previous months. Then, we pool the monthly returns of the two types of decile portfolios and run the following panel regressions,

$$y_{i,t} = \alpha_0 + \alpha_t + \beta_0 Treated_{t-1} + \beta_1 Treated_{t-1} \times JGTRRA_{t-1} + \varepsilon_{i,t}, \tag{3.2}$$

where the dependent variable, $y_{i,t}$, is the monthly return of a decile portfolio in month t, $Treated_{t-1}$ is a dummy variable that is equal to one if the portfolio is formed on stocks whose $DivR_{i,t-1} = 1$, and α_t is the time fixed effect. Similar to the stock-level DID regression, the coefficient of interest is β_1, which captures the difference between the dividend record months and the other months in their respective differences in portfolio returns after versus before the JGTRRA of 2003. In other words, it captures the differential impact of JGTRRA on anomaly returns after the dividend record months versus after the non-dividend record months. If our hypothesis is true, we would expect β_1 to be significantly positive (negative) for the long-short portfolio (short side) because of stronger mispricing in the dividend record months.

3.4 Shocks to short selling and mispricing

In this section, we examine the hypothesis that mispricing is stronger in dividend record months compared with other months after the JGTRRA of 2003 periods. We also investigate whether our hypothesis holds after the Reg SHO program periods.

3.4.1 Stock-level difference-in-differences analyses

In this subsection, we report the effect of the tax-driven shock to short selling on mispricing and anomalies in a stock-level DID panel regression framework as discussed in subsection 3.3.3. Recall that the three-way interaction term in Equation (3.1), $NOPS \times DivR \times JGTRRA$ captures the moderating effect of the JGTRRA of 2003 and the dividend record months on the predictive power of *NOPS*, and thus its coefficient, b_6, represents the differential impacts of JGTRRA on the predictive power of *NOPS* between the dividend record months and the non-dividend record months. If our hypothesis is true, we would expect b_6 to be significantly negative.

Table 3.2 reports the coefficients of Equation (3.1), from b_1 to b_6 for specifications with various fixed effects and clustering methods. We find that our overpricing measure, *NOPS*, is highly significantly negative in each column. In the last column, the coefficient of *NOPS* is -0.100 with a t-stat of -11.12, confirming the strong negative return predictability for *NOPS*. Interestingly, we also observe significantly negative coefficients of $DivR_{i,t-1}$, indicating a negative return after the dividend record month before JGTRRA. This result is consistent with Hartzmark and Solomon (2013) findings of high returns before the ex-dividend day and the subsequent reversals afterward.[9] Additionally, an insignificant coefficient on $DivR \times JGTRRA$ suggests that this price pattern is indifferent to the tax law change. Furthermore, the interaction term between the overpricing score and the dividend

[9] More specifically, they show that "a portfolio that longs companies in the month of their predicted dividend and shorts same companies in other months (within companies) earns abnormal returns of 37 basis points", thereby generating the dividend premium. Significant reversals are observed in the 40 days after the ex-dividend day. The effect is argued to be driven by price pressure from dividend-seeking investors in the lead-up to the ex-dividend day.

record month dummy variable is always significantly positive. This finding implies that before JGTRRA, the return predictability of *NOPS* becomes weaker after the dividend record months, likely due to the dividend month premium effect on the underpriced stocks. Next, we show a significant positive coefficient for *NOPS*×*JGTRRA* in each column, consistent with the findings that anomalies are weaker in the recent two decades (see, e.g, Chordia et al., 2014).

Consistent with our prediction, the coefficients on the three-way interaction term are significantly negative across all columns. For example, in the last column with firm and time fixed effects and double clustered standard errors, b_6 is -0.028 with a t-stat of -2.89, which is significant at the 1% level. These results indicate a significant increase in the predictive power of *NOPS* to future returns in response to the enactment of JGTRRA for the dividend record months relative to the other months.

We conduct several robustness checks. First, we repeat our analysis using several balanced sample periods including 5, 10, or 15 years before and after the JGTRRA of 2003. Next, we repeat our analyses using an alternative proxy for mispricing, which is the mispricing score in Stambaugh et al. (2015).[10] The results are reported in Table B.1 and B.2, respectively, in the Appendix. Apart from them, in the Online Appendix, we report results extending the sample period up to July 1965 or replacing the firm fixed effect with the industry fixed effect (3-digit SIC codes). All these results deliver the same message as Table 3.2. Collectively, we demonstrate that this tax-driven shock to short selling tightens short-sale constraints, thereby causing stocks to be more mispriced in the dividend record months than in the non-dividend record months.

3.4.2 Does this effect hold after the Reg SHO program period?

Reg SHO was in effect from May 2, 2005 to August 6, 2007. After the program period, the SEC eliminated short-sale price tests for all exchange-listed stocks. Consequently,

[10] *MISP* is constructed using 11 anomalies studied in Stambaugh et al. (2012). The data of *MISP* that ends in December 2016 is available on Robert Stambaugh's website: `http://finance.wharton.upenn.edu/~stambaug/`.

the short-sale constraint imposed by the uptick rule was removed for all stocks. In this subsection, we investigate whether the effect of the differential tax-driven shock to short selling on mispricing and anomalies still prevails after the Reg SHO program period. Doing so could avoid the confounding effects and further validate the robustness of this dividend taxation shock to short-sale constraints.

Since the JGTRRA was enacted at the end of May, 2003, we exclude sample periods between June, 2003 and July, 2007 and redo our DID analysis. Table 3.3 reports our regression results. We find that the coefficients on the three-way interaction term are always negatively significant. For example, using firm and time fixed effects and double clustering method, the coefficient of $NOPS \times DivR \times JGTRRA$ is -0.022 with a t-stat of -2.64. Our results indicate that the effect of the dividend taxation shock to short selling on mispricing and anomalies is powerful even after the Reg SHO, when the short-sale price tests are eliminated for all stocks.

3.4.3 Overpricing from the tax-driven shock to short selling

Anomalies could reflect mispricing. In the presence of limits to arbitrage such as short-sale constraints, negative information could be slowly incorporated into stock prices. Therefore, overpriced stocks earn lower future stock returns and contribute to return predictability. Because this tax-driven shock to short selling tightens short-sale constraints, its effect on anomalies should be mainly manifested on the overpriced stocks, which are concentrated in the short leg of the anomalies.

We construct two dummy variables, *Low NOPS* and *High NOPS*, based on the decile rank of *NOPS* each month. *Low NOPS* identifies the most underpriced stocks, while *High NOPS* represents the most overpriced stocks. In other words, *High NOPS* represents stocks in the short side of anomalies, whereas *Low NOPS* reflects stocks in the long side of anomalies. Next, we add *High NOPS* or *Low NOPS* individually or together, along with their respective interactions with *JGTRRA* and *DivR* to our DID regression in Equation (3.1). The new specification including both dummies is as follows,

$$
\begin{aligned}
ret_{i,t} = {} & \alpha_0 + \alpha_t + \alpha_i + b_1 DivR_{i,t-1} + b_2 DivR_{i,t-1} \times JGTRRA_{t-1} + b_3 NOPS\ DP_{i,t-1} + b_4 NOPS\ DP_{i,t-1} \times DivR_{i,t-1} \\
& + b_5 NOPS\ DP_{i,t-1} \times JGTRRA_{t-1} + b_6 NOPS\ DP_{i,t-1} \times DivR_{i,t-1} \times JGTRRA_{t-1} + b_7 Low\ NOPS_{i,t-1} \\
& + b_8 High\ NOPS_{i,t-1} + b_9 Low\ NOPS_{i,t-1} \times DivR_{i,t-1} + b_{10} Low\ NOPS_{i,t-1} \times JGTRRA_{t-1} \\
& + b_{11} High\ NOPS_{i,t-1} \times DivR_{i,t-1} + b_{12} High\ NOPS_{i,t-1} \times JGTRRA_{t-1} \\
& + b_{13} Low\ NOPS_{i,t-1} \times DivR_{i,t-1} \times JGTRRA_{t-1} + b_{14} High\ NOPS_{i,t-1} \times DivR_{i,t-1} \times JGTRRA_{t-1} + \varepsilon_{i,t}, \qquad (3.3)
\end{aligned}
$$

where $NOPS\ DP_{i,t-1}$ represents the decile rank of stock i based on its *NOPS* in month $t-1$.

The results are presented in Table 3.4. Consistent with our prediction, the effect mainly comes from the overpriced stocks. The first column reports the regression results for *Low NOPS* alone. The coefficient on the three-way interaction term, *Low NOPS*$\times$*DivR*$\times$*JGTRRA*, is positive but statistically insignificant. This result implies that underpriced stocks play little role in driving the differential changes in return predictability of anomalies across the dividend record months and the other months after the JGTRRA. The second column presents the result for *High NOPS* alone. The coefficient on the three-way interaction term, *High NOPS*$\times$*DivR*$\times$*JGTRRA*, is -0.738 with a t-stat of -2.30, significant at the 5% level. The last column reports the result in Equation (3.3) after considering *Low NOPS* and *High NOPS* together. We obtain similar results. The coefficient on *Low NOPS*$\times$*DivR*$\times$*JGTRRA* is insignificant, while the coefficient of *High NOPS*$\times$*DivR*$\times$*JGTRRA* is significantly negative. These results indicate that after JGTRRA, stocks in the short leg of anomalies become more overpriced in the dividend record month than in the other months compared to before JGTRRA. Consequently, this pattern drives the effect of the tax-driven shock to short selling on the strength of the anomalies.

3.4.4 The risk-based explanation

In this subsection, we examine whether exposure to systematic risks can explain why anomalies become stronger following the dividend record months after JGTRRA. In a dynamic risk premia model, our results could potentially hold if the risk premia of stocks change after the dividend record months, i.e., the betas of stocks change in the months after the dividend record months. For example, increases (decreases) in betas for *Low* (*High*) *NOPS* stocks after the dividend record months can result in stronger anomalies.

We consider the market factor (MKT) and five macroeconomic risk factors of Chen et al. (1986), including log-change in monthly industrial production index (MP), unexpected inflation (UI), change in expected inflation (DEI), change in term premium (UTS), and change in default premium (UPR).[11] To examine whether the dynamic risk can explain our results, we re-estimate our stock-level DID regression in the previous section by adding each factor and their corresponding interaction terms. Specifically, we modify Equation (3.3) by interacting each factor with the dummies, *High* or *Low NOPS*, $JGTRRA$, *DivR* individually or in combinations. Table 3.5 reports the results. We find similar results to Table 3.4 after controlling for the effects of the dynamic risk premia of various macro factors. The coefficients of *High NOPS*×*DivR*×*JGTRRA* are always significantly negative regardless of which factor is used, whereas the coefficients of *Low NOPS*×*DivR*×*JGTRRA* are always insignificant. Similar results are obtained when all six factors are included as shown in the last column of Table 3.5. For example, the coefficient is -1.242 with a t-stat of -3.49 for the triple interaction term *High NOPS*×*DivR*×*JGTRRA*, and 0.122 with a t-stat of 0.67 for *Low NOPS*×*DivR*×*JGTRRA*.

Overall, our findings in this subsection suggest that risk is unlikely to explain the findings that overpriced stocks largely contribute to the effect of the tax-driven shock on anomalies.

3.4.5 Placebo tests

Data-mining and repeated use of the same data have always been a concern in finance (see, e.g., Harvey et al., 2016; McLean and Pontiff, 2016; Linnainmaa and Roberts, 2018). For instance, Heath et al. (2022) show that the repeated use of the Reg SHO pilot program increases the likelihood of false discoveries. We alleviate this concern by exploiting a

[11] The first three factors data used in Liu and Zhang (2008) can be downloaded from Laura Liu's website: `http://lauraxiaoleiliu.gsm.pku.edu.cn/en_research.htm`. The data for term premium and default premium are obtained from Amit Goyal's website `http://www.hec.unil.ch/agoyal/`.

novel exogenous shock to short selling, and thus it is less likely to be spurious. However, to further guard against spurious results, we conduct several falsification tests for our main DID analysis.

First, we conduct various placebo tests by changing the timing of JGTRRA, while maintaining dividend record dates for each stock. We use the timing before and after 2003 for pseudo enactment of JGTRRA including July of 1997 and 2013, and January of 1999 and 2005. To avoid the actual effect of JGTRRA, we use two sample periods: (1) between July of 1985 and May of 2003; (2) between June of 2003 and December of 2019 for our placebo tests. We run the difference-in-differences regression as follows,

$$\begin{aligned} ret_{i,t} = & \alpha_0 + \alpha_t + \alpha_i + b_1 NOPS_{i,t-1} + b_2 DivR_{i,t-1} + b_3 NOPS_{i,t-1} \times DivR_{i,t-1} \\ & + b_4 NOPS_{i,t-1} \times PseudoJGTRRA_{t-1} + b_5 DivR_{i,t-1} \times PseudoJGTRRA_{t-1} \\ & + b_6 NOPS_{i,t-1} \times DivR_{i,t-1} \times PseudoJGTRRA_{t-1} + \varepsilon_{i,t}, \end{aligned} \tag{3.4}$$

where $PseudoJGTRRA_{t-1}$ is a dummy variable which equals one if month $t-1$ is after each of these four pseudo JGTRRA periods, respectively, and zero otherwise.

Table 3.6 shows that in none of these placebo tests, the coefficients of $NOPS \times DivR \times PseudoJGTRRA$ are significantly negative. Instead, they are significantly positive for pseudo events before 2003, and statistically insignificant for pseudo events after 2003.

Next, we conduct several placebo tests on the dividend record months. We consider two testing samples: (1) excluding only the dividend record months; (2) excluding dividend-paying stocks altogether. For each sample, in each month, we randomly choose 14% of firm-months observations (based on summary statistics in Table 3.1) to be the dividend record months. Consequently, we create a pseudo dividend record month dummy variable for each sample. We run the following regressions using different simulation seeds,

$$
\begin{aligned}
ret_{i,t} =& \alpha_0 + \alpha_t + \alpha_i + b_1 NOPS_{i,t-1} + b_2 PseudoDivR_{i,t-1} + b_3 NOPS_{i,t-1} \times PseudoDivR_{i,t-1} \\
&+ b_4 NOPS_{i,t-1} \times JGTRRA_{t-1} + b_5 PseudoDivR_{i,t-1} \times JGTRRA_{t-1} \\
&+ b_6 NOPS_{i,t-1} \times PseudoDivR_{i,t-1} \times JGTRRA_{t-1} + \varepsilon_{i,t}, \qquad (3.5)
\end{aligned}
$$

where $PseudoDivR_{i,t-1}$ is a dummy variable indicating the pseudo dividend record month of $t-1$ for stock i.

The results are presented in Table 3.7. The coefficient of interest is essentially zero and statistically insignificant in each column. In sum, our placebo tests indicate that our results are not spurious and that anomalies become stronger following the dividend record months after JGTRRA.

3.5 Portfolio-level DID and subsample analyses

In this section, we provide portfolio-level DID analyses to further confirm our findings in the previous sections. We first conduct portfolio-level DID regressions to re-examine our hypotheses. We then provide a series of subperiod or subsample analyses to further confirm the causal effect of short selling on anomalies.

3.5.1 Portfolio-level DID analyses

In this subsection, we investigate the causal effect of short selling on mispricing and anomalies in a portfolio-level DID panel regression framework described in Subsection 3.3.3.

Table 3.8 reports the results of the portfolio-level DID regressions for the long-short portfolio as well as the long side and short side, separately. For the long-short portfolio, the coefficient of the interaction term, β_1, is 1.677 with a t-stat of 4.04, further confirming our first hypothesis that anomalies become stronger after the dividend record months than after the other months in response to the JGTRRA of 2003.[12] Specifically, the results

[12] We also obtain qualitatively similar results using value-weighted and gross-return-weighted portfolios. The results are presented in the Online Appendix.

show that in response to JGTRRA, the change in the anomaly returns is on average 1.677% higher after the dividend record months than after the other months. It is worth noting that it does not mean that the anomaly return is higher after the dividend record months than after the other months. It merely signifies that JGTRRA has a stronger impact on anomalies following the dividend record months than following the other months, and this is because mispricing is stronger in the dividend record months due to the tax-driven shock to short selling. Indeed, β_0, the coefficient of $Treated$, is -2.083, highly significant and larger than β_1 in magnitude, suggesting that anomalies are much weaker after the dividend record months before JGTRRA. This result is consistent with that reported in Table 3.2. In contrast, anomalies are much less so following the dividend record months after JGTRRA.

We also find that the short side dominates the long side portfolio in driving the causal effect of short selling on anomalies, confirming our previous findings. The coefficient on $Treated \times JGTRRA$ in the short side is -1.284 compared with 0.393 in the long side. Economically, after JGTRRA, the short side contributes to around 77% of the difference in anomaly profit change between the dividend record months and the other months.[13]

Overall, our portfolio-level results confirm our findings in the stock-level analyses. We document that this dividend taxation shock imposes greater constraints to short selling in the dividend record months and thereby causing more overpricing in the short legs of the anomalies. As a result, anomalies become stronger following the dividend record months compared with the other time periods.

3.5.2 Investor sentiments

In this subsection, we examine how investor sentiment impacts the relation between short selling and anomalies. Stambaugh et al. (2012) argue that when investor sentiment is high, overpricing becomes more prevalent and thereby anomalies become stronger. If anomalies are driven by mispricing, then the causal effect of JGTRRA should be more pronounced in

[13]The significance of the coefficient on $Treated \times JGTRRA$ for the long side is not robust, and becomes insignificant in the subsequent analyses.

the periods of high investor sentiment.

We use two orthogonalized investor sentiment indices from Baker and Wurgler (2006) and Huang et al. (2015) to identify high and low sentiment periods.[14] We obtain the mean of each index using an expanding-window approach with at least twenty-four monthly observations. A high (low) sentiment month is the one in which the value of sentiment index at the end of previous month is above (below) the estimated mean value. Next, we re-run the portfolio-level DID regressions in Table 3.8 separately over the high and low sentiment periods, for the long side, short side, and long-short portfolio, respectively. Table 3.9 reports the coefficient of interest, β_1, for each regression.

The left panel describes the results using Baker and Wurgler (2006) sentiment index. For the long-short portfolios, the coefficient on $Treated \times JGTRRA$ is 3.195 (t-stat = 3.48) in the high sentiment periods compared with 0.909 (t-stat = 2.20) in the low sentiment periods. These results accord well with our hypotheses. For the short side, the high sentiment periods display a significantly negative coefficient on $Treated \times JGTRRA$, while the low sentiment periods accompany an insignificant β_1. After JGTRRA, when the investor sentiment is high, stocks in the short legs of the anomalies become more overpriced in the dividend record months than in the other months. In contrast, we do not observe such pattern in the low investor sentiment periods. Furthermore, we find little variation in β_1 between the high and low sentiment periods for the long side. We find qualitatively similar results using the sentiment index of Huang et al. (2015). In unreported analysis, we also identify the high or low sentiment periods using the full-sample median value and find similar results. For the long-short portfolios, β_1 is 1.858 (t-stat =3.08) in the high sentiment periods as opposed to 0.819 (t-stat =1.69) in the low sentiment periods.

Overall, we obtain a substantial variation in the causal effect of the shock to short selling on anomalies with respect to investor sentiment. These results also strengthen our second hypothesis that the effect mainly comes from the overpriced stocks.

[14] We thank Jeffrey Wurgler for sharing the investor sentiment data on his website at: `http://people.stern.nyu.edu/jwurgler/`.

3.5.3 Subsamples

We further explore the causal effect of the tax-driven shock to short selling on anomalies across stocks with various degrees of limits to arbitrage. If our results are driven by overpricing due to the negative shock to short selling, then the effect should be stronger for stocks with higher limits to arbitrage.

We consider various proxies for limits to arbitrage including firm age, firm size, idiosyncratic volatility, and size-adjusted institutional ownership. Stocks of young and small firms are faced with greater limits to arbitrage as they are more costly and difficult to arbitrage (Israel and Moskowitz, 2013). Idiosyncratic volatility could reflect risks that deter arbitrage (Stambaugh et al., 2015). Nagel (2005) shows that short-sale constraints are most likely to be binding among stocks with low size-adjusted IO, which is a proxy for short selling supply.[15]

For each of the limits to arbitrage proxies, we define high or low groups based on tercile portfolios. We then conduct the portfolio-level DID analysis for each subsample. For the analysis of size-adjusted IO, we exclude stocks in the lowest decile rank of IO because the dividend taxation shock might have a marginal effect on these stocks, which are already highly constrained due to the lack of short selling supply. Panel A of Table 3.10 reports the results for firm size and firm age, and Panel B reports the results for idiosyncratic volatility and size-adjusted IO. Consistent with our expectation, we observe substantially larger coefficients on $Treated \times JGTRRA$ in smaller and younger stocks, stocks with higher idiosyncratic volatility, and lower size-adjusted IO for the long-short portfolios. For example, the coefficient of $Treated \times JGTRRA$ is 1.510 (t-stat $= 3.78$) for small stocks compared with 0.833 (t-stat $= 2.54$) for large stocks. We also find that for the short legs of the anomalies, the DID coefficients are considerably more negative for stocks with smaller size, younger age, higher idiosyncratic volatility, and lower size-adjusted IO. For instance, the coefficient

[15] To calculate size-adjusted IO, we first obtain the logit of IO and then run cross-sectional regressions of the logit(IO) on the logarithm of firm size and squared logarithm of firm size each quarter. The residuals in the regressions are referred to size-adjusted IO. Following Nagel (2005), we lag IO by two quarters.

of $Treated \times JGTRRA$ is -1.299 (t-stat $= -3.06$) for small stocks compared with -0.543 (t-stat $= -1.35$) for large stocks. However, all the DID coefficients are insignificant for the long legs of the anomalies.

In summary, the evidence in Table 3.10 provides additional support for the causal effect of short selling on anomalies.

3.5.4 Dividend stocks only

One of the advantages of a DID regression framework is the weak requirement of exogeneity. For example, it only requires that the selection into dividend record months does not affect the evolution of anomalies over time. Since firms rarely change their dividend policy, this exogeneity condition should hold.

Nevertheless, we test the robustness of our results using only firms that issue dividends so as to have a matched sample. Specifically, in this sample, the treated and control observations are all from the same dividend stocks. Panel A of Table 3.11 reports the results. The DID coefficient, β_1, is 0.753 (t-stat $= 3.06$) for the long-short portfolio, -0.550 (t-stat $= -2.50$) for the short side, and 0.203 (t-stat $= 1.33$) for the long side, respectively. These results are consistent with the evidence in Table 3.8.

One potential issue about the exogeneity of JGTRRA is that some firms may start to pay dividends because of the dividend tax cut. For example, Chetty and Saez (2005) show an increase in dividend initiations immediately after JGTRRA for non-financial and non-utility firms. On the other hand, Brav et al. (2008) argue that the tax cut merely imposes a marginal effect on a firm's dividend policy. They show that dividend initiations indeed temporarily spike after the act, but then return to pre-JGTRRA levels.[16] In addition to the low number of tax cut-induced dividend initiations, this one-time change should not affect the evolution of anomalies over time, and thus it does not violate the weak exogeneity requirement of the DID analysis. Nevertheless, further excluding these stocks creates an

[16]They find that between 2002 and 2005, 76 out of 265 firms initiated dividends after the act. Only a few firms occasionally mention the dividend tax cut as the reason for their initiations in their press releases. In our sample, we observe an increase of 66 dividend stocks to a total of 1555 from 2003Q3 to 2004Q3.

interesting setting resembling a controlled experiment – identical samples are randomly chosen either to be the treated (dividend record months) or the control (the other months) before the shock.[17] We repeat the above analysis by keeping only stocks that pay dividends before JGTRRA. We assume that it takes three months for a firm to change its dividend policy after JGTRRA. Thus, we drop stocks that initiated dividends after September 2003. Panel B of Table 3.11 reports the new DID results. We observe virtually the same results: β_1 is 0.517 (t-stat $= 2.01$) for the long-short portfolio and -0.514 (t-stat $= -2.54$) for the short side, while it is only 0.03 (t-stat $= 0.02$) for the long side. It is worth noting that β_0, the coefficient of the *treated* on the short side is insignificant in both panels, contrary to the significant coefficients in Table 3.8, confirming that the treated and the control behave the same before the act.[18]

We also use the within-firm calendar-time portfolio method in Hartzmark and Solomon (2013) to test the robustness of our results. Hartzmark and Solomon (2013) argue that the within-firm portfolios are likely to have zero loadings on risk factors, making the results less likely driven by risk. Specifically, each month, we sort all dividend-paying firms into quintile portfolios based on *NOPS*. The treated group consists of firms that report dividend record dates in the previous months, while the control group includes firms that do not report dividend record dates in the previous months, but have reported dividend record dates in the previous 12 months. Then we obtain the respective long-short portfolios for the treated and control groups. Finally, we regress the difference in the long-short portfolio returns between the treated and control groups on the post-JGTRRA dummy. The results are reported in Table B.3 of the Appendix. We find that the difference in the anomaly returns between the treated and control groups is significantly positive after the JGTRRA periods, although they are significantly negative before JGTRRA, consistent with our other

[17]Despite the above stated benefits of using this sample, there are also potential issues such as survivorship bias and too few observations. Dividend-paying stocks account for about 33% in our sample.

[18]The reversal after the dividend record month (Hartzmark and Solomon, 2013) may account for the negative β_0 coefficient for the long sides in both panels. In an unreported analysis, we obtain similar results after excluding stocks that suspend paying dividends after the JGTRRA.

results. When adding risk factors to our regressions, we indeed obtain insignificant loadings on them.

Overall, the results in this subsection further strengthen the causal effect of the differential tax-driven shock to short selling on mispricing and anomalies.

3.6 Anomaly Types

In this section, we investigate whether different types of anomalies respond differently to the dividend tax cut shock. We categorize the 182 anomalies based on four types in McLean and Pontiff (2016) and Engelberg et al. (2018): (1) event, (2) market, (3) fundamentals, and (4) valuation. Specifically, Event anomalies are based on corporate events and changes in performance, such as share issues and investment growth; market anomalies are constructed using only market data such as price momentum and idiosyncratic volatility; fundamentals anomalies are firm accounting attributes; valuation anomalies consist of accounting fundamentals scaled by market information, such as book-to-market and earnings-to-price ratios. In total, we have 52 event anomalies, 62 market anomalies, 51 fundamentals anomalies, and 17 valuation anomalies. In addition to these four types of anomalies, we consider the 11 anomalies in Stambaugh et al. (2012). We label this type of anomalies as *SYY*.

We first construct *NOPS* separately for each of these five groups of anomalies. Next, we re-run our portfolio-level DID regressions. Table 3.12 presents the results separately for each anomaly type. We find consistent results for all but *valuation* anomalies. The coefficients on $Treated \times JGTRRA$ are significantly positive for the long-short portfolios, confirming our hypothesis that anomalies respond to JGTRRA more strongly after the dividend record months compared with the other months, and are substantially negative for the short legs, whereas these coefficients are insignificant for the long legs. These results provide further support for our hypothesis that the effect mainly comes from the overpriced stocks.

Our findings that most anomalies are likely driven by mispricing are surprising and yet important. Mispricing is often associated with behavioral biases, and thus our findings highlight the prevalence of behavioral biases in stock markets. A few fundamentals anomalies such as accruals (Hirshleifer et al., 2012, 2011) and asset growth (Lam and Wei, 2011; Lipson et al., 2011) are argued to be related to mispricing in the prior literature. Additionally, Yan and Zheng (2017) find that many fundamental-based anomalies are stronger following higher sentiment periods and among stocks with greater limits to arbitrage. They argue that these anomalies are likely to be driven by mispricing rather than random chance or data mining. We provide more definite evidence supporting this argument. We also follow Hou et al. (2020) to categorize our 182 anomalies into six types: (1) momentum, (2) value & growth, (3) profitability, (4) investment, (5) intangibles, and (6) trading frictions. Our results remain strong for momentum, profitability, investment and intangibles, and marginally significant for trading frictions, but insignificant for value & growth anomalies.

Overall, we demonstrate that the causal effect of the tax-driven shock to short selling on mispricing is robust to various anomaly types except for *valuation* or *value & growth* anomalies, which are likely driven by risk. We emphasize this new result is not possible to obtain due to the small number of anomalies considered in Chu et al. (2020). Our result highlights the importance of using a comprehensive set of anomalies.

3.7 Conclusion

There are numerous studies on anomalies, but the causes of anomalies are still in debate. In this study, we investigate the causal effect of short selling on mispricing and anomalies using a robust and plausibly exogenous shock after the Job and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003.

Using 182 anomalies and the DID regression framework, we find that mispricing becomes stronger in the dividend record months after the JGTRRA, and as a result, anomalies are stronger after the dividend record months. Moreover, we show that the effect mainly comes from the overpriced stocks. Our findings are robust after controlling for various risk factors. Moreover, our various falsification tests indicate that data-mining is unlikely to drive our results. We further demonstrate that our results are stronger during high investor sentiment periods when overpricing is more prevalent, and in stocks that are more short-sale constrained.

We further divide anomalies into four types and examine each type separately. We find that the effect of the shock is significant in all but valuation anomalies, suggesting that most anomalies are driven by mispricing while valuation anomalies are likely driven by risk. Taking advantage of the unique setting of the JGTRRA shock, we also consider dividend stocks only, and find virtually the same results. Taken together, this study offers a novel test of the causal effect of short selling on mispricing and anomalies and provides solid evidence that anomalies mainly reflect mispricing.

Table 3.1: Summary statistics

Panel A describes summary statistics of $DivR_{i,t}$, which equals one if stock i reports a dividend record date in month t and zero otherwise. We restrict our sample to ordinary taxable cash dividends (CRSP distribution code = 1232) of $0.01 or greater that are paid by ordinary common shares listed on the NYSE/AMEX/NASDAQ. We exclude stocks with prior month prices below $5 per share. The sample period is 1985:7 to 2019:12.

Panel B provides descriptive statistics for the aggregated mispricing measure which is the average of those at the stock level. The net overpriced score (*NOPS*) for each stock is defined as *NShort* - *NLong*, where *NShort* (*NLong*) is the total number out of 182 anomalies that the stock is in the short (long) legs of the decile portfolios. The sample period is 1985:7 to 2019:12.

Panel A: Firm-month observations with dividend record dates

	DivR = 1	*DivR* = 0	Total
# of firm-month observations	225,631	1,362,850	1,588,481
Percentage	14.20%	85.80%	100%

Panel B: Summary statistics of *NOPS*

	Mean	Std.Dev	Min	p25	p50	p75	Max
NShort	12.54	8.94	0	6	10	16	77
NLong	14.84	7.88	0	9	14	19	66
NOPS	-2.29	10.12	-61	-8	-2	3	63

Table 3.2: Difference-in-differences results

This table reports results from the stock-level difference-in-differences regression, where the dependent variable $ret_{i,t}$ is the monthly stock return (in percentage). t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. *** denotes two-tail statistical significance at the 1% level.

Fixed Effects	Month	Month	Firm & Month	Firm & Month
S.E. Clusters	Month	Firm & Month	Month	Firm & Month
NOPS	-0.095***	-0.095***	-0.100***	-0.100***
	(-9.02)	(-9.00)	(-11.07)	(-11.12)
DivR	0.077	0.077	-0.273***	-0.273***
	(0.37)	(0.37)	(-2.78)	(-2.77)
NOPS×*DivR*	0.046***	0.046***	0.031***	0.031***
	(4.55)	(4.55)	(4.76)	(4.76)
NOPS×*JGTRRA*	0.055***	0.055***	0.062***	0.062***
	(4.56)	(4.55)	(5.76)	(5.79)
DivR×*JGTRRA*	-0.148	-0.148	0.081	0.081
	(-0.62)	(-0.62)	(0.43)	(0.43)
NOPS×*DivR*×*JGTRRA*	-0.035***	-0.035***	-0.028***	-0.028***
	(-2.83)	(-2.82)	(-2.89)	(-2.89)

Table 3.3: DID results after Regulation SHO

This table provides DID results for testing whether the causal effect of dividend taxation shock to short-selling on anomalies is robust after the Reg SHO. t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12 excluding the period 2003:6 to 2007:07. * and *** denote two-tail statistical significance at the 10% and 1% levels, respectively.

Fixed Effects	Month	Month	Firm & Month	Firm & Month
S.E. Clusters	Month	Firm & Month	Month	Firm & Month
NOPS	-0.089***	-0.089***	-0.098***	-0.098***
	(-12.88)	(-12.86)	(-17.22)	(-17.26)
DivR	-0.056	-0.056	-0.300***	-0.300***
	(-0.52)	(-0.52)	(-6.32)	(-6.19)
NOPS×*DivR*	0.031***	0.031***	0.021***	0.021***
	(5.33)	(5.29)	(4.86)	(4.82)
NOPS×*JGTRRA*	0.052***	0.052***	0.071***	0.071***
	(5.12)	(5.11)	(8.69)	(8.72)
DivR×*JGTRRA*	0.063	0.063	0.140	0.140
	(0.35)	(0.35)	(0.99)	(0.99)
NOPS×*DivR*×*JGTRRA*	-0.019*	-0.019*	-0.022***	-0.022***
	(-1.87)	(-1.86)	(-2.66)	(-2.64)

Table 3.4: Overpricing from the tax-driven shock to short selling

This table tests whether the mispricing effect comes from the long or short side. We create two dummy variables, *High NOPS* and *Low NOPS* , based on the decile rank of *NOPS* each month. We add firm and month fixed effects and cluster standard errors on both firm and time. t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. *, **, and *** denote two-tail statistical significance at the 10%, 5%, and 1% levels, respectively.

DivR	-0.927***	-0.682***	-0.654***
	(-7.53)	(-6.75)	(-5.92)
DivR×JGTRRA	0.642***	0.451***	0.435***
	(3.77)	(2.90)	(2.66)
NOPS DP	-0.284***	-0.232***	-0.210***
	(-10.44)	(-11.96)	(-9.94)
NOPS DP×DivR	0.108***	0.054***	0.049***
	(5.31)	(3.39)	(2.88)
NOPS DP×JGTRRA	0.178***	0.141***	0.129***
	(5.11)	(5.74)	(4.72)
NOPS DP×DivR×JGTRRA	-0.094***	-0.049**	-0.047*
	(-2.99)	(-1.99)	(-1.77)
Low NOPS	0.133		0.350***
	(1.60)		(4.93)
High NOPS		-1.171***	-1.238***
		(-6.99)	(-7.57)
Low NOPS×DivR	0.158		-0.045
	(1.41)		(-0.41)
Low NOPS×JGTRRA	-0.045		-0.189*
	(-0.35)		(-1.68)
High NOPS×DivR		0.857***	0.866***
		(4.08)	(4.12)
High NOPS×JGTRRA		0.742***	0.780***
		(3.55)	(3.88)
Low NOPS×DivR×JGTRRA	-0.132		0.025
	(-0.80)		(0.16)
High NOPS×DivR×JGTRRA		-0.738**	-0.744**
		(-2.30)	(-2.33)

Table 3.5: Controlling for dynamic risk factors

This table investigates whether the causal effect of dividend taxation shock to short-selling on anomalies comes from the short legs after controlling for dynamic risk factors. We consider the market factor (MKT), and five macroeconomic risk factors from Chen et al. (1986): the growth rate of industrial production (MP), unexpected inflation (UI), change in expected inflation (DEI), term premium (UTS), and default premium (UPR). We interact each source of risk with the dummies, *High* or *Low NOPS*, *JGTRRA*, *DivR* individually or in combinations. In the first six columns, we add one factor at a time. In the last column, we include all six risk factors. We add firm and month fixed effects and cluster standard errors on both firm and time. t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. *, **, and *** denote two-tail statistical significance at the 10%, 5%, and 1% levels, respectively.

	MKT	MP	UI	DEI	UTS	UPR	All
DivR	-0.504***	-0.738***	-0.635***	-0.664***	-0.670***	-0.649***	-0.505***
	(-3.93)	(-5.47)	(-5.68)	(-6.00)	(-6.01)	(-5.82)	(-3.52)
DivR×JGTRRA	0.398**	0.506***	0.380**	0.407**	0.413**	0.409**	0.331*
	(2.43)	(2.76)	(2.25)	(2.43)	(2.47)	(2.47)	(1.85)
NOPS DP	-0.209***	-0.210***	-0.210***	-0.210***	-0.210***	-0.210***	-0.210***
	(-9.91)	(-9.99)	(-9.98)	(-9.98)	(-9.98)	(-9.95)	(-9.94)
NOPS DP×DivR	0.052***	0.051***	0.050***	0.049***	0.051***	0.050***	0.053***
	(2.95)	(2.96)	(2.91)	(2.84)	(2.97)	(2.90)	(3.05)
NOPS DP×JGTRRA	0.128***	0.126***	0.121***	0.120***	0.121***	0.125***	0.119***
	(4.73)	(4.61)	(4.38)	(4.35)	(4.39)	(4.58)	(4.37)
NOPS DP×DivR×JGTRRA	-0.047*	-0.048*	-0.042	-0.042	-0.043	-0.047*	-0.044
	(-1.78)	(-1.80)	(-1.58)	(-1.55)	(-1.60)	(-1.78)	(-1.61)
Low NOPS	0.409***	0.315***	0.359***	0.352***	0.350***	0.352***	0.427***
	(5.01)	(4.12)	(5.03)	(4.95)	(4.97)	(4.88)	(5.18)
High NOPS	-1.532***	-1.168***	-1.244***	-1.240***	-1.234***	-1.239***	-1.734***
	(-7.16)	(-4.35)	(-7.47)	(-7.61)	(-7.57)	(-7.45)	(-6.76)
Low NOPS×DivR	-0.125	-0.014	-0.072	-0.037	-0.031	-0.051	-0.142
	(-1.10)	(-0.12)	(-0.68)	(-0.35)	(-0.29)	(-0.47)	(-1.15)
Low NOPS×JGTRRA	-0.213*	-0.163	-0.183	-0.176	-0.174	-0.182	-0.237*
	(-1.69)	(-1.40)	(-1.58)	(-1.53)	(-1.52)	(-1.59)	(-1.82)
High NOPS×DivR	0.949***	0.907***	0.891***	0.883***	0.857***	0.862***	1.175***
	(4.22)	(3.49)	(4.22)	(4.24)	(4.10)	(4.05)	(4.29)
High NOPS×JGTRRA	0.852***	0.748**	0.767***	0.769***	0.781***	0.764***	1.135***
	(3.66)	(2.55)	(3.67)	(3.76)	(3.85)	(3.71)	(4.12)
Low NOPS×DivR×JGTRRA	0.088	0.014	0.050	0.009	0.011	0.027	0.122
	(0.53)	(0.08)	(0.31)	(0.06)	(0.07)	(0.17)	(0.67)
High NOPS×DivR×JGTRRA	-0.814**	-0.941***	-0.841**	-0.828**	-0.794**	-0.760**	-1.242***
	(-2.41)	(-2.64)	(-2.57)	(-2.54)	(-2.50)	(-2.34)	(-3.49)

Table 3.6: Placebo tests with pseudo JGTRRA

This table reports various placebo tests where the timing of the tax code is arbitrarily changed. We use the timing before and after 2003 for pseudo enactment of JGTRRA including July of 1997 and 2013, and January of 1999 and 2005. Consequently, we use two sample periods: (1) between July 1985 and May 2003; (2) between June 2003 and December 2019. We consider $PseudoJGTRRA_t$, a dummy variable which equals one if month t is after the pseudo date specified, and zero otherwise. We add firm and month fixed effects and cluster standard errors on both firm and time. t-statistics are presented in parentheses below the coefficient estimates. *, **, and *** denote two-tail statistical significance at the 10%, 5%, and 1% levels, respectively.

Sample period	1985:7 - 2003:5		2003:6 - 2019:12	
	1997:7	*1999:1*	*2005:1*	*2013:7*
NOPS	-0.090***	-0.093***	-0.019	-0.034***
	(-12.48)	(-12.04)	(-1.19)	(-5.22)
DivR	-0.334***	-0.291**	-0.581*	-0.387***
	(-2.69)	(-2.35)	(-1.80)	(-4.38)
NOPS×*DivR*	0.013**	0.016**	-0.015	-0.000
	(2.10)	(2.41)	(-0.77)	(-0.03)
NOPS×*PseudoJGTRRA*	-0.050**	-0.058**	-0.013	0.011
	(-2.52)	(-2.30)	(-0.77)	(1.04)
DivR×*PseudoJGTRRA*	0.322	0.248	0.410	0.463***
	(0.87)	(0.47)	(1.13)	(2.76)
NOPS×*DivR*×*PseudoJGTRRA*	0.048***	0.055***	0.021	0.013
	(3.03)	(2.77)	(1.00)	(1.22)

Table 3.7: Placebo tests with pseudo dividend record months

This table reports placebo tests with pseudo dividend record dates. We consider two testing samples: (1) excluding only the dividend record months; (2) excluding dividend-paying stocks altogether. For each sample, in each month, we randomly choose 14% of firm-month observations (based on summary statistics in Table 1) to be stocks with dividend record dates over the previous month. $PseudoDivR_{i,t-1}$ is a dummy variable that equals one if stock i has a pseudo dividend record date in month $t-1$, and zero otherwise. *Simulation 1* and *Simulation 2* use two different simulation seeds. We add both firm and month fixed effects and cluster standard errors on both firm and time. t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. ** and *** denote two-tail statistical significance at the 5% and 1% levels, respectively.

Subsample	Drop if *DivR* = 1		Drop dividend stocks	
	Simulation 1	Simulation 2	Simulation 1	Simulation 2
NOPS	-0.051***	-0.053***	-0.113***	-0.113***
	(-8.66)	(-8.97)	(-11.58)	(-11.38)
PseudoDivR	-0.027	0.063	0.109	-0.080
	(-0.35)	(0.88)	(1.33)	(-1.13)
NOPS×*PseudoDivR*	-0.006	0.008	0.002	-0.001
	(-0.65)	(0.84)	(0.26)	(-0.14)
NOPS×*JGTRRA*	0.022**	0.022**	0.074***	0.073***
	(2.32)	(2.32)	(6.26)	(6.19)
PseudoDivR×*JGTRRA*	0.010	0.094	-0.092	0.184**
	(0.09)	(0.86)	(-0.89)	(1.97)
NOPS×*PseudoDivR*×*JGTRRA*	0.007	0.007	0.002	0.004
	(0.46)	(0.48)	(0.29)	(0.43)

Table 3.8: Portfolio-level DID

This table reports the portfolio-level DID results. For stocks that report dividend record dates in the previous month, we sort them into ten decile portfolios based on *NOPS* and then compute the monthly portfolio returns for the long, short, and long-short portfolios. We repeat the procedure for stocks that do not report dividend record dates in the previous month. Robust t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. ** and *** denote two-tail statistical significance at the 5% and 1% levels, respectively.

	Long-Short	Long side	Short side
Treated	-2.083***	-0.545***	1.539***
	(-6.21)	(-3.67)	(3.73)
Treated×*JGTRRA*	1.677***	0.393**	-1.284***
	(4.04)	(2.12)	(-2.68)
Constant	2.351***	1.476***	-0.875***
	(15.82)	(22.23)	(-5.06)

Table 3.9: Subperiods of investor sentiment

This table reports the main DID coefficients β_1 from subperiods regression with respect to investor sentiment. We use two orthogonalized investor sentiment indices from Baker and Wurgler (2006) and Huang et al. (2015), respectively, to identify the high or low sentiment periods. We first obtain the mean of each investor sentiment index using a recursive-window with at least twenty-four monthly observations. A high-sentiment month is the one in which the value of sentiment index at the end of previous month is above the mean value in the recursive-window, or vice versa. We re-run the DID regression for the two subperiods, respectively. Robust t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. *, **, and *** denote two-tail statistical significance at the 10%, 5%, and 1% levels, respectively.

	Baker and Wurgler (2006)		Huang et al. (2015)	
	High	Low	High	Low
β_1 (*Long-Short*)	3.195***	0.909**	3.424***	0.795**
	(3.48)	(2.20)	(3.45)	(2.14)
β_1 (*Long side*)	0.303	0.402*	0.860*	0.174
	(0.80)	(1.90)	(1.92)	(0.92)
β_1 (*Short side*)	-2.892***	-0.507	-2.565**	-0.621
	(-2.77)	(-1.12)	(-2.14)	(-1.64)

Table 3.10: Subsamples of limits to arbitrage proxies

This table reports the main DID coefficients β_1 from subsample regression with respect to different limits to arbitrage proxies, including firm size, firm age, idiosyncratic volatility (IVOL), and size-adjusted institutional ownership (IO). For each of limits to arbitrage proxies, we define high or low groups based on tercile portfolios. Robust t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. *, **, and *** denote two-tail statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Firm size and firm age				
	Large Firm	Small Firm	Mature Firm	Young Firm
β_1 (*Long-Short*)	0.833**	1.510***	0.729***	1.840***
	(2.54)	(3.78)	(3.36)	(3.48)
β_1 (*Long side*)	0.289	0.211	0.231	0.381
	(1.55)	(0.96)	(1.38)	(1.30)
β_1 (*Short side*)	-0.543	-1.299***	-0.497**	-1.459**
	(-1.35)	(-3.06)	(-2.09)	(-2.50)
Panel B: IVOL and IO				
	High IVOL	Low IVOL	High IO	Low IO
β_1 (*Long-Short*)	1.594***	0.502***	0.981***	1.378***
	(2.86)	(2.80)	(2.60)	(3.40)
β_1 (*Long side*)	0.520	0.143	0.099	0.376
	(1.47)	(1.14)	(0.43)	(1.63)
β_1 (*Short side*)	-1.073*	-0.359**	-0.882**	-1.002**
	(-1.86)	(-2.14)	(-2.33)	(-2.01)

Table 3.11: Dividend stocks only

This table investigates the causal effect of dividend taxation shock to short-selling on anomalies for dividend-paying stocks. In Panel A, we exclude non-dividend-paying stocks during the entire sample period. In Panel B, we exclude both non-dividend-paying stocks and stocks that initiated dividends after September 2003. We choose September of 2003 because we assume that it might take three months for a firm to react to the act. Robust t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. *, **, and *** denote two-tail statistical significance at the 10%, 5%, and 1% levels, respectively.

	Long-Short	Long side	Short side
Panel A: Dividend-paying stocks only			
Treated	-0.715***	-0.487***	0.228
	(-4.00)	(-4.57)	(1.30)
Treated×JGTRRA	0.753***	0.203	-0.550**
	(3.06)	(1.33)	(-2.50)
Constant	1.423***	1.509***	0.085
	(16.32)	(27.94)	(1.08)
Panel B: Dividend-paying stocks before JGTRRA only			
Treated	-0.470***	-0.348***	0.122
	(-3.05)	(-3.62)	(0.97)
Treated×JGTRRA	0.517**	0.003	-0.514**
	(2.01)	(0.02)	(-2.54)
Constant	1.309***	1.430***	0.121*
	(14.52)	(25.69)	(1.71)

Table 3.12: Types of anomalies

This table investigates the different types of anomalies. We split the 182 significant anomalies into four groups based on McLean and Pontiff (2016); Engelberg et al. (2018): (1) Event, (2) Market, (3) Fundamentals, and (4) Valuation. We separately compute *NOPS* for each of these four types of anomalies. We also consider *NOPS* constructed from Stambaugh et al. (2012) 11 anomalies, labeled as *NOPS* SYY. Robust t-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2019:12. ** and *** denote two-tail statistical significance at the 5% and 1% levels, respectively.

NOPS by Type:	SYY	Event	Market	Fundamentals	Valuation
	11	52	62	51	17
Panel A: Long-Short					
Treated	-1.589***	-1.085***	-1.964***	-1.924***	-0.693
	(-6.29)	(-5.36)	(-6.00)	(-6.14)	(-1.56)
Treated×JGTRRA	1.258***	1.092***	1.064***	1.440***	0.296
	(3.89)	(4.13)	(2.81)	(3.60)	(0.55)
Constant	1.617***	1.128***	2.194***	1.605***	1.206***
	(14.00)	(11.97)	(16.03)	(11.23)	(6.26)
Panel B: Long side					
Treated	-0.459**	-0.215	-0.423***	-0.498***	-0.110
	(-2.44)	(-1.01)	(-2.60)	(-2.69)	(-1.03)
Treated×JGTRRA	0.342	0.221	0.024	0.356	0.001
	(1.60)	(0.88)	(0.12)	(1.62)	(0.01)
Constant	1.133***	1.057***	1.327***	1.215***	1.095***
	(14.62)	(11.68)	(18.62)	(15.37)	(18.23)
Panel C: Short side					
Treated	1.130***	0.870***	1.541***	1.427***	0.583
	(3.33)	(3.13)	(3.85)	(3.38)	(1.19)
Treated×JGTRRA	-0.916**	-0.871***	-1.040**	-1.085**	-0.295
	(-2.26)	(-2.64)	(-2.31)	(-2.17)	(-0.52)
Constant	-0.483***	-0.072	-0.866***	-0.390**	-0.111
	(-3.31)	(-0.60)	(-5.31)	(-2.17)	(-0.54)

REFERENCES

Ali, A., Hwang, L.-S., and Trombley, M. A. (2003). Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics*, 69(2):355–373.

Ang, A., Piazzesi, M., and Wei, M. (2006). What does the yield curve tell us about gdp growth? *Journal of Econometrics*, 131(1-2):359–403.

Aruoba, S. B. and Diebold, F. X. (2010). Real-time macroeconomic monitoring: Real activity, inflation, and interactions. *American Economic Review*, 100(2):20–24.

Avramov, D., Kaplanski, G., and Subrahmanyam, A. (2021). Moving average distance as a predictor of equity returns. *Review of Financial Economics*, 39(2):127–145.

Babaoğlu, K., Christoffersen, P., Heston, S., and Jacobs, K. (2018). Option valuation with volatility components, fat tails, and nonmonotonic pricing kernels. *The Review of Asset Pricing Studies*, 8(2):183–231.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4):1645–1680.

Bakshi, G., Crosby, J., Gao, X., and Zhou, W. (2021). Negative correlation condition and the dark matter property of asset pricing models. *Working paper*.

Bakshi, G., Gao, X., and Xue, J. (2022). Recovery with applications to forecasting equity disaster probability and testing the spanning hypothesis in the treasury market. *Journal of Financial and Quantitative Analysis*, forthcoming.

Bakshi, G., Kapadia, N., and Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16(1):101–143.

Bakshi, G. and Madan, D. (2000). Spanning and derivative-security valuation. *Journal of Financial Economics*, 55(2):205–238.

Bali, T. G., Engle, R. F., and Murray, S. (2016). *Empirical Asset Pricing: The Cross Section of Stock Returns*. John Wiley & Sons, Hoboken, New Jersey.

Ball, R., Gerakos, J., Linnainmaa, J. T., and Nikolaev, V. (2020). Earnings, retained earnings, and book-to-market in the cross section of expected returns. *Journal of Financial Economics*, 135(1):231–254.

Baltussen, G., van Bekkum, S., and Da, Z. (2019). Indexing and stock market serial dependence around the world. *Journal of Financial Economics*, 132(1):26–48.

Bansal, R., Miller, S., Song, D., and Yaron, A. (2021). The term structure of equity risk premia. *Journal of Financial Economics*, 142(3):1209–1228.

Barberis, N. and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128.

Bates, J. M. and Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.

Bauer, M. D. and Rudebusch, G. D. (2020). Interest rates under falling stars. *American Economic Review*, 110(5):1316–1354.

Bekaert, G., Cho, S., and Moreno, A. (2010). New keynesian macroeconomics and the term structure. *Journal of Money, Credit and Banking*, 42(1):33–62.

Berk, J. B., Green, R. C., and Naik, V. (1999). Optimal investment, growth options, and security returns. *The Journal of Finance*, 54(5):1553–1607.

Binsbergen, J. H. v., Brandt, M., and Koije, R. (2012). On the timing and pricing of dividends. *American Economic Review*, 102:1596–1618.

Binsbergen, J. H. V. and Koijen, R. S. (2017). The term structure of returns: Facts and theory. *Journal of Financial Economics*, 124(1):1–21.

Blocher, J., Reed, A. V., and Van Wesep, E. D. (2013). Connecting two markets: An equilibrium framework for shorts, longs, and stock loans. *Journal of Financial Economics*, 108(2):302–322.

Borovička, J., Hansen, L. P., and Scheinkman, J. A. (2016). Misspecified recovery. *The Journal of Finance*, 71(6):2493–2544.

Bowles, B., Reed, A. V., Ringgenberg, M. C., and Thornock, J. R. (2017). Anomaly time. *Working paper*.

Brav, A., Graham, J. R., Harvey, C. R., and Michaely, R. (2008). Managerial response to the May 2003 dividend tax cut. *Financial Management*, 37(4):611–624.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics*, 18(2):373–399.

Campbell, J. Y. (2017). *Financial decisions and markets: a course in asset pricing*. Princeton University Press.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Campbell, J. Y. and Vuolteenaho, T. (2004). Bad beta, good beta. *American Economic Review*, 94(5):1249–1275.

Carr, P., Ellis, K., and Gupta, V. (1998). Static hedging of exotic options. *The Journal of Finance*, 53(3).

Carr, P. and Laurence, P. (2011). Multi-asset stochastic local variance contracts. *Mathematical Finance*, 21(1):21–52.

Carr, P. and Madan, D. (1999). Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2(4):61–73.

Chabi-Yo, F. (2019). What is the conditional autocorrelation on the stock market? *Working paper*.

Chabi-Yo, F. and Loudis, J. (2020). The conditional expected market return. *Journal of Financial Economics*, 137(3):752–786.

Chen, A. Y. (2021). The limits of p-hacking: Some thought experiments. *The Journal of Finance*, 76(5):2447–2480.

Chen, A. Y. and Zimmermann, T. (2021). Open source cross-sectional asset pricing. *Critical Finance Review,*, forthcoming.

Chen, N.-F., Roll, R., and Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, 59(3):383–403.

Chetty, R. and Saez, E. (2005). Dividend taxes and corporate behavior: Evidence from the 2003 dividend tax cut. *The Quarterly Journal of Economics*, 120(3):791–833.

Chordia, T., Subrahmanyam, A., and Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, 58:41–58.

Christoffersen, P., Heston, S., and Jacobs, K. (2013). Capturing option anomalies with a variance-dependent pricing kernel. *The Review of Financial Studies*, 26(8):1963–2006.

Chu, Y., Hirshleifer, D., and Ma, L. (2020). The causal effect of limits to arbitrage on asset pricing anomalies. *The Journal of Finance*, 75(5):2631–2672.

Cieslak, A. and Povala, P. (2015). Expected returns in treasury bonds. *The Review of Financial Studies*, 28(10):2859–2901.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cochrane, J. H. (1991). Production-based asset pricing and the link between stock returns and economic fluctuations. *The Journal of Finance*, 46(1):209–237.

Cochrane, J. H. and Saa-Requejo, J. (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy*, 108(1):79–119.

Cong, L. W., Tang, K., Wang, J., and Zhang, Y. (2021). Deep sequence modeling: Development and applications in asset pricing. *The Journal of Financial Data Science*, 3(1):28–42.

De la O, R. and Myers, S. (2022). Which subjective expectations explain asset prices? *Working paper*.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial studies*, 22(5):1915–1953.

Dixon, P. N., Fox, C. A., and Kelley, E. K. (2021). To own or not to own: Stock loans around dividend payments. *Journal of Financial Economics*, 140(2):539–559.

Dong, X., Li, Y., Rapach, D. E., and Zhou, G. (2022). Anomalies and the expected market return. *The Journal of Finance*, 77(1):639–681.

Dow, C. H. (1920). *Scientific stock speculation*. Magazine of Wall Street.

Duffie, D. and Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *The Review of Financial Studies*, 12(4):687–720.

Engelberg, J., McLean, R. D., and Pontiff, J. (2018). Anomalies and news. *The Journal of Finance*, 73(5):1971–2001.

Fama, E. F. and French, K. R. (1988a). Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1):3–25.

Fama, E. F. and French, K. R. (1988b). Permanent and temporary components of stock prices. *Journal of Political Economy*, 96(2):246–273.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The Journal of Finance*, 50(1):131–155.

Fama, E. F. and French, K. R. (1998). Value versus growth: The international evidence. *The Journal of Finance*, 53(6):1975–1999.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fama, E. F. and Schwert, G. W. (1977). Asset returns and inflation. *Journal of Financial Economics*, 5(2):115–146.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, 56(1):329–352.

Gao, L., Han, Y., Li, S. Z., and Zhou, G. (2018). Market intraday momentum. *Journal of Financial Economics*, 129(2):394–414.

Gerakos, J. and Linnainmaa, J. T. (2018). Decomposing value. *The Review of Financial Studies*, 31(5):1825–1854.

Gormsen, N. J. (2021). Time variation of the equity term structure. *The Journal of Finance*, 76(4):1959–1999.

Goyal, A. and Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.

Goyal, A., Welch, I., and Zafirov, A. (2021). A comprehensive look at the empirical performance of equity premium prediction ii. *Working paper*.

Goyenko, R. and Zhang, C. (2020). The joint cross section of option and stock returns predictability with big data and machine learning. *Working paper*.

Green, J., Hand, J. R., and Zhang, X. F. (2017). The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies*, 30(12):4389–4436.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273.

Han, Y., Zhou, G., and Zhu, Y. (2016). A trend factor: Any economic gains from using information over investment horizons? *Journal of Financial Economics*, 122(2):352–375.

Hartzmark, S. M. and Solomon, D. H. (2013). The dividend month premium. *Journal of Financial Economics*, 109:640–660.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *The Review of Financial Studies*, 29(1):5–68.

Harvey, D. I., Leybourne, S. J., and Newbold, P. (1998). Tests for forecast encompassing. *Journal of Business & Economic Statistics*, 16(2):254–259.

Haugen, R. A. and Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3):401–439.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, pages 77–105.

Heath, D., Ringgenberg, M., Samadi, M., and Werner, I. M. (2022). Reusing natural experiments. *Journal of Finance*, forthcoming.

Heston, S. L. (2021). Recovering the variance premium. *Working paper*.

Hirshleifer, D., Hou, K., and Teoh, S. H. (2012). The accrual anomaly: risk or mispricing? *Management Science*, 58(2):320–335.

Hirshleifer, D., Teoh, S. H., and Yu, J. J. (2011). Short arbitrage, return asymmetry, and the accrual anomaly. *The Review of Financial Studies*, 24(7):2429–2461.

Hong, H. and Stein, J. C. (2007). Disagreement and the stock market. *Journal of Economic Perspectives*, 21(2):109–128.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705.

Hou, K., Xue, C., and Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies*, 33(5):2019–2133.

Hu, G. and Jacobs, K. (2020). Expected and realized returns on volatility. *Working paper*.

Huang, D., Jiang, F., Li, K., Tong, G., and Zhou, G. (2021). Scaled pca: A new approach to dimension reduction. *Management Science*, pages 1–18.

Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, 28(3):791–837.

Huang, D., Zhang, H., Zhou, G., and Zhu, Y. (2020). Fundamental extrapolation and stock returns. *Working paper*.

Israel, R. and Moskowitz, T. J. (2013). The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics*, 108(2):275–301.

Jensen, C. S., Lando, D., and Pedersen, L. H. (2019). Generalized recovery. *Journal of Financial Economics*, 133(1):154–174.

Jobson, J. D. and Korkie, B. M. (1981). Performance hypothesis testing with the sharpe and treynor measures. *Journal of Finance*, pages 889–908.

Kadan, O. and Tang, X. (2020). A bound on expected stock returns. *Review of Financial Studies*, 33(4):1565–1617.

Keim, D. B. and Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, 17(2):357–390.

Kelly, B. and Pruitt, S. (2013). Market expectations in the cross-section of present values. *The Journal of Finance*, 68(5):1721–1756.

Kelly, B. and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2):294–316.

Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. *3rd International Conference for Learning Representations, Conference paper*.

Kothari, S. P. and Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics*, 44(2):169–203.

Kozicki, S. and Tinsley, P. A. (2001). Shifting endpoints in the term structure of interest rates. *Journal of Monetary Economics*, 47(3):613–652.

Kremens, L. and Martin, I. (2019). The quanto theory of exchange rates. *American Economic Review*, 109(3):810–843.

Lam, F. E. C. and Wei, K. J. (2011). Limits-to-arbitrage, investment frictions, and the asset growth anomaly. *Journal of Financial Economics*, 102(1):127–149.

LeRoy, S. F. (1973). Risk aversion and the martingale property of stock prices. *International Economic Review*, pages 436–446.

Lettau, M. and Ludvigson, S. (2001). Resurrecting the (c) capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109(6):1238–1287.

Lewellen, J. (2010). Accounting anomalies and fundamental analysis: An alternative view. *Journal of Accounting and Economics*, 50(2-3):455–466.

Linnainmaa, J. T. and Roberts, M. R. (2018). The history of the cross-section of stock returns. *The Review of Financial Studies*, 31(7):2606–2649.

Lipson, M. L., Mortal, S., and Schill, M. J. (2011). On the scope and drivers of the asset growth effect. *Journal of Financial and Quantitative Analysis*, 46(6):1651–1682.

Liu, L. X. and Zhang, L. (2008). Momentum profits, factor pricing, and macroeconomic risk. *The Review of Financial Studies*, 21(6):2417–2448.

Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1):41–66.

Lo, A. W. and MacKinlay, A. C. (1990). An econometric analysis of nonsynchronous trading. *Journal of Econometrics*, 45(1-2):181–211.

Martin, I. (2013). Consumption-based asset pricing with higher cumulants. *Review of Economic Studies*, 80(2):745–773.

Martin, I. (2017). What is the expected return on the market? *The Quarterly Journal of Economics*, 132(1):367–433.

Martin, I. (2021). On the autocorrelation of the stock market. *Journal of Financial Econometrics*, 19(1):39–52.

Martin, I. W. and Wagner, C. (2019). What is the expected return on a stock? *The Journal of Finance*, 74(4):1887–1929.

McLean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1):5–32.

Mencia, J. and Sentana, E. (2013). Valuation of vix derivatives. *Journal of Financial Economics*, 108(2):367–391.

Merton, R. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–87.

Miller, M. H. (1999). The history of finance. *The Journal of Portfolio Management*, 25(4):95–101.

Moskowitz, T. J., Ooi, Y. H., and Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, 104(2):228–250.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics*, 78(2):277–309.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science*, 60(7):1772–1791.

Nelson, C. R. (1976). Inflation and rates of return on common stocks. *The Journal of Finance*, 31(2):471–483.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation. *Econometrica*, 55(3):703–708.

Pesaran, M. H. and Timmermann, A. (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics*, 137(1):134–161.

Pettenuzzo, D., Timmermann, A., and Valkanov, R. (2014). Forecasting stock returns under economic constraints. *Journal of Financial Economics*, 114(3):517–553.

Pontiff, J. and Schall, L. D. (1998). Book-to-market ratios as predictors of market returns. *Journal of Financial Economics*, 49(2):141–160.

Porta, R. L., Lakonishok, J., Shleifer, A., and Vishny, R. (1997). Good news for value stocks: Further evidence on market efficiency. *The Journal of Finance*, 52(2):859–874.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, 22(1):27–59.

Rapach, D. and Zhou, G. (2022). Asset pricing: Time-series predictability. *Oxford Research Encyclopedia of Economics and Finance*, forthcoming.

Rapach, D. E., Ringgenberg, M. C., and Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics*, 121(1):46–65.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2):821–862.

Ross, S. A. (2015). The recovery theorem. *The Journal of Finance*, 70(2):615–648.

Rozeff, M. S. (1984). Dividend yields are equity risk premiums. *Journal of Portfolio Management*, pages 68–75.

Rudebusch, G. D. and Wu, T. (2008). A macro-finance model of the term structure, monetary policy and the economy. *The Economic Journal*, 118(530):906–926.

Shleifer, A. and Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 52(1):35–55.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2):288–302.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle. *The Journal of Finance*, 70(5):1903–1948.

Stock, J. and Watson, M. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61(4):783–820.

Stock, J. H. and Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44(2):293–335.

Stock, J. H. and Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41(3):788–829.

Stock, J. H. and Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6):405–430.

Stock, J. H. and Watson, M. W. (2006). Forecasting with many predictors. *Handbook of Economic Forecasting*, 1:515–554.

Theil, H. (1966). *Applied economic forecasting*. Amsterdam: North-Holland.

Thornock, J. (2013). The effects of dividend taxation on short selling and market quality. *The Accounting Review*, 88(5):1833–1856.

Tian, W. (2014). Spanning with indexes. *Journal of Mathematical Economics*, 53:111–118.

Tian, W. (2019). The financial market: not as big as you think. *Mathematics and Financial Economics*, 13(1):67–85.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Yan, X. and Zheng, L. (2017). Fundamental analysis and the cross-section of stock returns: A data-mining approach. *The Review of Financial Studies*, 30(4):1382–1423.

Yang, P. R. (2020). Using the yield curve to forecast economic growth. *Journal of Forecasting*, 39(7):1057–1080.

Zhang, L. (2005). The value premium. *The Journal of Finance*, 60(1):67–103.

Zhu, Y. and Zhou, G. (2009). Technical analysis: An asset allocation perspective on the use of moving averages. *Journal of Financial Economics*, 92(3):519–544.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

APPENDIX A: EQUITY FORWARD RETURN FROM DERIVATIVES

A.1 Proof of Proposition 2.2.1

The proof is divided into several steps.

Step 1. We first derive an approximation formula of VIX as follows

$$VIX^2_{t\to t+T} \sim \frac{1}{T}\left(\mathbb{E}^Q_t\left[\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right)^2\right]-1\right). \tag{A1}$$

By using the second-order expansion of $log(1+x)\sim x-\frac{1}{2}x^2$ when x closes to zero, and $\frac{R_{t\to t+T}}{R_{f,t\to t+T}}$ sufficiently closes to one, we obtain

$$log\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right)\sim\frac{R_{t\to t+T}}{R_{f,t\to t+T}}-1-\frac{1}{2}\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}-1\right)^2. \tag{A2}$$

By taking the conditional expectation under the risk-neutral probability measure $\mathbb{Q}$, and using the relation that $\mathbb{E}^{\mathbb{Q}}_t\left[\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right]=1$, we obtain

$$\mathbb{E}^{\mathbb{Q}}_t\log\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right)\sim\frac{1}{2}-\frac{1}{2}\mathbb{E}^{\mathbb{Q}}_t\left[\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right)^2\right]. \tag{A3}$$

Recall the definition of VIX as a risk-neutral entropy

$$VIX^2_{t\to t+T}=\frac{2}{T}L^{\mathbb{Q}}_t\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right) \tag{A4}$$

where $L^{\mathbb{Q}}_t(X)\equiv\log\mathbb{E}^{\mathbb{Q}}_tX-\mathbb{E}^{\mathbb{Q}}_t\log X$. By Equation (A3), we obtain

$$VIX^2_{t\to t+T} \sim \frac{1}{T}\left(\mathbb{E}^{\mathbb{Q}}_t\left[\left(\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right)^2\right]-1\right), \tag{A5}$$

since $\log\mathbb{E}^{\mathbb{Q}}_t\left[\frac{R_{t\to t+T}}{R_{f,t\to t+T}}\right]=0$.

Step 2. We derive the result for $\mathbb{E}^{\mathbb{Q}}_t[R^n]$ when $n=2$. We use the formula (A1) for the time period from $t+T_1$ to $t+T_1+T_2$,

$$VIX^2_{t+T_1\to t+T_1+T_2} \sim \frac{1}{T_2}\left(\mathbb{E}^{\mathbb{Q}}_{t+T_1}\left[\left(\frac{R}{R_f}\right)^2\right]-1\right).$$

Here, to simplify notation, we write $R = R_{t+T_1\to t+T_1+T_2}, R_f = R_{f,t+T_1\to t+T_1+T_2}$.

By applying the conditional expectation of the last equation at time t under the $\mathbb{Q}$-measure , we have

$$\mathbb{E}^{\mathbb{Q}}_t[VIX^2_{t+T_1\to t+T_1+T_2}] \sim \frac{1}{T_2}\left(\mathbb{E}^{\mathbb{Q}}_{t}\left[\left(\frac{R}{R_f}\right)^2\right]-1\right), \tag{A6}$$

where the left-hand side can be expressed as

$$\mathbb{E}^{\mathbb{Q}}_t[VIX^2_{t+T_1\to t+T_1+T_2}] = \left(\underbrace{\mathbb{E}^{\mathbb{Q}}_t[VIX_{t+T_1\to t+T_1+T_2}]}_{FVIX_{t,t+T_1\to t+T_1+T_2}}\right)^2 + Var^{\mathbb{Q}}_t(VIX_{t+T_1\to t+T_1+T_2}).$$

Here, the first term on the right-hand side of the last equation is the square of the VIX future by the risk-neutral pricing formula, and the second term is the conditional variance $Var^{\mathbb{Q}}_t(VIX_{t+T_1\to t+T_1+T_2})$.

We now consider the VIX option with maturity $t+T_1$ and the underlying is $VIX_{t+T_1\to t+T_1+T_2}$. Since the VIX is a tradable asset, by the fundamental pricing theorem in derivative theory, its future value process under the $\mathbb{Q}$-measure is a martingale. Then, the conditional variance $Var^{\mathbb{Q}}_t(VIX_{t+T_1\to t+T_1+T_2})$ equals $\left(FVIX^2_{t,t+T_1\to t+T_1+T_2}\right)\sigma_t^2 T_1$, where σ_t is the implied volatility of the at-the-money VIX option. Therefore,

$$\mathbb{E}^{\mathbb{Q}}_t[VIX^2_{t+T_1\to t+T_1+T_2}] = FVIX^2_{t,t+T_1\to t+T_1+T_2}(1+\sigma_t^2 T_1). \tag{A7}$$

Plug back into Equation (A6) and we obtain

$$F_t^2(1+\sigma_t^2 T_1) \sim \frac{1}{T_2}\left(\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{R}{R_f}\right)^2\right]-1\right), \tag{A8}$$

where $F_t = FVIX_{t,t+T_1\to t+T_1+T_2}$ denotes the futures prices on VIX index.

Step 3. We derive the result for $\mathbb{E}_t^{\mathbb{Q}}[R^n]$ for $n \geq 3$. By the n-th order approximation of $\log(1+x)$, we have

$$\log(1+x) \sim \sum_{i=1}^{n}(-1)^{i-1}\frac{1}{i}x^i.$$

For $x = \frac{R}{R_f} - 1$, we obtain

$$\log\left(\frac{R}{R_f}\right) \sim \sum_{i=1}^{n}(-1)^{i-1}\frac{1}{i}\left(\frac{R}{R_f}-1\right)^i.$$

By using the same method in Step 2, we have

$$\mathbb{E}_{t+T_1}^{\mathbb{Q}}\left[\log(\frac{R}{R_f})\right] \sim \sum_{i=1}^{n}(-1)^{i-1}\frac{1}{i}\mathbb{E}_{t+T_1}^{\mathbb{Q}}\left[\left(\frac{R}{R_f}-1\right)^i\right].$$

Since

$$VIX_{t+T_1\to t+T_1+T_2}^2 = -\frac{2}{T_2}\mathbb{E}_{t+T_1}^{\mathbb{Q}}\left[\log(\frac{R}{R_f})\right] \sim \frac{2}{T_2}\sum_{i=1}^{n}(-1)^i\frac{1}{i}\mathbb{E}_{t+T_1}^{\mathbb{Q}}\left[\left(\frac{R}{R_f}-1\right)^i\right],$$

By taking expectation conditional on t, the iterated law of expectation implies

$$\mathbb{E}_t[VIX_{t+T_1\to t+T_1+T_2}^2] \sim \frac{2}{T_2}\sum_{i=1}^{n}(-1)^i\frac{1}{i}\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{R}{R_f}-1\right)^i\right],$$

and we obtain,

$$\frac{T_2}{2}F_t^2\left(1+\sigma^2 T_1\right) \sim \sum_{i=1}^{n}(-1)^i\frac{1}{i}\mathbb{E}_t^{\mathbb{Q}}\left[\left(\frac{R}{R_f}-1\right)^i\right], \quad n \geq 3. \tag{A9}$$

□

Approximation error: We next explain why this approximation is sufficiently tight for empirical applications. For simplicity, we use $x = \frac{R_{t,t\to t+T}}{R_{ft,t\to t+T}} - 1$. Let $a \equiv \sup_x |\log(1+x) - (x - \frac{x^2}{2})|$ for all possible scenarios of x. The number a is very small in magnitude because x is closes to zero. Moreover, for any $c > 0$,

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\left|\log(1+x) - (x - \frac{x^2}{2})\right|\right] &= \mathbb{E}^{\mathbb{Q}}\left[\left|\log(1+x) - (x - \frac{x^2}{2})\right| : |x| \leq c\right] \\
&\quad + \mathbb{E}^{\mathbb{Q}}\left[\left|\log(1+x) - (x - \frac{x^2}{2})\right| : |x| > c\right] \\
&\leq \frac{c^3}{3} + \mathbb{E}^{\mathbb{Q}}\left[\left|\log(1+x) - (x - \frac{x^2}{2})\right| : |x| > c\right] \\
&\leq \frac{c^3}{3} + aP(|x| > c).
\end{aligned}$$

Clearly, the smaller the parameter c, the smaller the first term $\frac{c^3}{3}$. Although the probability $P(|x| \geq c)$ can become larger given a smaller value of c, this probability itself is usually very small. In total, the upper bound of $\mathbb{E}^{\mathbb{Q}}\left[|\log(1+x) - (x - \frac{x^2}{2})|\right]$ is very small.

Numerically, if choose $|x| \leq 1\%$ for the monthly return (annual return bound is 12 percent), and the average VIX is 15%, then

$$\mathbb{E}_t^{\mathbb{Q}}\left[\left|\log(1+x) - (x - \frac{x^2}{2})\right|\right] \leq \frac{1}{3}(0.01)^3,$$

and

$$\left|\mathbb{E}_t^{\mathbb{Q}}[\log(1+x)]\right| = \frac{T}{2} VIX^2 = \frac{1}{2 \times 12}(0.15)^2.$$

Therefore,

$$\left|\frac{\mathbb{E}_t^{\mathbb{Q}}\left[\left|\log(1+x) - (x - \frac{x^2}{2})\right|\right]}{\mathbb{E}_t^{Q}[\log(1+x)]}\right| \leq \frac{1}{3}(0.01)^3 \times (2 \times 12)\frac{1}{0.15^2} = 0.04\%.$$

If we choose a large number for the month return, $|x| \leq 2\%$, which means the annual return is bounded between $[-24\%, 24\%]$, and VIX = 20%, then

$$\left| \frac{\mathbb{E}_t^{\mathbb{Q}}\left[\left|\log(1+x) - (x - \frac{x^2}{2})\right|\right]}{\mathbb{E}_t^{\mathbb{Q}}[\log(1+x)]} \right| \leq 0.16\%.$$

Therefore, the approximation formula is sufficiently accurate for the market data.

A.2 Proof of Proposition 2.4.1

Before proving Proposition 2.4.1, we prove two results first. The first one presents an alternative expression of forward return in terms of expected value of future options' values. The second one is on a relationship between option gamma and strike gamma for a general option.

Proposition A.2.1 *Suppose that interest rates are deterministic. Then*

$$\mathbb{E}_t[R_{t+1\to t+2}] = \frac{2}{R_{f,t\to t+1} S_t} \int_0^\infty \mathbb{E}_t^{\mathbb{Q}}\left[\frac{C_{t+1}(S_{t+1}, K)}{S_{t+1}}\right] dK, \tag{A10}$$

where

- S_t = *underlying index price observed at time* t;
- C_{t+1} = *the call option price at time* $t+1$.

Proof: Suppose that interest rates are deterministic. By Equation (2.1), the expected future return under the real-world probability measure $\mathbb{P}$ can be written as

$$\mathbb{E}_t[R_{t+1\to t+2}] = \frac{1}{R_{f,t\to t+2}} \mathbb{E}_t^{\mathbb{Q}}\left[(R_{t+1\to t+2})^2 \times R_{t\to t+1}\right], \tag{A11}$$

$$= \frac{1}{R_{f,t\to t+2}} \mathbb{E}_t^{\mathbb{Q}}\left\{\mathbb{E}_{t+1}^{Q}\left[(R_{t+1\to t+2})^2\right] \times R_{t\to t+1}\right\}, \tag{A12}$$

where

$$\mathbb{E}^{\mathbb{Q}}_{t+1}\left[(R_{t+1\to t+2})^2\right] = \mathbb{E}^{\mathbb{Q}}_{t+1}\left[\left(\frac{S_{t+2}}{S_{t+1}}\right)^2\right] = \frac{1}{S^2_{t+1}}\mathbb{E}^{\mathbb{Q}}_{t+1}\left[S^2_{t+2}\right]. \tag{A13}$$

Plug back into the equation and we have

$$\mathbb{E}_t[R_{t+1\to t+2}] = \left(\frac{1}{R_{f,t\to t+2}}\right)\left(\frac{1}{S_t}\right)\mathbb{E}^{Q}_t\left\{\frac{1}{S_{t+1}}\mathbb{E}^{Q}_{t+1}\left[S^2_{t+2}\right]\right\}. \tag{A14}$$

We use Equation (2.3) for $\mathbb{E}^{Q}_{t+1}\left[S^2_{t+2}\right]$, obtaining

$$\frac{1}{R_{f,t+1\to t+2}}\mathbb{E}^{Q}_{t+1}\left[S^2_{t+2}\right] = 2\int_0^\infty \frac{1}{R_{f,t+1\to t+2}}\mathbb{E}^{Q}_{t+1}\left[(S_{t+2}-K)^+\right]dK \tag{A15}$$

$$= 2\int_0^\infty C_{t+1}(S_{t+1},K)dK, \tag{A16}$$

where $C_{t+1}(S_{t+1},K)$ denotes the price of a call option at time $t+1$ that will expire at time $t+2$ with a strike price K.

Then, by Fubini's theorem, we have,

$$\mathbb{E}_t[R_{t+1\to t+2}] = \left(\frac{1}{R_{f,t\to t+2}}\right)\left(\frac{1}{S_t}\right)\mathbb{E}^{Q}_t\left\{\frac{1}{S_{t+1}}\left[2R_{f,t+1\to t+2}\int_0^\infty C_{t+1}(S_{t+1},K)dK\right]\right\}$$

$$= \frac{2}{R_{f,t\to t+1}S_t}\int_0^\infty \mathbb{E}^{Q}_t\left[\frac{C_{t+1}(S_{t+1},K)}{S_{t+1}}\right]dK. \tag{A17}$$

□

The following result could be known as folklore in derivative literature. However, since we do not find an appropriate reference for this result, we present its complete proof.

Lemma A.2.1 *Let C'' denote the second-order partial derivative of call option price with respect to the underlying price, and $\ddot{C}$ the second-order partial derivative with respect to* *the strike, then we have*

$$S^2C''(S,K) = K^2\ddot{C}(S,K). \tag{A18}$$

Proof:

Let C' denote the partial derivative of call option price with respect to (w.r.t.) the underlying price, $\dot{C}$ the partial derivative w.r.t. strike, and $\dot{C}'$ the second-order partial derivative w.r.t. strike and underlying.

We first demonstrate that $C(S,K)$ is homogeneous of degree 1. In other words, $C(aS,aK) = aC(S,K)$, for all real numbers $a > 0$. To see it, by the risk-neutral pricing equation,

$$C(S,K) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[(S_T-K)^+|S_t = S\right]. \tag{A19}$$

Using the formula $(ax)^+ = ax^+$, for all x and $a > 0$, the payoff is $(aS_T - aK)^+ = a(S_T - K)^+$. Then by the risk-neutral pricing equation again,

$$C(aS,aK) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[(aS_T-aK)^+|aS_t = aS\right] = ae^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[(S_T-K)^+|S_t = S\right] = aC(S,K). \tag{A20}$$

Accordingly, for any $a,b > 0$, we have

$$abC(S,K) = C(abS,abK). \tag{A21}$$

Take $\frac{\partial}{\partial a}$ on both sides of Eq. (A21) and set $a = 1$

$$bC(S,K) = bSC'(bS,bK) + bK\dot{C}(bS,bK). \tag{A22}$$

First, evaluate Equation (A22) at $b = 1$

$$C(S,K) = SC'(S,K) + K\dot{C}(S,K). \tag{A23}$$

Take partial derivative $\frac{\partial}{\partial K}$

$$\dot{C}(S,K) = S\dot{C}'(S,K) + K\ddot{C}(S,K) + \dot{C}(S,K), \tag{A24}$$

and we obtain

$$S\dot{C}'(S,K) + K\ddot{C}(S,K) = 0. \tag{A25}$$

Second, take first $\frac{\partial}{\partial a}$ on both sides of Equation (A22), and then $\frac{\partial}{\partial b}$ on the resulting equation, and set $a = b = 1$, we obtain

$$C(S,K) = SC'(S,K) + S^2C''(S,K) + 2S\dot{C}'(S,K)K + K\dot{C}(S,K) + K^2\ddot{C}(S,K). \tag{A26}$$

We next equate the right-hand sides of Equations (A23) and (A26) and obtain

$$S^2C''(S,K) + 2S\dot{C}'(S,K)K + K^2\ddot{C}(S,K) = 0. \tag{A27}$$

Plug the Equation (A25) into the equation above,

$$S^2C''(S,K) = K^2\ddot{C}(S,K), \tag{A28}$$

and we obtain Lemma A.2.1. □

Now we are ready to prove Proposition 2.4.1.

Proof of Proposition 2.4.1

To compute the right-hand side of Equation (A17), and thus $\mathbb{E}_t[R_{t+1\to t+2}]$, write

$$C(S,K) = \frac{1}{R_f}\int_K^{\infty}(z-K)q(z|S)dz, \tag{A29}$$

where $q(\cdot)$ is the conditional density of S_{t+1} under the risk-neutral probability measure $\mathbb{Q}$ and R_f denotes the gross risk-free return. Now let $\dot{C}$ denote the partial derivative with respect to strike. Then

$$\dot{C}(S,K) = \frac{1}{R_f}\left(\int_0^K q(z|S)dz - 1\right). \tag{A30}$$

Write $\Pi(K|S) = \int_0^K q(z|S)dz$ for the conditional distribution under $\mathbb{Q}$, that is, $\Pi(K|S) = Q(S_{t+1} \leq K|S_t = S)$. Then, $\Pi(K|S) = 1 + R_f\dot{C}(S,K)$, and $d\Pi(K|S) = R_f\ddot{C}(S,K)dK$.

Hence,

$$\begin{aligned}
\int_0^\infty \mathbb{E}_t^Q\left[\frac{C_{t+1}(S_{t+1},K)}{S_{t+1}}\right] dK &= \int_0^\infty \int_0^\infty \frac{C_{t+1}(L,K)}{L}\left(R_{f,t\to t+1}\ddot{C}_t(S_t,L)\,dL\right) dK \\
&= R_{f,t\to t+1}\int_0^\infty \int_0^\infty \frac{C_{t+1}(L,K)}{L}\ddot{C}_t(S_t,L)\,dLdK \\
&= R_{f,t\to t+1}S_t^2\int_0^\infty \int_0^\infty \frac{C_{t+1}(L,K)}{L}\frac{C_t''(S_t,L)}{L^2}\,dLdK, \qquad \text{(A31)}
\end{aligned}$$

where the last line substitutes gamma for strike-gamma using $C''(S,K) = \frac{K^2}{S^2}\ddot{C}(S,K)$, as specified by Lemma A.2.1.

Plug back into Equation (A17)

$$\mathbb{E}_t[R_{t+1\to t+2}] = 2S_t\int_0^\infty \frac{C_t''(S_t,L)}{L^3}\left(\underbrace{\int_0^\infty C_{t+1}(L,K)\,dK}_{\text{inside-integral, } I(L)}\right) dL, \tag{A32}$$

and we obtain Proposition 2.4.1. □

A.3 Nonzero Risk-neutral Relation

In this section, we provide a simple example to demonstrate that the risk-neutral correlation between the spot return and the future return square, $corr_t^{\mathbb{Q}}\left[R_{t\to t+1}, R^2_{t+1\to t+2}\right]$, can be nonzero.

At time $t = 0,1,2$, let the risk-free rate of return be zero, and the risky asset returns during the two consecutive periods be $R_{t\to t+1} = R_1 = 1+\varepsilon$ and $R_{t+1\to t+2} = R_2 = 1+\varepsilon\eta$, respectively. Suppose $\mathscr{F}_1$ is generated by ε, and $\mathscr{F}_2$ is by $\{\varepsilon,\eta\}$, where both ε and η are mean zero and independent of one another. Immediately, we have

$$\mathbb{E}[R_1] = 1 + \mathbb{E}[\varepsilon] = 1, \tag{A33}$$

$$\mathbb{E}_1[R_2] = 1 + \mathbb{E}_1[\varepsilon\eta] = 1 + \varepsilon\mathbb{E}_1[\eta] = 1, \tag{A34}$$

since $\mathbb{E}_1[\eta] = \mathbb{E}[\eta] = 0$. Therefore, the conditional expectation operator, $\mathbb{E}[.]$, is under the risk-neutral probability measure $\mathbb{Q}$.

We next compute the (risk-neutral) covariance between the spot return and the future return square, $Cov(R_1, R_2^2)$. First,

$$\mathbb{E}\left[R_2^2\right] = \mathbb{E}\left[1 + 2\varepsilon\eta + \varepsilon^2\eta^2\right] = 1 + \mathbb{E}\left[\varepsilon^2\eta^2\right] \tag{A35}$$

Second,

$$\mathbb{E}\left[R_1R_2^2\right] = \mathbb{E}\left[(1+\varepsilon)(1 + 2\varepsilon\eta + \varepsilon^2\eta^2)\right] = \mathbb{E}\left[1 + 2\varepsilon\eta + \varepsilon^2\eta^2\varepsilon + 2\varepsilon^2\eta + \varepsilon^3\eta^2\right], \tag{A36}$$

$$= 1 + \mathbb{E}\left[\varepsilon^2\eta^2\right] + \mathbb{E}\left[\varepsilon^3\eta^2\right]. \tag{A37}$$

Hence,

$$Cov(R_1, R_2^2) = \mathbb{E}\left[R_1R_2^2\right] - \mathbb{E}[R_1]\,\mathbb{E}\left[R_2^2\right] = \mathbb{E}\left[\varepsilon^3\eta^2\right], \tag{A38}$$

$$= \mathbb{E}\left[\varepsilon^3\right]\mathbb{E}\left[\eta^2\right]. \tag{A39}$$

In other words,

$$corr(R_1, R_2^2) \neq 0 \quad \text{if and only if} \quad \mathbb{E}\left[\varepsilon^3\right] \neq 0. \tag{A40}$$

If we choose $\varepsilon \sim \eta$ (the same distribution), in theory, the risk-neutral correlation $corr(R_1, R_2^2)$ can be any number (with the same sign as $\mathbb{E}\left[\varepsilon^3\right]$), as long as $\mathbb{E}\left[\varepsilon^3\right] \neq 0$.

APPENDIX B: MISPRICING AND ANOMALIES: AN EXOGENOUS SHOCK TO SHORT SELLING FROM JGTRRA

Table B.1: DID results over balanced sample periods

This table reports the main DID results over the balanced sample periods over $[-5,+5]$, $[-10,+10]$, and $[-15,+15]$ years before and after the JGTRRA 2003. We add both firm and month fixed effects and cluster standard errors on both firm and time. t-statistics are presented in parentheses below the coefficient estimates. *, **, and *** denote two-tail statistical significance at the 10%, 5%, and 1% levels, respectively.

Sample periods	$[-5,+5]$	$[-10,+10]$	$[-15,+15]$
NOPS	-0.125***	-0.108***	-0.103***
	(-4.92)	(-7.94)	(-9.99)
DivR	-0.207	-0.267*	-0.292***
	(-0.75)	(-1.71)	(-2.63)
NOPS×*DivR*	0.058***	0.041***	0.035***
	(3.79)	(4.46)	(4.94)
NOPS×*JGTRRA*	0.059**	0.059***	0.061***
	(2.37)	(3.95)	(5.18)
DivR×*JGTRRA*	-0.045	-0.035	0.092
	(-0.10)	(-0.12)	(0.44)
NOPS×*DivR*×*JGTRRA*	-0.059***	-0.043***	-0.032***
	(-2.79)	(-3.14)	(-3.05)

Table B.2: DID results with *MISP*

This table reports the main DID results after replacing *NOPS* by Stambaugh et al. (2015) mispricing score, *MISP*. The *MISP* data are obtained from Robert F. Stambaugh's website. *t*-statistics are presented in parentheses below the coefficient estimates. The sample period is 1985:7 to 2016:12 due to the availability of *MISP* data. *** denotes two-tail statistical significance at the 1% level.

Fixed Effects	Month	Month	Firm & Month	Firm & Month
S.E. Clusters	Month	Firm & Month	Month	Firm & Month
MISP	-0.049***	-0.049***	-0.041***	-0.041***
	(-8.54)	(-8.53)	(-5.98)	(-6.00)
DivR	-1.767***	-1.767***	-1.171***	-1.171***
	(-8.09)	(-8.00)	(-5.63)	(-5.57)
$MISP \times DivR$	0.034***	0.034***	0.017***	0.017***
	(5.75)	(5.73)	(3.98)	(3.97)
$MISP \times JGTRRA$	0.034***	0.034***	0.039***	0.039***
	(4.38)	(4.38)	(5.35)	(5.37)
$DivR \times JGTRRA$	1.372***	1.372***	1.255***	1.255***
	(4.56)	(4.58)	(4.20)	(4.23)
$MISP \times DivR \times JGTRRA$	-0.029***	-0.029***	-0.024***	-0.024***
	(-3.68)	(-3.69)	(-3.67)	(-3.69)

Table B.3: Within-firm portfolios

This table reports results using the within-firm calendar-time portfolio method in Hartzmark and Solomon (2013). Each month, we sort dividend-paying firms into quintile portfolios based on $NOPS$. The treated group consists of firms that report dividend record dates in the previous months, while the control group includes firms that do not report dividend record dates in the previous months, but reported dividend record dates in the previous 12 months. Then we obtain the respective long-short portfolios for the treated and control groups. Finally, we regress the difference in the long-short portfolio returns between the treated and control groups on $JGTRRA_{t-1}$. We also control for Fama and French (2015) five factors and Hou et al. (2015) four factors. Newey-west t-statistics are presented in parentheses below the coefficient estimates. * and ** denote two-tail statistical significance at the 10% and 5% levels, respectively.

	Long-Short	Long-Short	Long-Short
JGTRRA	0.308**	0.279*	0.276*
	(2.11)	(1.95)	(1.87)
MKT		-0.021	
		(-0.79)	
SMB		-0.020	
		(-0.62)	
HML		-0.005	
		(-0.10)	
RMW		-0.049	
		(-1.09)	
CMA		-0.054	
		(-0.98)	
MKT (HXZ)			-0.019
			(-0.87)
ME			-0.003
			(-0.08)
IA			-0.065
			(-1.15)
ROE			-0.004
			(-0.09)
Constant	-0.228**	-0.168	-0.181
	(-2.22)	(-1.62)	(-1.63)