Variance risk in aggregate stock returns and time-varying return predictability

JFE2019.04

主讲人：王 航

Author： Sungjune Pyun

Assistant Professor at NUS Business School Department of Finance

Education

Ph.D. , Business Administration (Finance) , University of Southern California , 2011-2017

M.S. , Statistics , Stanford University , 2008-2010

B.A , Applied Statistics ; Business Administration , Yonsei University , 2001-2008

Research & Teaching Interests

Empirical Asset Pricing, Variance Risk, Financial Derivatives, Financial Econometrics

Working Paper

Variance Risk Premium in Individual Stocks and the Exposure to Factor Variance Risk

Abstract

- This paper introduces a new out-of-sample forecasting methodology for monthly market returns using the variance risk premium (VRP)
- This methodology is motivated by the 'beta representation'
- When the slope of the contemporaneous regression of market returns on variance innovation is larger, future returns are more sharply related to the current VRP
- Predictions are more accurate when market returns are highly correlated to variance shocks

The structure of this paper

1. Introduction
2. The variance risk premium and the expected market returns
3. Data and estimation
4. Out-of-sample predictions
5. In-sample predictions
6. Robustness
7. Conclusion

1. Introduction

P 1-2 （ Background ）

- predictive relationships appear to change over time， (Fama and French, 1988a; Dangl and Halling, 2012)
- predictors that perform well in sample often fail out of sample， (Goyal and Welch, 2008; Campbell and Thompson, 2008)
- return predictions typically perform poorly at shorter horizons, (Fama and French, 1988a)
- monthly or quarterly market returns are predictable by the one-month variance risk premium (VRP), (Bollerslev et al. 2009)

$$R_{m,t+1} = \beta_0 + \beta_p VRP_t + \epsilon_{t+1} \quad (1)$$

P 3-5 （A new approach）

The new methodology is derived from two theoretical observations:

- The one-month market risk premium should be related to the VRP by the market's exposure to variance risk. Then,

$$R_{m,t} = \beta_{v,0} + \beta_v (RV_t - E_{t-1}[RV_t]) + \epsilon_{0,t} \qquad (2)$$

- When variance risk is responsible for a larger fraction of market risk, the VRP should explain a greater share of the market risk premium.

P 6-10 (Comparison with traditional methods)

- Defects of traditional methods
- Advantages of the new methodology:

1. the contemporaneous regression of returns on variance innovations has a much higher R^2

2. the new approach only uses the most recent month of data to determine the parameters

3. the new approach strictly outperforms the traditional way of return forecasting at the monthly horizon

4. the out-of-sample predictive power of the VRP depends strongly on the degree of correlation between market returns and variance innovations

P 11-12 (The in-sample results and Cause analysis)

- The predictive beta estimated from in-sample regressions decreases in the contemporaneous variance beta
- The ability to more directly estimate the contribution of variance risk to the market risk premium follows from three unique characteristics of the VRP :

1. the VRP precisely measures the price of variance risk

2. unexpected changes in market variance is estimable relatively accurately using high-frequency data

3. variance risk comprises a large part of the variation in market returns

1.1 Related literature

This paper is related to at least four different areas of research:

1. the leverage effect, Carr and Wu (2016), Bandi and Reno (2016)

2. time-varying return predictability, Henkel et al. (2011) and Dangl and Halling (2012), Lettau and Van Nieuwerburgh (2008), Johannes et al. (2014), Rapach et al. (2010)

3. the role of the price of variance risk across various asset classes, Martin and Wagner (2016), Londono (2014) and Bollerslev et al. (2014), Londono and Zhou (2017), Wang et al. (2013) (credit default swaps) and Choi et al. (2017) (bonds)

4. downside risk, Kelly and Jiang (2014), Feunou et al. (2017) and Bollerslev et al. (2015), Carr and Wu (2016), Bekaert and Hoerova (2014), Chen et al. (2018)

2. The variance risk premium and the expected market returns

variance risk premium (VRP):

$$VRP_T = Cov_T\left(SDF_{T,T+1}, \int_T^{T+1} dV_t\right)$$
$$\approx E_T^Q\left[\int_T^{T+1} dV_t\right] - E_T\left[\int_T^{T+1} dV_t\right] \quad (3)$$

Bollerslev et al. (2009) find that the VRP predicts short-term market returns. They run predictive regressions of monthly, quarterly, and semi-annual market returns ($R_{m,t+1}$) on the VRP_t,

$$R_{m,t+1} = \beta_0 + \beta_p VRP_t + \epsilon_{t+1} \quad (4)$$

the VRP is unique for several other reasons：

- VRP is actually a price of risk
- the factor on which it is based, namely, variance innovations

the market price and variance tend to move in the opposite direction：

- the “leverage” effect
- “volatility feedback”

2.1 A simple model

$$\frac{dS_t}{S_t} = \mu_t d_t + \sqrt{V_t}(\rho_t dW_t^v + \sqrt{1-\rho_t^2} dW_t^0) \quad (5)$$

$$dV_t = \theta_t dt + \sigma_v dW_t^v \quad (6)$$

$$\frac{dS_t}{S_t} = \mu_t d_t + \rho_t \frac{\sqrt{V_t}}{\sigma_v}(dV_t - \theta_t dt) + \sqrt{(1-\rho_t^2)V_t} dW_t^0 \quad (7)$$

$$Cov_T\left(-SDF_{T,T+1}, \int_T^{T+1} \frac{dS_t}{S_t}\right) = -\rho_t \frac{V_t}{\sigma_v} Cov_T\left(SDF_{T,T+1}, \int_T^{T+1} dV_t\right) - \sqrt{(1-\rho_t^2)V_t} Cov_T\left(SDF_{T,T+1}, \int_T^{T+1} dW_t^0\right) \quad (8)$$

2.2 Empirical implications

- The slope that determines the relation between the short-term market risk premium and the risk premium on market variance is largely determined by the amount of variance risk present in the market portfolio.

$$R_{m,t} = \beta_{v,0} + \beta_v (RV_t - E_{t-1}[RV_t]) + \epsilon_{o,t} \quad (9)$$

$$E_T[R_{m,T+1}] = -\beta_v VRP_T + O_T, \quad (10)$$

- The equation that describes the relation between the VRP and the market risk premium suggests that market returns are more accurately predictable when the index and the variance of returns move closely together. If the orthogonal premium is unpriced or unrelated to the VRP, the orthogonal premium will appear as noise in a predictive regression in which the VRP is the sole predictor. If the premium is priced and related to the VRP, this premium will bias the predictive beta. In either case, as the contemporaneous correlation (ρ_t) gets closer to zero, predictions will become less accurate. On the other hand, when correlations are close to −1, the VRP should almost entirely identify the market risk premium.

Contemporaneous beta approach：

The method is based on the close relationship between the predictive and contemporaneous betas and implemented by using the beta of the contemporaneous regression in place of the predictive beta to form the out-of-sample forecast.

Reasons：

- The contemporaneous relation between returns and changes in variance is much stronger than the predictive relationship between the VRP and future returns.
- Both returns and estimates of realized variance are available at the daily frequency. Hence, this new approach can be used even when the predictive relation changes rapidly over time.

3. Data and estimation

3.1 Forecasting variance

This paper use intraday, high-frequency, return-based RV to model the forecasts.

The high-frequency intraday trading data for the S&P 500 Index are obtained from Tickdata. This paper requires the data between 1989 and0 2016, since the first component used to estimate the VRP, the VIX2, is only available from 1990.

To estimate the second component, RV is computed by first calculating squared log returns from the last tick of each five-minute interval.

The constructed RV series is then used to compute the variance forecasts.

This paper use a variation of Corsi's(2009) model and forecast the RV instead of the volatility.

$$RV_{\tau+k} = a_0 + a_d RV_\tau + a_w \sum_{j=0}^{4} RV_{\tau-j} + a_m \sum_{j=0}^{4} RV_{\tau-j} + e_{\tau+k} \tag{11}$$

The one-day RV forecast ($\widehat{RV}_{\tau+1|\tau}$) can then be constructed using the loadings on the daily, weekly, and monthly components. The forecast of monthly variance at day τ (($\widehat{RV}_{\tau+1,\quad\tau+22|\tau}$) is estimated by averaging the 22 daily forecasts (k = 1, . . . , 22). The forecasts are estimated using daily observations on a 12-month rolling window to ac- count for the possibility that the forecast relation changes over time.

3.2 Estimation of the VRP

The VRP is measured by taking the difference between the square of VIX and the monthly forecast of RV.

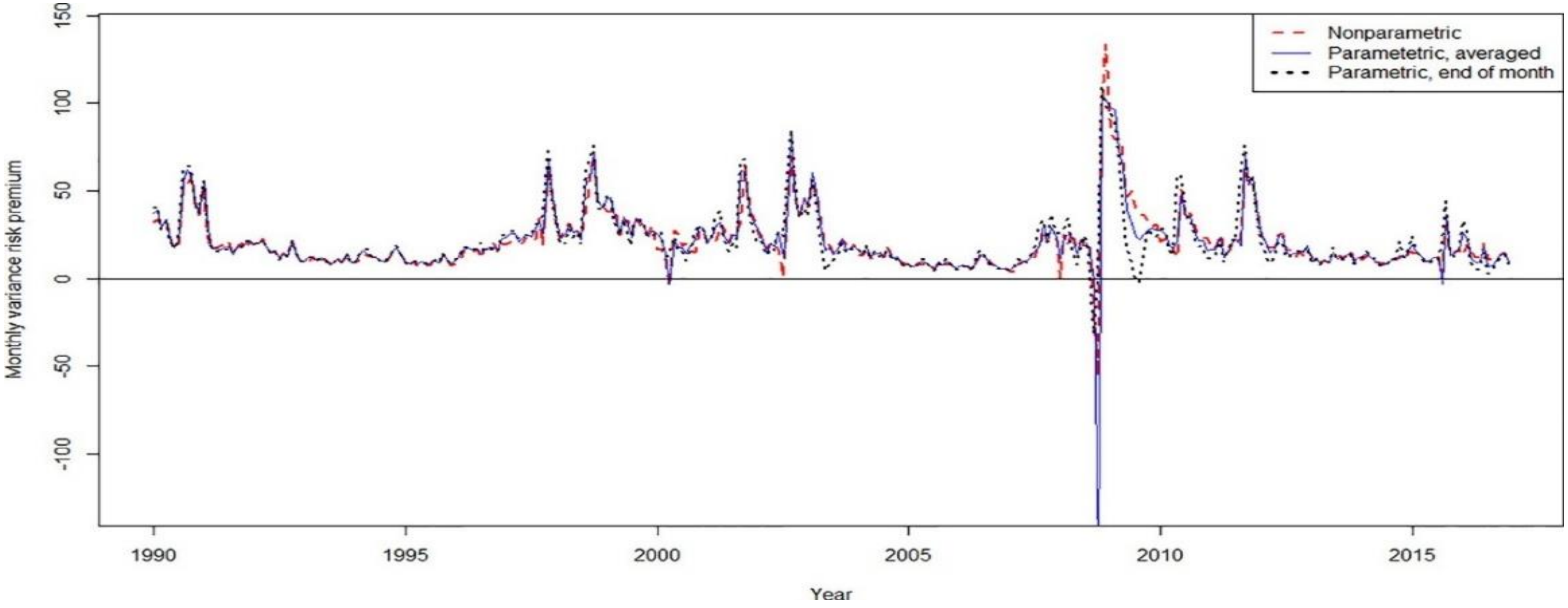
To deal with the mismatch, I either average the daily observations or take the end-of-month values. These two measures are parametric and denoted by $VRP_{\bar{P}}$ andVRP_{PE}, respectively.

$$VRP_{\bar{p},t} = \sum_{\tau \in t} \left(\frac{VIX_{\tau}^{2}}{252} - \frac{\widehat{RV}_{\tau+1,\tau+22|\tau}}{22}\right) \quad (12)$$

$$VRP_{P^{E},t} = \frac{VIX_{m(t)}^{2}}{12} - \widehat{RV}_{m(t)+1,m(t)+22|m(t)} \quad (13)$$

This paper supplement these two measures with a third nonparametric one. Denoted by VRP_N, the nonparametric VRP is the difference between the scaled VIX and the historical RV, both averaged over the entire month.

Fig. 1 compares the time series of the three VRPs used in this paper.



The daily innovation of market variance is calculated by computing the unexpected changes in RV scaled so that it matches the one-month interval. Then, the monthly contemporaneous beta is estimated from the regression of market returns on variance innovations, using only observations that belong to that particular month.

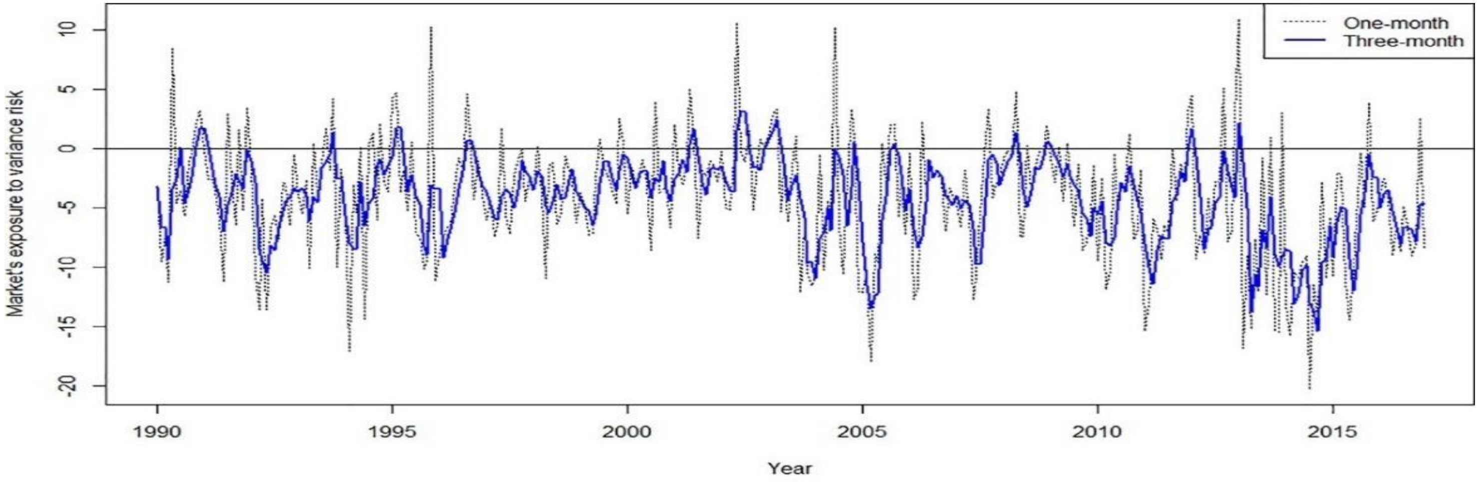
$$R_{m,\Gamma} = \beta_{v,0,t} + \beta_{v,t}(RV_{\Gamma} - R\tilde{V}_{\Gamma|\Gamma-1}) + \epsilon_{\Gamma} \quad (14)$$

To deal with possible heteroscedasticity, I consider weighted least squares (WLS) in addition to ordinary least squares (OLS).

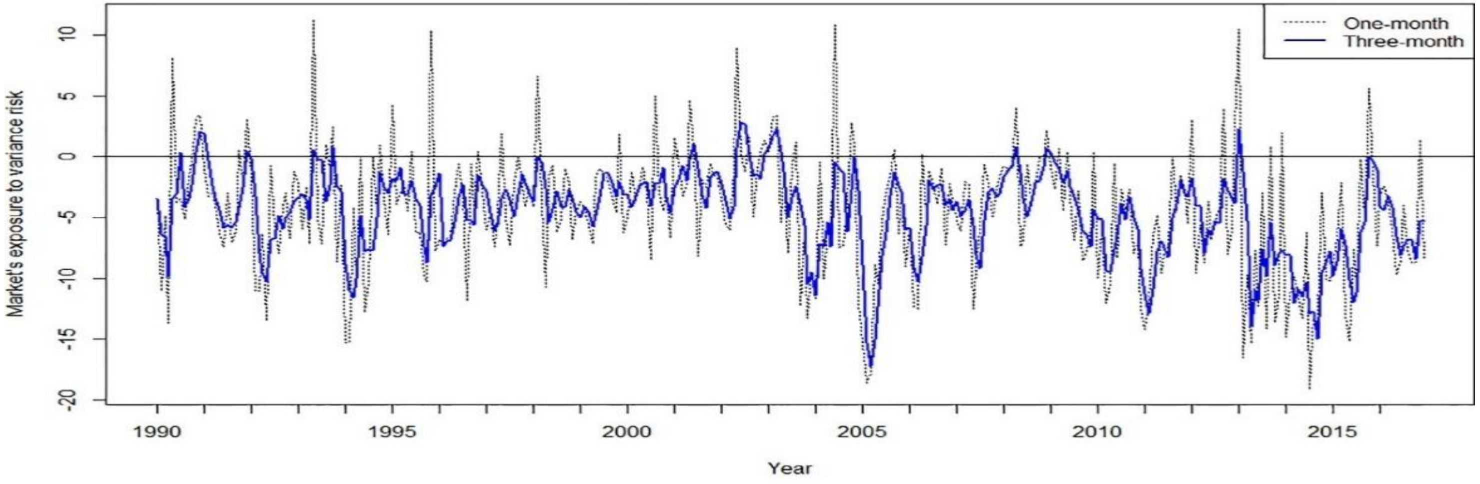
The contemporaneous correlation $(\hat{\rho}_t)$ is the correlation between the two variables in the above equation. The correlations are closely connected to the betas because they are transformations of each other.

$$\hat{\rho}_t = \hat{\beta}_{v,t} \times \frac{\hat{\sigma}_t(\mathrm{RV}_{\tau} - \widehat{RV}_{\tau|\tau-1})}{\hat{\sigma}_t(R_{m,\tau})} \quad (15)$$

The time series of the betas is provided in Fig. 2.



(a) Weighted least squares



(b) Ordinary least squares

The time series of the correlations is in Fig. 3.

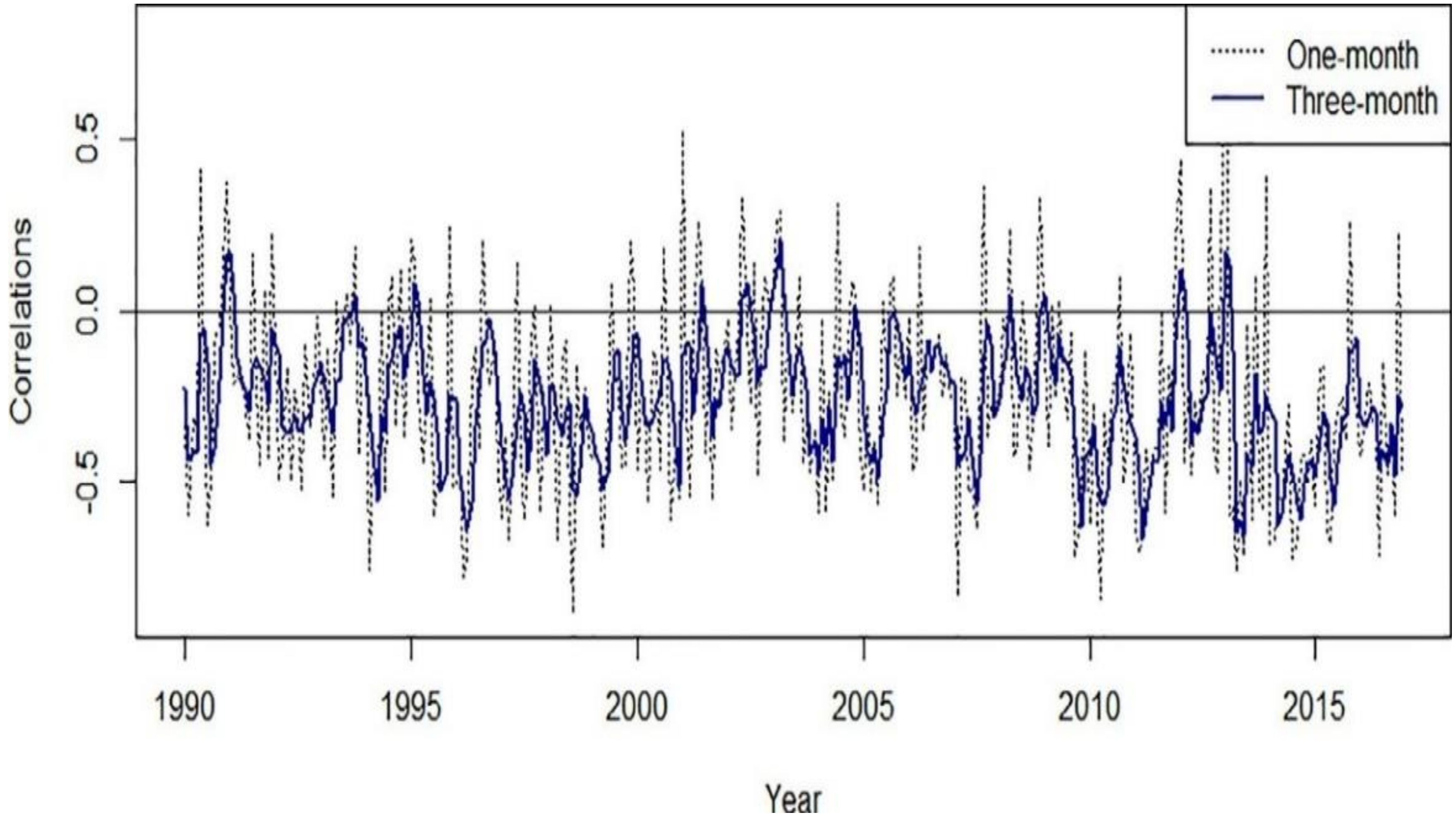


Table 1

Panel A: Summary statistics

	Mean	StDev	$\hat{\rho} \leq$ median Mean	$\hat{\rho} \leq$ median StDev	NBER recession Mean	NBER recession StDev	Autocorr.
RV (monthly)	16.43	29.51	16.67	34.68	47.00	72.52	0.643
Implied variance (monthly)	19.71	7.49	19.89	7.34	29.02	10.30	0.840
VRP_N	20.96	16.74	19.01	14.63	38.13	32.58	0.696
$VRP_{\overline{P}}$	21.74	17.86	21.43	16.38	39.76	31.01	0.764
VRP_{PE}	21.30	18.77	21.86	19.98	31.81	29.04	0.601
$\hat{\beta}_v$	−4.19	5.10	−7.72	3.92	−1.49	3.05	0.200
$\hat{\beta}_{v,WLS}$	−4.56	5.08	−7.72	3.81	−1.60	3.01	0.203
$\hat{\rho}$	−0.259	0.281	−0.486	0.132	−0.130	0.226	0.124
Number of months	324		162		37		

Panel B: The leverage effect and correlations

	$\hat{\beta}_{v,t+1}$	$\hat{\rho}_{t+1}$
Contemporaneous returns ($R_{M,t+1}$)	0.120	0.280
Lagged annual returns ($\sum_{k=0}^{11} R_{M,t-k}$)	−0.261	−0.287
RV_t	0.196	0.142
VIX_t	0.297	0.165
VIX trend_t	−0.121	−0.259
$SKEW_t$	−0.300	−0.222
Tail risk_t	−0.128	−0.131
$VRP_{N,t}$	0.199	0.152
$VRP_{\overline{P},t}$	0.274	0.214
$VRP_{PE,t}$	0.220	0.128

4. Out-of-sample predictions

This section documents two finding:

- The beta that explains the predictive relationship is close to the negative of the contemporaneous beta.
- Predictions perform better when the contemporaneous correlation between market returns and variance innovations is more negative.

4.1 Out-of-sample predictions

The traditional approach to providing OOS forecasts:

(1) I run a predictive regression using the past k months of historical data (from time T − k + 1 to T) as

$$R_{m,t} = \beta_0 + \beta_p VRP_{t-1} + \epsilon_t \quad (16)$$

Then, the one-step-ahead predicted value of the excess market returns ($\hat{R}_{m,T+1|T}$) is given as $\hat{\beta}_{0,T} + \hat{\beta}_{p,T} VRP_T$.

(2) using the OOS-R^2 to evaluate the OOS predictive performance

$$1 - \frac{\sum_t \left(\hat{R}_{m,t+1|t} - R_{m,t+1}\right)^2}{\sum_t \left(\bar{R}_{m,t} - R_{m,t+1}\right)^2} \quad (17)$$

The Wald statistic is given as

$$W = T\left(T^{-1}\sum_{t=1}^{T} \Delta L_{t+1}\right)\hat{\Omega}^{-1}\left(T^{-1}\sum_{t=1}^{T} \Delta L_{t+1}\right) \quad (18)$$

Where $\triangle L_{t+1}=(R_{m,t}-R_{m,t+1})^2-(R_{m,t+1|t}-R_{m,t+1})^2$ and $\Omega=\sum_{t=1}^{T}(\Delta L_{t+1}-\overline{\Delta L})^2$. Asymptotically, this Wald statistic follows a Chi-square distribution with degrees of freedom equal to the difference in the number of predictors.

"contemporaneous beta" approach:

$$\hat{R}_{m,T+1|T} = -\hat{\beta}_{v,T} VRP_T \qquad (19)$$

The product of the negative variance risk exposure and the VRP predicts excess market returns with a zero intercept. This relation is based on the assumption that the orthogonal component is either unpriced or is too noisy to determine in the short-run.

"hybrid approach" (a combination of the contemporaneous beta and traditional approaches):

$$R_{m,t+!} = -\hat{\beta}_{v,t} VRP_t + \delta_0 + \delta_1 \sqrt{1-\hat{\rho}_t^2} X_t + \eta_{t+1} \qquad (20)$$

$$\hat{R}_{m,T+1|T} = -\hat{\beta}_{v,t} VRP_t + \hat{\delta}_0 + \hat{\delta}_1 \sqrt{1-\hat{\rho}_t^2} X_t \qquad (21)$$

Table 2 summarizes the OOS-R^2s and the Wald statistics, along with p-values

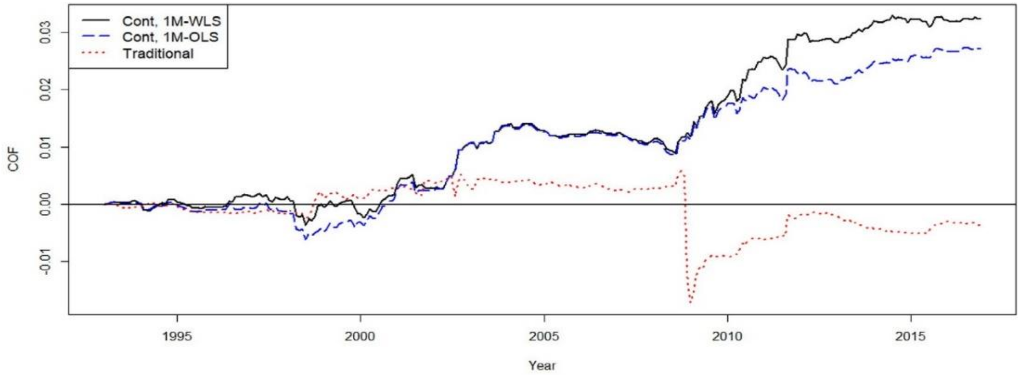
		VRP measures					
		VRP_N		$\text{VRP}_{\bar{P}}$		VRP_{pE}	
		Unconstrained	Constrained	Unconstrained	Constrained	Unconstrained	Constrained
Panel A: The traditional approach							
	OOS-R^2	0.010	−0.007	0.002	−0.008	0.052	0.032
	Wald	0.064	0.033	0.004	0.060	1.501	0.875
	p-value	(0.800)	(0.857)	(0.952)	(0.807)	(0.221)	(0.350)
Panel B: The contemporaneous beta approach							
B-1: No intercept							
1-month	OOS-R^2	0.065	0.061	0.079	0.076	0.084	0.079
WLS	Wald	5.027	6.836	7.956	10.838	5.996	6.984
	p-value	(0.025)	(0.009)	(0.005)	(0.001)	(0.014)	(0.008)
1-month	OOS-R^2	0.054	0.051	0.068	0.066	0.085	0.069
OLS	Wald	3.686	5.325	8.056	8.964	4.544	5.665
	p-value	(0.055)	(0.021)	(0.005)	(0.003)	(0.033)	(0.017)
3-month	OOS-R^2	0.064	0.066	0.049	0.052	0.064	0.064
WLS	Wald	11.460	12.232	3.554	3.818	4.130	4.122
	p-value	(0.001)	(0.000)	(0.059)	(0.051)	(0.042)	(0.042)
3-month	OOS-R^2	0.053	0.060	0.041	0.047	0.054	0.059
OLS	Wald	10.096	11.462	2.735	3.552	3.339	3.862
	p-value	(0.001)	(0.001)	(0.098)	(0.059)	(0.068)	(0.049)
B-2: Including intercept							
1-month	OOS-R^2	0.059	0.053	0.074	0.068	0.080	0.073
WLS	Wald	4.078	5.317	3.668	8.991	5.096	6.018
	p-value	(0.043)	(0.021)	(0.055)	(0.003)	(0.024)	(0.014)
1-month	OOS-R^2	0.047	0.045	0.062	0.061	0.065	0.064
OLS	Wald	4.920	4.084	6.485	7.543	5.096	4.899
	p-value	(0.027)	(0.043)	(0.011)	(0.006)	(0.024)	(0.027)

To better understand cumulative improvements in the loss function over the benchmark. I define the Cumulative Outperformance of the Forecast (COF) as : when the new approach performs especially better over the traditional approach, I develop a measure that computes the

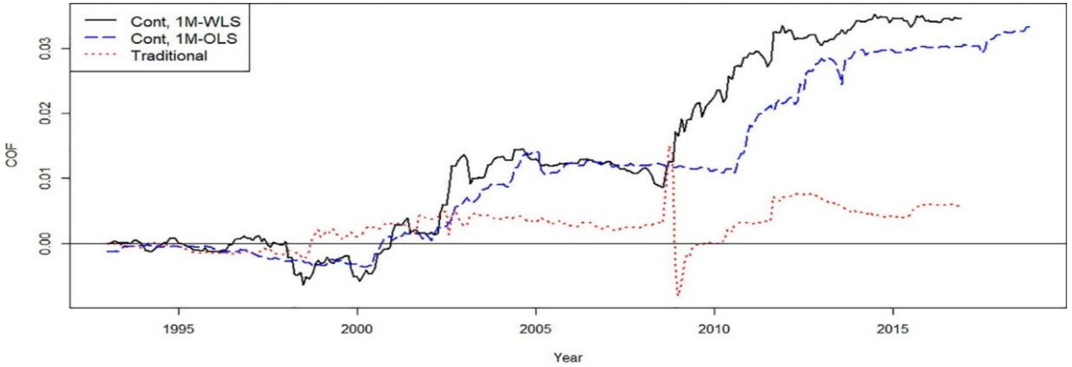
$$COF_T = \sum_{t=1}^{T} \Delta L_t \qquad (22)$$

where Lt is the square loss function given in Eq. (18).

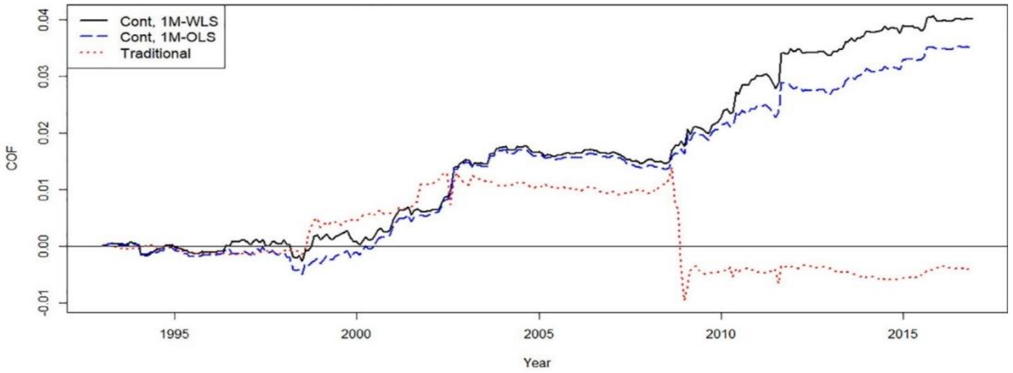
Fig. 4 plots the COF of the constrained forecast, and Fig. 5 that of the unconstrained forecast.



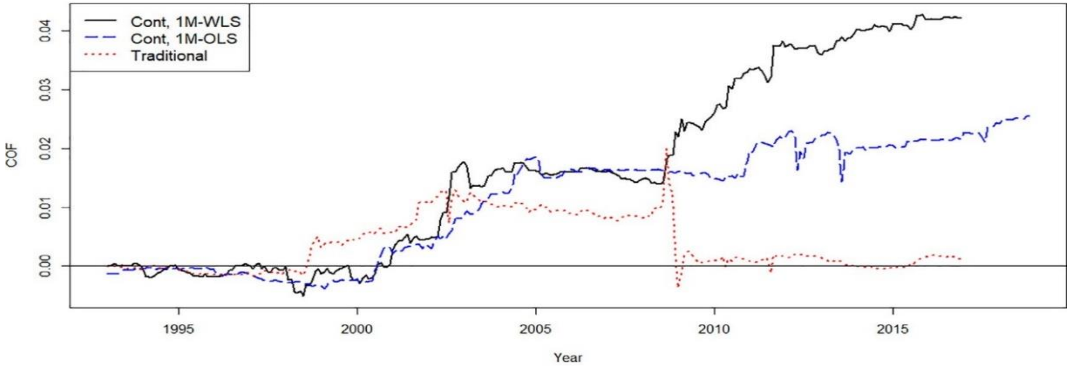
(a) Variance risk premium (nonparametric)



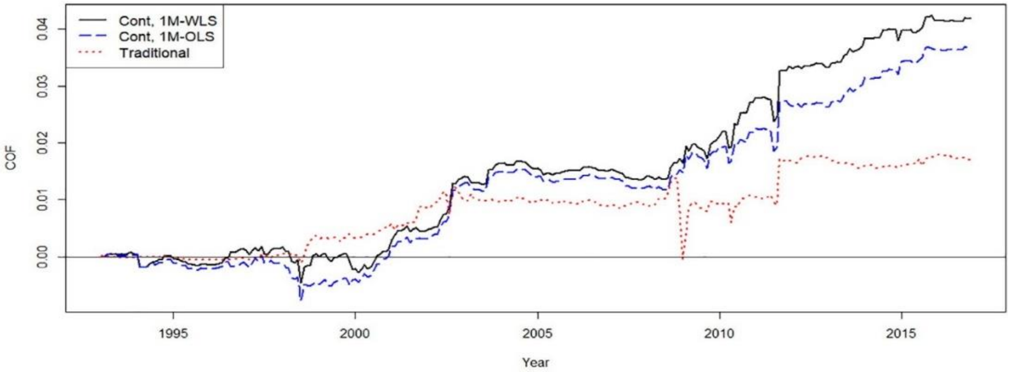
(a) Variance risk premium (nonparametric)



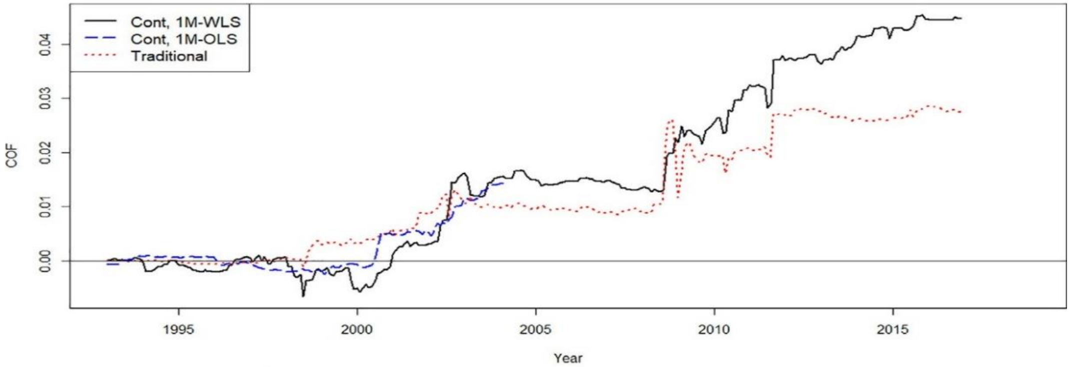
(b) Variance risk premium (parametric, averaged)



(b) Variance risk premium (parametric, averaged)



(c) Variance risk premium (parametric, end of month)



(c) Variance risk premium (parametric, end of month)

4.2 Time-varying out-of-sample predictability

This paper also study the connection between contemporaneous correlations and predictive R^2s. I do so by dividing the full sample into different non-overlapping subsamples.

Each of the 288 months in the full sample period of 1993–2016 is classified into one of three groups according to the monthly series of the contemporaneous correlations between market returns and variance innovations.

When the correlation during a particular month is more negative than the first tercile of the historical distribution of past values, the month is classified as a “high” month. When the correlation is more positive than the second tercile, it is classified as a “lower” month. Therefore, the classifications are made without any look-ahead bias.

Then, the OOS-R^2s are computed separately for each of these groups.

Table 3 summarizes the OOS-R^2s for each of the subsamples.

		OOS-R^2					
		VRP_N		$VRP_{\overline{P}}$		VRP_{PE}	
		1-month	3-month	1-month	3-month	1-month	3-month
Panel A: The traditional approach							
	High	0.068	0.162	0.096	0.065	0.164	0.207
	Medium	0.049	0.069	−0.035	0.057	−0.050	0.060
	Low	−0.124	−0.135	−0.097	−0.079	0.000	−0.056
	High-low	0.193	0.298	0.193	0.144	0.164	0.263
Panel B: The contemporaneous beta approach							
B-1: No intercept							
WLS	High	0.082	0.131	0.122	0.065	0.161	0.128
	Medium	0.056	0.054	0.053	0.048	0.031	0.053
	Low	0.050	0.047	0.045	0.059	0.036	0.048
	High-low	0.032	0.085	0.077	0.006	0.125	0.080
OLS	High	0.079	0.126	0.120	0.064	0.156	0.123
	Medium	0.042	0.042	0.038	0.037	0.009	0.041
	Low	0.029	0.039	0.024	0.050	0.020	0.040
	High-low	0.050	0.088	0.096	0.014	0.136	0.083
B-2: Including intercept							
WLS	High	0.074	0.147	0.114	0.084	0.139	0.149
	Medium	0.057	0.033	0.055	0.024	0.050	0.031
	Low	0.040	0.035	0.035	0.050	0.024	0.038
	High-low	0.034	0.112	0.079	0.034	0.115	0.111
OLS	High	0.070	0.140	0.111	0.080	0.131	0.141
	Medium	0.041	0.021	0.037	0.014	0.032	0.019
	Low	0.021	0.025	0.016	0.039	0.004	0.028
	High-low	0.049	0.114	0.095	0.040	0.128	0.112

- When market prices and variance move closely together, the VRP is a very powerful predictor of short-horizon market returns.

- When they move independently, it is hard to predict market returns using the VRP, since the market portfolio is less exposed to variance risk.

4.3 Explaining the orthogonal premium

Return predictors other than the VRP may also complement the VRP for two reasons.

- First, the predictive power of the VRP is strong for monthly and quarterly returns.
- Second, the predictive strength of many common predictors tends to decrease for the post-1993 period. In contrast, the VRP has been demonstrated to be a strong predictor of market returns in the post-1990 period.

Additional variables (X_t)		1993–2016 VRP$_N$	VRP$_{\overline{P}}$	VRP$_{PE}$	Subsamples High	Medium	Low	H-L
Intercept only	OOS-R^2	0.047	0.062	0.065	0.070	0.041	0.021	0.049
	p-value	(0.027)	(0.011)	(0.024)				
D/Y	OOS-R^2	0.016	0.042	0.033	0.017	0.019	0.013	0.004
	p-value	(0.464)	(0.167)	(0.235)				
TERM	OOS-R^2	0.021	0.053	0.028	0.058	0.019	−0.021	0.079
	p-value	(0.260)	(0.061)	(0.234)				
DEF	OOS-R^2	−0.015	0.018	0.002	0.000	0.045	−0.079	0.079
	p-value	(0.639)	(0.614)	(0.962)				
Short rate	OOS-R^2	−0.010	0.008	-0.310	0.007	0.031	−0.063	0.069
	p-value	(0.731)	(0.839)	(1.000)				
Short interest	OOS-R^2	0.020	0.052	0.023	0.048	0.028	−0.020	0.068
	p-value	(0.361)	(0.096)	(0.397)				
cay	OOS-R^2	0.025	0.054	0.038	0.089	−0.067	0.023	0.066
	p-value	(0.359)	(0.125)	(0.240)				
NO/S	OOS-R^2	0.015	0.048	0.025	0.036	0.027	−0.019	0.055
	p-value	(0.434)	(0.106)	(0.300)				
LJV	OOS-R^2	−0.104	−0.189	−0.103	−0.013	−0.108	−0.212	0.199
	p-value	(0.092)	(0.276)	(0.205)				
VRP	OOS-R^2	−0.001	−0.070	−0.074	0.069	0.023	−0.100	0.169
	p-value	(0.995)	(0.312)	(0.150)				

4.4 Evaluating economic significance– a trading strategy

- This paper also evaluate whether the closeness between the two betas can be used to form a trading strategy. Following Goyal and Welch (2008), I use the one-step-ahead OOS forecasts to calculate optimal weight on the stock market as

$$w_T = \frac{\hat{R}_{m,T+1|T}}{\gamma \hat{\sigma}_T^2} \quad (23)$$

The certainty equivalent (CE) of the return is computed as

$$CE = \overline{R_p} - \frac{\gamma}{2} \widehat{Var}(R_p) \quad (24)$$

- Consider an alternative strategy, in which a fraction of the allocation of stocks depends on the model-based predicted returns and the rest on the historical average of past returns. The weight invested in the risky asset becomes

$$w_T' = \frac{\hat{R}_{m,T+1|T}}{\gamma \hat{\sigma}_T^2} \sqrt{\hat{\rho}_T^2} + \frac{\bar{R}_{m,T}}{\gamma \hat{\sigma}_T^2} \sqrt{1 - \hat{\rho}_T^2} \quad (25)$$

Table 5 summarizes the resulting gains/losses in the annualized SRs and CEs.

	Unconditional weighting				Conditional weighting			
	SR	CE	Δ SR	Δ CE	SR	CE	Δ SR	Δ CE
Fixed weight	0.527	0.046						
Average (benchmark)	0.632	0.040						
The traditional approach								
VRP_N	0.524	0.033	−0.108	−0.007	0.673	0.048	+0.042	+0.009
$VRP_{\overline{P}}$	0.667	0.044	+0.035	+0.004	0.739	0.054	+0.107	+0.014
VRP_{PE}	0.672	0.053	+0.040	+0.013	0.741	0.059	+0.109	+0.019
The contemporaneous beta approach with no intercept								
VRP_N	0.760	0.078	+0.129	+0.039	0.729	0.071	+0.097	+0.031
$VRP_{\overline{P}}$	0.922	0.098	+0.290	+0.058	0.836	0.084	+0.204	+0.044
VRP_{PE}	0.782	0.090	+0.151	+0.050	0.749	0.078	+0.117	+0.039
The contemporaneous beta approach including intercept								
VRP_N	0.753	0.081	+0.121	+0.041	0.715	0.070	+0.083	+0.030
$VRP_{\overline{P}}$	0.902	0.098	+0.270	+0.059	0.817	0.081	+0.185	+0.042
VRP_{PE}	0.812	0.097	+0.180	+0.057	0.750	0.079	+0.119	+0.039

Conclusion

- Predictions under the traditional approach could be highly misleading during periods when returns and variance innovations are unrelated. During these times, investors appear to perceive variance risk as unrelated to market risk. The VRP, therefore, provides little information about the market risk premium.

- When the correlation is highly negative, the VRP and the market risk premium are also highly related because market and variance risk are closely related. Moreover, they are connected in a particular way, so that the market's exposure to variance risk can replace the predictive beta. The contemporaneous beta approach predicts the one-month market also in an economically significant manner.

5. In-sample predictions

This section confirm that the key results also hold in sample.

I first summarize the results of the classical predictive regressions, replicating that of Bollerslev et al. (2009).

Next, I examine properties of the time-varying predictive beta and whether the predictive beta can be inferred from the past contemporaneous relation between returns and variance innovations.

Then, I show that the in-sample predictive beta is approximately proportional to the contemporaneous beta.

Finally, I investigate the performance of the predictive regressions over time and demonstrate that their accuracy is related to the correlation between market returns and variance innovations.

This paper suggests that this predictive beta may change over time. They must be higher when the market portfolio loads more on variance risk, and lower when the market does not load on variance risk. This hypothesis can be directly tested by running the regression of

$$R_{m,t+1} = \gamma_0 + \gamma_p VRP_t + \gamma_I VRP_t \times \hat{\beta}_{v,t} + \varepsilon_{p,t+1} \tag{26}$$

Table 6 summarizes the regression coefficients, t-statistics, and the adjusted-R^2s of the simple predictive regression as well as the interactive regressions.

Panel A: Prediction using ordinary least squares

$\hat{\beta}_{v,t}$	VRP_N			$VRP_{\overline{P}}$			VRP_{PE}		
		OLS	WLS		OLS	WLS		OLS	WLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VRP_t	4.485*	4.030*	3.931*	3.333*	2.542	2.396*	5.497***	4.210**	3.874**
	(1.85)	(1.82)	(1.80)	(1.65)	(1.31)	(1.84)	(3.15)	(2.18)	(1.96)
$VRP_t \times \hat{\beta}_{v,t}$		-0.751***	-0.843***		-0.883***	-0.973***		-0.579**	-0.660**
		(2.67)	(3.02)		(3.15)	(3.50)		(2.29)	(2.56)
Adj-R^2	0.028	0.054	0.062	0.017	0.051	0.060	0.056	0.073	0.079

Panel B: Prediction using weighted least squares

$\hat{\beta}_{v,t}$	VRP_N			$VRP_{\overline{P}}$			VRP_{PE}		
		OLS	WLS		OLS	WLS		OLS	WLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VRP_t	3.697*	3.276*	3.096*	3.279**	2.556	2.376	4.922***	0.380**	3.472**
	(1.91)	(1.76)	(1.68)	(2.23)	(1.51)	(1.41)	(3.13)	(2.26)	(2.01)
$VRP_t \times \hat{\beta}_{v,t}$		-0.052**	-0.626**		-0.663***	-0.740***		-0.497**	-0.579**
		(2.31)	(2.64)		(2.75)	(3.06)		(2.23)	(2.55)
Adj-R^2	0.014	0.027	0.033	0.012	0.038	0.038	0.034	0.046	0.050

*** denotes significance at 1%, ** at 5%, and * at 10% level.

Table 7 reports the R^2s, coefficients, and t-statistics for each of the predictive regressions run separately for subsamples.

		Classification			
	Number of months	High 113	Medium 103	Low 108	High-low
VRP_N	In-sample R^2	0.117***	0.047**	0.004	0.113
		(3.83)	(2.26)	(0.69)	
	Predictive beta (β_p)	11.441	5.054	1.427	
$VRP_{\overline{P}}$	In-sample R^2	0.131***	−0.003	0.007	0.124
		(4.11)	(−0.12)	(0.78)	
	Predictive beta (β_p)	10.017	−0.282	1.580	
VRP_{pE}	In-sample R^2	0.179***	0.001	0.027	0.152
		(4.91)	(0.36)	(0.71)	
	Predictive beta (β_p)	8.743	3.936	0.026	

*** denotes significance at 1%, ** at 5%, and * at 10% level.

Conclusion

The results show that the contemporaneous and predictive relations are linked in a very specific manner, such that the predictive beta depends on the contemporaneous beta.

Moreover, the predictive performance, measured by R^2, increases as the correlation between market returns and variance innovations becomes more negative.

6. Robustness

6.1 Alternative measures of the variance risk premium

This paper construct several other measures of VRP that have been used in previous research:

- consider the measure of Bollerslev et al. (2009) and denote this by VRP_{BTZ}
- consider the measure of Bekaert and Hoerova (2014):
$$\widetilde{RV}_t = \alpha_0 + \alpha_1 \widetilde{RV}_{t-1} + \alpha_2 VIX_{t-1}^2 + e_t \quad (27)$$
modify their original measure and let VRP_{BH} be the difference between the end-of-month value of VIX and the RV forecast of the above model
- $VRPV_{XO}$ denotes the case in which both option-implied variance (VXO) and high-frequency realized variance are estimated using the S&P 100 Index

Table 8 reports the key results of this paper using these alternative measures of the VRP.

Panel A: OOS performance of the traditional approach using alternative measures

	VRP_{BTZ}	VRP_{BH}	$VRP_{VXO,N}$		$VRP_{VXO,\bar{P}}$		VRP_{VXO,P^E}	
	1993–2016	1993–2016	1991–2016	1993–2016	1991–2016	1993–2016	1991–2016	1993–2016
OOS-R^2	0.037	0.024	−0.015	−0.009	−0.010	−0.007	0.011	0.016
Wald	2.176	0.367	0.233	0.082	0.233	0.082	0.186	0.339
p-value	(0.140)	(0.545)	(0.629)	(0.775)	(0.629)	(0.775)	(0.666)	(0.561)

Panel B: OOS performance of the contemporaneous beta approach

	Statistics	VRP_{BTZ}	VRP_{BH}	$VRP_{VXO,N}$		$VRP_{VXO,\bar{P}}$		VRP_{VXO,P^E}	
		1993–2016	1993–2016	1991–2016	1993–2016	1991–2016	1993–2016	1991–2016	1993–2016
One-month	OOS-R^2	0.050	0.073	0.068	0.070	0.083	0.086	0.084	0.087
WLS	Wald	2.245	3.877	5.993	5.723	9.145	8.781	6.906	6.666
	p-value	(0.134)	(0.049)	(0.014)	(0.017)	(0.002)	(0.003)	(0.009)	(0.010)
One-month	OOS-R^2	0.041	0.054	0.055	0.057	0.071	0.073	0.072	0.073
OLS	Wald	2.907	2.863	4.098	3.960	6.885	6.602	4.982	4.619
	p-value	(0.088)	(0.091)	(0.043)	(0.047)	(0.009)	(0.010)	(0.026)	(0.032)
Three-month	OOS-R^2	0.060	0.072	0.059	0.064	0.046	0.049	0.053	0.054
WLS	Wald	4.130	4.951	10.298	10.890	3.404	3.568	3.063	2.979
	p-value	(0.042)	(0.026)	(0.001)	(0.001)	(0.065)	(0.059)	(0.080)	(0.084)
Three-month	OOS-R^2	0.052	0.057	0.049	0.053	0.041	0.044	0.047	0.049
OLS	Wald	4.574	3.997	7.637	8.120	2.898	3.034	2.649	2.516
	p-value	(0.032)	(0.046)	(0.006)	(0.004)	(0.089)	(0.082)	(0.104)	(0.113)

Panel C: Conditional OOS performance (traditional approach, 1993–2016)

	OOS-R^2				
	VRP_{BTZ}	VRP_{BH}	$VRP_{VXO,N}$	$VRP_{VXO,\bar{P}}$	VRP_{VXO,P^E}
C-1: One-month correlations					
High	0.060	0.129	0.054	0.062	0.083
Medium	0.043	−0.057	0.045	0.011	−0.001
Low	0.002	−0.002	−0.124	−0.101	−0.049
High-low	0.058	0.131	0.178	0.163	0.131
C-2: Three-month correlations					
High	0.135	0.121	0.075	0.036	0.078
Medium	0.060	0.034	0.034	0.039	0.027
Low	−0.041	−0.040	−0.121	−0.089	−0.047
High-low	0.176	0.161	0.196	0.125	0.125

6.2 Alternative specifications for the traditional approach

		VRP_N	$VRP_{\overline{P}}$	VRP_{PE}
Panel A: The traditional approach				
WLS, ten-year rolling	OOS-R^2	0.016	0.005	0.050
1993–2016	Wald	0.319	0.026	1.806
	p-value	(0.572)	(0.872)	(0.179)
Five-year rolling	OOS-R^2	−0.046	−0.039	0.015
1993–2016	Wald	0.805	0.664	0.001
	p-value	(0.370)	(0.415)	(0.978)
Seven-year rolling	OOS-R^2	−0.030	−0.029	0.018
1993–2016	Wald	0.053	0.144	0.563
	p-value	(0.818)	(0.704)	(0.453)
Expanding window	OOS-R^2	0.015	0.005	0.057
1998–2016	Wald	0.112	0.016	1.488
	p-value	(0.738)	(0.899)	(0.223)
Panel B: The contemporaneous beta approach (no intercept)				
One-month	OOS-R^2	0.073	0.090	0.093
WLS beta	Wald	5.309	8.728	5.621
1998–2016	*p*-value	(0.021)	(0.003)	(0.018)
One-month	OOS-R^2	0.066	0.083	0.085
OLS beta	Wald	4.736	8.728	6.158
1998–2016	*p*-value	(0.030)	(0.003)	(0.013)

7. Conclusion

- This article shows that the slope that determines the contemporaneous relationship between market and variance risk resembles the relationship between the risk premium of the market and market variance. As a result, when the beta of the contemporaneous regression of market returns on changes in its variance is used as the predictive slope for the VRP, one-month market returns can be predicted in a statistically and economically significant manner, even out of sample.
- The predictive power strongly depends on the degree of the contemporaneous correlation between returns and variance innovations. When correlations are highly negative, predictions can be made more accurately. Since the predicted strength of the leverage effect can be estimated ex ante, we can anticipate this predictive power. The combination of the contemporaneous beta and the VRP outperforms the average returns consistently over time, regard- less of the strength of the asymmetry in the market.
- Although the VRP is constructed from option prices on the index as well as index returns, its ability to predict future returns is not necessarily restricted to the equity index.

Thank You!