Variance Risk in Aggregate Stock Returns and Time-Varying Return Predictability

Sungjune Pyun*

Job Market Paper
October, 2016

ABSTRACT

This paper introduces a new out-of-sample forecasting methodology for asset returns using the variance risk premium (VRP). While Bollerslev, Tauchen, and Zhou (2009) demonstrate that the VRP predicts market returns up to six months, I show that the out-of-sample performance is comparably weak. In contrast, my new approach produces an out-of-sample forecast that is both highly statistically and economically significant. Specifically, I find a monthly out-of-sample R-squared of 8% and a trading strategy that produces a 0.1 gain in the annual Sharpe ratio. This new approach is motivated by the 'beta representation,' which implies that the market risk premium is related to the price of variance risk by the exposure to variance risk (beta). Hence, empirically, 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. Also, when variance shocks explain a greater fraction of market returns, the predictions are more accurate. Furthermore, these results are extendable to the cross-section of exchange rates. The VRP of the S&P 500 Index only predicts returns of currencies that are highly correlated with changes in market variance. For those currencies, the new methodology offers considerable improvements in out-of-sample performance. These results suggest that the variance risk exposure is a key factor that determines whether and how returns are predictable by the VRP.

JEL classification: G10, G11, G12, G15 and G17.
Keywords: Variance Risk Premium, Leverage Effect, Return Predictability, Beta Representation

---

*Marshall School of Business, University of Southern California. First draft: August, 2015. This paper was previously under the title: "What the Variance Risk Premium Tells Us about the Expected Market Returns". I thank my advisor Christopher Jones for extremely thoughtful comments and continuous encouragement. I also thank Wayne Ferson, Scott Joslin, Arthur Korteweg, Juhani Linnainmaa, Selale Tuzel, Fernando Zapatero and seminar participants at the University of Southern California, AFBC (Sydney), EFA (Baltimore), and MFA (Atlanta) for helpful comments. All errors are my own.

I. Introduction

Whether market returns are predictable using public information is of interest to both practitioners and academics. Although studies show that a number of variables can forecast future market returns, several problems have also been observed. First, predictive relationships appear to change over time, with some variables being successful in certain periods (Fama and French 1988a) or at a specific periods of the business cycle (Dangl and Halling 2012). Second, predictors that perform well in sample often fail out of sample (Goyal and Welch 2008, Campbell and Thompson 2008). Lastly, return predictions typically perform worse for shorter horizons (Fama and French 1988a) with many well-known predictors failing to forecast returns at the horizon below six months. Statistical inference on long horizon predictions is unreliable, raising concerns that some findings could be spurious.[1]

However, a recent study by Bollerslev, Tauchen, and Zhou (2009) suggests that even monthly or quarterly market returns are predictable by the one-month variance risk premium (VRP), measured as the difference between option-implied variance and realized variance. They report a positive and statistically significant slope coefficient for the regression

$$R_{m,t+1} = \beta_{p0} + \beta_p VRP_t + \epsilon_{t+1}, \tag{1}$$

where $R_{m,t+1}$ is the leading excess market return. Commonly, the VRP is interpreted to be a proxy of time-varying aggregate risk aversion. For example, the VRP is considered to embed critical information about the moments of the stochastic discount factor (SDF) that is also useful in explaining variation in the risk premium of the market.[2]

This paper proposes a new out-of-sample approach to monthly return predictions using the VRP that performs well both in terms of statistical and economic significance. The new methodology is derived from two theoretical observations. First, the one-month market risk

[1]See, for example, Hodrick (1992), Stambaugh (1999), Ang and Bekaert (2007), and Pastor and Stambaugh (2009).

[2]See, for example, Todorov (2010), Drechsler and Yaron (2011), Bekaert, Hoerova, and LoDuca (2013), and Bekaert and Hoerova (2014), among others.

premium should be related to the VRP by the market's exposure to variance risk. This logic follows intuitively from what is known as the "beta representation," i.e., that the risk premium of an asset is related to the price of risk by the size of risk exposure. The second observation is that when variance risk is responsible for a larger fraction of market risk, the VRP should explain a greater share of the market risk premium.

The beta representation suggests that the risk premium of an asset can be represented as the product of the price of risk and the risk exposure. However, as both of these are typically latent, assessing the value of the risk premium is not straightforward. However, when we decompose market risk into a component that is related to variance risk - and another that is orthogonal to variance a fraction of the risk premium is estimable. Under this decomposition, the expected market returns $\mathrm{E}_t[R_{m,t+1}]$ can be expressed as

$$\mathrm{E}_t[R_{m,t+1}] = \beta_{v,t}\lambda_{v,t} + \lambda_{o,t},$$

where the $\lambda_{v,t}$ is the price of variance risk and $\lambda_{o,t}$ is due to orthogonal risk. Orthogonal risk can arise for multiple reasons such as dividend shocks, consumption shocks, productivity shocks, other news, or even noise. While the price of orthogonal risk cannot be easily evaluated, the first component, the risk premium due to market variance can be estimated.

The ability to directly estimate the contribution of variance risk to the market risk premium follows from three unique characteristics of the VRP. Firstly, unlike many other predictors that are merely related to the price of risk in an unknown manner, the VRP precisely measures the price of variance risk $(\lambda_{v,t})$. Moreover, the underlying risk factor, namely unexpected changes in market variance, is estimable relatively accurately using high-frequency data. Therefore, the market's exposure to variance risk $(\beta_{v,t})$ is also observable, albeit with some estimation error. Finally, variance risk comprises a large part of the variation in market returns. The high negative relation between market risk and variance risk is well known by what is called the "leverage" effect. A strong negative relation means that the observable component $(\beta_{v,t}\lambda_{v,t})$ is likely an essential element of the market risk premium.

One of the main results of this paper is that the slope coefficient of the monthly predictive regression of (1) can be estimated more accurately by the market's exposure to variance risk ($\beta_{v,t}$). This exposure can be estimated by the slope of the regression of market returns on the unexpected contemporaneous changes in realized variance (RV)

$$R_{m,t} = \beta_{v0} + \beta_v(RV_t - \mathrm{E}_{t-1}[RV_t]) + \epsilon_{o,t}. \tag{2}$$

In fact, in some respects this estimate is far superior to the slope obtained in the traditional manner in which the predictive regression (1) is estimated directly.

The traditional way of forming an out-of-sample forecast is by running the predictive regression of (1) on a rolling basis for a relatively long sample. The estimated coefficients of the predictive regression are then used further to form a one-step-ahead out-of-sample forecast. This traditional methodology relies on the assumption that the predictive relation remains relatively stable for an extended amount of time so that past values of the predictive slope provide a reasonably good approximation of the predictive relationship today. However, as noted in the first paragraph, studies suggest that the predictive relationship changes over time. To be adaptive to the time-varying predictive relation, we need a shorter estimation period. However, this may also e problematic since reducing the estimation period will increase the estimation error of the coefficients.

The new methodology of predicting monthly market returns out of sample I propose in this paper, is based on the close equivalence of variance risk exposure and the predictive slope. The new approach directly uses the contemporaneous variance beta ($\beta_{v,t}$) in place of the predictive beta(β_p). As both returns and the series of the realized variances are available at the daily frequency, the contemporaneous relationship can be estimated on a monthly basis using observations from the first to the last day of the month. The size of the slope can then be multiplied by the VRP calculated at the end of the corresponding month to form a return forecast for the following month.

This new methodology is potentially superior for several reasons. First, the regression of returns on variance innovations has a much higher R^2 compared to that of the traditional predictive regressions. Higher R^2s imply that the coefficients used for the out-of-sample predictions are estimated much more accurately. Moreover, the new approach only uses the most recent month of data to determine the parameters. Hence, the proposed out-of-sample forecast methodology is applicable even when economic conditions change rapidly over time, due to a much shorter estimation period.

The empirical analysis on market returns confirms the hypothesis that the market's contemporaneous variance beta is very close to the predictive beta. As anticipated, the predictive beta estimated from in-sample regressions decreases in the contemporaneous variance beta. On average, a single unit decline in the contemporaneous beta leads to an approximate increase of 0.7 units in the one-month predictive beta and of a single unit in the three-month predictive beta. Moreover, the predictive power of the VRP depends strongly on the degree of correlation between market returns and variance innovations. When correlations are highly negative, VRP-based forecasts explain a considerable share of future market returns. On the contrary, when correlations are close to zero, market returns are essentially unpredictable by the VRP. Since these correlations can be measured ex-ante, we can determine the predictive accuracy in advance. The in-sample predictive R^2 of one-month predictions ranges from 7.0% - 18.0% during periods when market returns and variance innovations are highly correlated, compared to 0.4%-0.6% when the correlations are close to zero.

The out-of-sample results substantially improves when we use the new approach. In particular, the new proposed approach predicts one-month market returns in a statistically and economically significant manner. Specifically, the out-of-sample forecasts formed in the traditional way, by running a series of rolling predictive regressions, are unable to produce accurate forecasts of one-month returns. Although some of the out-of-sample R^2s are positive (-1.0 to 4.4%) depending on how the VRP is measured, they are far from being statistically significant. However, when we forecast future returns using the contemporaneous variance beta of the market, the R^2s are always much higher (7.5%-10.4%) than those of the traditional approach. Moreover, the corresponding Wald statistics are statistically significant across every VRP measure considered. These results are robust regardless of whether a constant or zero premium on the orthogonal component is assumed. Finally, there is a gain of approximately

0.10 (18%) in annual Sharpe Ratios and 0.02 (40%) in Certainty Equivalent when forming a trading strategy based on the new proposed approach.

The analysis also shows that the out-of-sample R^2s are higher when the price and variance of the market move closely together. The R^2s of the traditional approach are 9.4% – 18.2% during months in which the price and variance move closely together, but decreases to -2.5% – -6.9% when they are unrelated. The gap between the high and low periods decreases when using the new proposed approach, but the R^2s are always higher when the contemporaneous correlations are high in magnitude. These results imply that the VRP provides more information about the market risk premium when returns and variance innovations move together.

Although the VRP is constructed from option prices on the index and the index returns, its ability to predict future returns is not necessarily restricted to the equity index. Rather, the VRP should simply predict returns that highly correlate with changes in stock market variance. Assets whose returns change more sensitively to changes in stock variance should be predictable by the equity VRP with a higher beta. This article tests this hypothesis using 21 different currencies. The empirical results confirm that the VRP-based predictions are more accurate for currencies that are more exposed to changes in stock market variance. While only the returns of several currencies are predicted well by the VRP of the equity index, those currencies are the ones that are most highly correlated with equity variance. Currencies whose returns are contemporaneously related to market variance with a higher beta are predictable, with a higher slope, by the VRP. Also, the predictive power is associated with the degree of contemporaneous correlations. Lastly, the out-of-sample R^2s are also higher for currencies that highly correlate with equity variance. In sum, the cross-sectional results confirm that it is the amount of market variance risk that leads to returns being predictable by the VRP.

II. Related Literature

The hypothesis that the predictive power of the VRP might be related to the negative correlation between returns and variance has also been suggested by Carr and Wu (2016). Bandi and Reno

(2016) investigate whether co-jumps in returns and volatility have a simultaneous impact on the market risk premium and the VRP. However, to my knowledge, this is the first paper that formally tests the hypothesis. Also, the relation between the predictive and contemporaneous betas has not previously been explored.

This paper is closely related to a stream of research that studies time-varying return predictability. For example, Henkel, Martin, and Nardari (2011) and Dangl and Halling (2012) find that the power of well-known return predictors is business-cycle dependent, and that the market is mainly predictable only during recessions[3]. Furthermore, Lettau and Van Nieuwerburgh (2008) argue that there may be shifts from the steady state, which makes the in-sample coefficients too unstable to use for out-of-sample forecasts. Johannes, Korteweg, and Polson (2014) also show that the predictive coefficients are time-varying. Based on the idea that the predictive relation changes over time, Rapach, Strauss, and Zhou (2010) suggest combining multiple predictors to form an optimal forecast. These papers mainly study the time-varying predictability of valuation ratios (e.g. dividend yields, P/D, P/E ratio), which are known to predict returns for a longer horizon. This paper focuses on the VRP, which predicts market returns over a shorter horizon.

There are other articles that use information from the options market to evaluate the equity risk premium. For instance, Martin and Wagner (2016) claim that a combination of index option implied variances is a tight lower bound of the equity risk premium. Bollerslev, Xu, and Zhou (2015) consider predicting dividend growth and the equity premium jointly, using the VRP as one of the predictors. Also, option prices of different underlying assets are often used to predict asset returns of different classes. Londono (2014) and Bollerslev, Marrone, Xu, and Zhou (2014) study the predictive power of the VRP in an international context. Londono and Zhou (2014) build a variance risk premium measure from currency options and show that both the equity VRP and the currency variance risk premium is a determinant of the cross-section of stock returns. Applications of the variance risk premium to different asset markets also include studies by Wang, Zhou, and Zhou (2013) (credit default swaps) and Choi, Mueller, and Vedolin

[3]See also Garcia (2013), Chen (2009), Lustig, Roussanov, and Verdelhan (2014), and Cujean and Hasler (2015) among others.

(2015) (bonds). Returns from various assets are often predicted by the same underlying asset of options in which the risk premium is computed. However, the present paper suggests that this need not be the case, as long as the corresponding asset is exposed to variance risk of the U.S. stock market.

It is common to use downside risk as a return predictor. This paper is potentially related because the variance beta is related to market skewness, as they are different transformations of the third moment. For example, Kelly and Jiang (2014) propose a downside risk measure that predicts market returns. Feunou, Jahan-Parvar, and Okou (2015) and Bollerslev, Todorov, and Xu (2015) suggest that it is downside risk portion of the variance risk that mostly contributes to return predictability. Carr and Wu (2016) propose an alternative measure of the VRP using an option volatility surface that better predicts market returns. Bekaert and Hoerova (2014) decompose the VIX index into two parts, one that predicts market returns and the other that proxies for financial instability. Chen, Joslin, and Ni (2016) argue that increased financial intermediary constraints, measured using trading activities of deep out-of-the-money puts, lead to a higher risk premium.

III. The Variance Risk Premium and the Expected Market Returns

In stochastic discount factor (SDF) representation, the risk premium on one-month integrated market variance, the so-called variance risk premium (VRP), can be expressed as[4]

$$\begin{aligned} VRP_t &= \mathrm{Cov}_T\left(SDF_{T,T+1}, \int_T^{T+1} dV_t\right) \\ &\approx \mathrm{E}_T^Q\left[\int_T^{T+1} dV_t\right] - \mathrm{E}_T\left[\int_T^{T+1} dV_t\right], \end{aligned} \tag{3}$$

[4]Note the "+" sign in front of the SDF. Although the VRP can be defined as the difference between the real-world measure and the risk-neutral measure, I follow the sign convention of Bollerslev, Tauchen, and Zhou (2009). Also, the approximation comes from ignoring the effect of risk-free rates.

where $SDF_{T,T+1}$ is the integrated SDF over the same one-month horizon from T. The risk-neutral expectation is commonly measured using the square of the Volatility Index (VIX), available from the Chicago Board Options Exchange (CBOE). The VIX is the standard deviation of S&P 500 Index returns under the risk-neutral measure, computed using a model-free approach. The index is interpolated so that it matches the expectation of one-month integrated variance. Consequently, estimating the real-world expectation component over the same horizon is a natural choice. Notably, the VRP is the negative of the unit price of variance risk. Since option-implied variance is typically higher than the realized variance, according to the above definition, the VRP is positive for the most of the sample.

In a recent paper, Bollerslev, Tauchen, and Zhou (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,

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

and report a positive and statistically significant β_p. They provide a model that suggests that the VRP may proxy for parameter uncertainty. Drechsler and Yaron (2011) further argue that the VRP is a measure of economic uncertainty. Other studies also commonly interpret the VRP as a gauge of time-varying risk aversion (Todorov 2010, Bollerslev, Gibson, and Zhou 2011, Bekaert, Hoerova, and LoDuca 2013).

Besides being a short-term predictor, the VRP is a unique predictor of market returns for several reasons. One is that it is actually a price of risk, rather than a variable that merely encodes information about risk prices (e.g. the dividend yield) in an unknown manner. Second, the factor on which it is based, namely variance innovations, can be observed with a tolerable amount of estimation error. Most importantly, those variance innovations are highly correlated with market returns, which means that the market has substantial exposure to variance risk.

Prior research often ignores the important aspect of the equity market, namely that the market variance and price tend to move in the opposite direction. One well-known explanation, known as the "leverage" effect (Black 1976, Christie 1982), hypothesizes that an adverse shock in the market causes the overall leverage to increase, leading to higher volatility. Hence, the level of variance and price are negatively related. A more popular explanation, known as "volatility feedback," is that risk-averse investors require a higher premium for being in a high-volatility state. Therefore, investors demand more compensation in the future when market variance increases unexpectedly. Since a higher risk premium implies lower values today, prices must drop when variance increases (Pindyck 1984, French, Schwert, and Stambaugh 1987, Bollerslev, Litvinova, and Tauchen 2006).

This negative relationship implies that the market portfolio is subject to variance risk, which naturally suggests that VRP is the premium on an important source of aggregate variation in the stock market, which would affect the required return of the market. Moreover, since the VRP can be measured relatively accurately, a fraction of the market risk premium can be inferred from the VRP, essentially in real time. Specifically, the market risk premium would depend on the price of variance risk by a slope which measures the risk exposure. The simple model provided below demonstrates this relationship.

1. A Simple Model

To build some intuition, consider a stochastic volatility model in which the correlation $\rho_t < 0$ between market returns and changes in market variance is assumed to be time-varying. When S_t is the price of the aggregate market portfolio, approximately represented by the S&P 500 Index, and V_t is its variance, we see

$$\frac{dS_t}{S_t} = \mu_t dt + \sqrt{V_t}(\rho_t dW_t^v + \sqrt{1-\rho_t^2} dW_t^o) \tag{5}$$

$$dV_t = \theta_t dt + \sigma_v dW_t^v. \tag{6}$$

By construction, the two Brownian motions dW_t^v and dW_t^o are independent. These processes assume that the return and variance process follow a bivariate Gaussian process with a negative correlation. The drifts are not specified but are assumed to be time-varying. The volatility of the variance is assumed to be constant, but can be time-varying. Solving the first equation in terms of variance innovations and dW_t^o yields

$$\frac{dS_t}{S_t} = \mu_t dt + \rho_t \frac{\sqrt{V_t}}{\sigma_v}(dV_t - \theta_t dt) + \sqrt{(1-\rho_t^2)V_t}\, dW_t^o. \tag{7}$$

This two-factor structure indicates that market movements can be decomposed into two parts. First, market prices can vary as market variance moves. According to volatility feedback, or as in the inter-temporal model of Merton (1973), an unexpected increase in the market variance must directly lead to lower returns. The second part reflects price movements due to all other reasons. By rotational indeterminacy, one can always transform the other sources of variation into a variable that is orthogonal to the first one. For simplicity, I refer to the second term as the 'orthogonal' component and the premium associated with this component as the orthogonal premium.

The SDF representation can be used to match the one-month market risk premium with the one-month VRP.

$$\begin{aligned}\mathrm{Cov}_T\left(-SDF_{T,T+1}, \int_T^{T+1} \frac{dS_t}{S_t}\right) &= -\rho_T \frac{V_T}{\sigma_v}\, \mathrm{Cov}_T\left(SDF_{T,T+1}, \int_T^{T+1} dV_t\right) \\ &\quad - \sqrt{(1-\rho_T^2)V_T}\, \mathrm{Cov}_T\left(SDF_{T,T+1}, \int_T^{T+1} dW_t^o\right). \end{aligned} \tag{8}$$

Equation (8) represents the relation between the two premia in continuous time and is the key to understanding my approach. The left-hand side equals the market risk premium over a single month period and can be decomposed into a linear combination of two prices of risk. The first governs how the one-month VRP relates to the market risk premium. Notably, the size of the slope that connects the VRP to the market risk premium ($-\rho_t \frac{V_t}{\sigma_v}$) is the negative of the market's exposure to variance risk. It is also the negative beta of the regression in which market returns are regressed on contemporaneous variance innovations. Thus, the relation between returns and unexpected changes in variance determines the slope that connects the VRP and *future* market returns.

In fact, the slope measures how the market responds to unexpected changes in market variance. From the perspective of an investor who wants to reduce exposure to variance risk, this term represents the part of stock return variance that can be hedged in the variance swap market. A variance swap exchanges future realized variance for a notional amount. Carr and Wu (2009) show that the risk-neutral expectation of variance is the notional amount of the swap. Therefore, the VRP is essentially the expected unit cost of variance risk, and the contemporaneous beta is the number of swap contracts required to hedge against exposure to variance risk.

The second term of the decomposition in (8) represents how the price of orthogonal component relates to the market premium. The orthogonal premium could potentially affect how the VRP and the market risk premium are related, because the VRP and the orthogonal premium can also be related. For example, an increase in aggregate risk aversion could both affect the VRP and the orthogonal premium simultaneously. Therefore, depending on the degree to which the two premia are linked, it is possible that the orthogonal premium modifies how the VRP relates to the market risk premium.

There are at least two reasons to believe that orthogonal risk is largely unrelated to the VRP. First, while the predictive power of the VRP decreases as the forecast horizon increases, the opposite is true for other well-known predictors, such as the dividend yield (Fama and French 1988a), E/P ratio (Campbell and Shiller 1988), term spread (Campbell 1987), and *cay* (Lettau and Ludvigsen 2001). Moreover, these predictors tend to perform well during recessions. For example, Rapach, Strauss, and Zhou (2010), Henkel, Martin, and Nardari (2011) and Dangl and Halling (2012) demonstrate that the predictive power is strong only during recessions. As will be shown in the following section, these periods do not coincide with periods in which is a strong negative relation between market returns and market variance, which is when the VRP has its strongest predictive power.

Although not entirely realistic, there are at least two circumstances where the linear relationship between the VRP and the market premium would be perfect. Under these assumptions, the slope that governs the relation between the premia is exactly the same as the slope that connects returns to variance innovations. The first case is when the orthogonal component is unpriced. The other is when the orthogonal premium is uncorrelated to the VRP.

2. Empirical Implications of the Model

The remainder of this paper thoroughly discusses several important aspects of the simple model presented. First, 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. We can run a contemporaneous regression of the daily excess market returns on the unexpected change in realized variance (RV) over a fixed interval as

$$R_{m,t} = \beta_{c0} + \beta_v(RV_t - \mathrm{E}_{t-1}[RV_t]) + \epsilon_{o,t}.$$

The slope of this regression measures how much the market reacts to unexpected changes in market variance. This is also the coefficient on the VRP in (8). The equation that relates the VRP to the market risk premium can then be represented as

$$\mathrm{E}_T[R_{m,T+1}] = -\beta_{v,t} VRP_T + O_T,$$

where O_T comes from the premium due to the orthogonal component[5]. The orthogonal premium is theoretically equivalent to $\sqrt{V_T(1-\rho_T^2)}\ \mathrm{Cov}_T(SDF_{T,T+1}, \int_T^{T+1} dW_t^o)$ from (8).

Second, the equation that describes the relation between the VRP and the market risk premium suggests that we can predict market returns more accurately when the index and the variance of returns move closely together. The model indicates that the proportion of the total market variation related to variance risk is ρ_t^2. 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, the

[5]The relationship between the one-month market risk premium and the one-month VRP can also be roughly observed by taking expectations on both sides of the contemporaneous regression with respect to the risk-neutral measure and the real-world measure, and subtracting one from the other. The one-month risk premium is exactly the sum of the premium on the two components.

orthogonal premium will create a bias when the VRP is used as a single predictor. 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 $\pm$ 1, the VRP should almost entirely identify the market risk premium. One way to understand this is to consider variance swaps on the market portfolio. A variance swap can perfectly hedge the market portfolio. Under no arbitrage, a perfectly hedged position should not generate anything more than the risk-free rate. Conclusively, since ρ_t^2 is the R^2 of the contemporaneous regression, the predictive R^2 should depend on the R^2 of the contemporaneous regression.

The link between the predictive and the contemporaneous relations is extremely useful when forecasting market returns out of sample. As discussed above, accurate predictions of market returns are extremely hard. In-sample R^2s rarely exceed 5% for most common predictors (Goyal and Welch 2008) at the annual horizon. A low R^2 also implies that the parameter estimates are likely to be more inaccurate, which induces a poor out-of-sample forecast. To compensate for the high percentage of noise in returns data, we need to rely on an extended estimation sample. However, this is only possible when we assume that the predictive relation remains constant. However, recent research suggests that the predictive relationship does not remain constant over time. As noted above, for valuation ratios such as the dividend yields, the predictive power is particularly higher during recessions. Also, Johannes, Korteweg, and Polson (2014) argue that the parameters that govern the predictive relation are time-varying. As the traditional way of using predictive regressions assumes a constant predictive relationship, the forecasts can be particularly misleading when the relation varies rapidly over time, or when there is a structural break in the economy. The econometrician may attempt to address parameter instability by using a short estimation window. However, when the first-stage estimation period is too short, as mentioned, the coefficients of the first-stage predictive regression will be imprecise.

There is an alternative approach that can be used specifically for the VRP, which I refer to as the "contemporaneous beta approach." The method is based on the close relation between the predictive and contemporaneous betas, and implemented by using the beta of the contemporaneous regression in place of the predictive beta to form an out-of-sample forecast of market returns. There are two reasons why the contemporaneous beta could potentially be a much more accurate estimate than the predictive beta of a rolling-window predictive regressions. First of all, the contemporaneous relation between returns and changes in variance is much stronger than the predictive relationship between the VRP and future returns. While the predictive R^2s hardly exceed 5%, the average of the R^2s of the contemporaneous regression is slightly above 15%. Furthermore, both returns and estimates of realized variance are available at the daily frequency. Being able to use data at a higher frequency implies that there is more data, which no longer necessitates relying on an extended estimation period. Using a short estimation period, such as a month, might be enough to generate a slope coefficient that is sufficiently accurate to form an out-of-sample forecast. Depending on what we assume for the orthogonal premium and variance forecast, it is essentially possible to get an estimate of the monthly equity premium using a single month of data. Hence, the new proposed approach can be used even when the predictive relation changes rapidly over time.

Finally, a similar logic should also apply to asset classes other than the equity index. It is not necessary that we use a VRP that is based on the asset whose returns we are trying to predict. As long as those returns are correlated with S&P 500 Index volatility, the VRP should have some predictive power. Assets that are highly correlated with variance innovations should be predictable with higher accuracy, and those that relate to changes in market variance with a higher beta should be predictable using the VRP with a higher slope. In the empirical section of this paper, I test whether this hypothesis holds for exchange rates.

One may ask whether there are any other risk premia for which we have a relative good estimate of the price of risk, and for which we can estimate the exposure to its risk relatively accurately. A potentially related application is to the bond market since inflation rates are observable, inflation risk directly affects bond prices. Moreover, the price of inflation risk can be estimated using Treasury Inflation Protected Securities (TIPS). However, as noted by d'Amico, Kim, and Wei (2014), TIPS is largely impacted by the liquidity risk premium, which makes estimating the risk premium challenging. The application to the inflation risk premium is beyond the scope of this paper.

The next section examines supporting empirical evidence that shows that the contemporaneous and predictive relations for market returns are in fact closely connected.

IV. Data and Estimation

Theoretically, the one-month VRP (of the S&P 500 Index) is the difference between the risk-neutral expectation and the real-world expectation of one-month return variance. Although it is relatively straightforward to infer the risk-neutral counterpart using the square of the VIX, estimation of the real-world expectation is model-dependent and subject to specification error. Moreover, there might be a mismatch, for example, when we use the monthly averaged value for one and the end-of-the-month value for the other. This could be especially problematic when the market volatility is trending during a month. This section discusses how we measure the VRP and the contemporaneous betas (denoted by $\hat{\beta}_v$) and correlations (denoted by $\hat{\rho}$) from daily returns and realized variances.

1. Forecasting Variance

For the purpose of estimating the VRP selecting a good variance forecast model is important. As Bekaert and Hoerova (2014) argue, the VRP's ability to predict returns may depend on the particular model used to compute the real-world expectation component. I use intra-day, high-frequency, return-based RV to model the forecasts. High-frequency RV models have advantages over standard generalized autoregressive conditional heteroscedasticity (GARCH) or stochastic volatility (SV) models, which typically rely on daily returns. First, traditional GARCH or SV models are somewhat difficult to estimate. Distributional assumptions are required for either model. Moreover, RV-based models are known to outperform standard GARCH or SV models when forecasting variance. This outperformance is partly possible because high-frequency data enables us to make a more accurate measurement of the latent variance process. Finally, using RV allows us to fit complex multivariate models that capture the long memory feature of the latent variance process.[6]

[6] Under several conditions, Andersen, Bollerslev, Diebold, and Labys (2001) show that the RV converges in probability to the true variance. Andersen and Bollerslev (1997) indicate that variance estimate based on intraday returns provide information about long-run volatility dependencies. For advantages using high frequency based RV see Andersen, Bollerslev, Diebold, and Labys (2003). Also see Andersen, Bollerslev, and Diebold (2007), Chen and Ghysels (2011), and Busch, Christensen, and rregaard Nielsen (2011) among others for details about the performances and extensions of the HAR-RV type models.

This paper obtains the RV of the S&P 500 Index from a combination of versions 0.1 and 0.2 of the Oxford-Man Institute's Realized Library (Gerd, Lunde, Shephard, and Sheppard 2009), which is publicly available. Version 0.1 covers 1996 - 2009 while version 0.2 covers the post-2000 period. The first part of 0.1 (1996-1999) is added to the 0.2 version to create the daily series of RV between 1996 and 2015. This variable is computed by first calculating squared log returns from the last tick of each five-minute interval. A subsampling scheme at one-minute intervals (Zhang, Mykland, and Ait-Sahalia 2005) is used to reduce microstructure noise. Hansen and Lunde (2006), for example, study the impact of subsampling and note that subsampling, theoretically, is always beneficial in reducing microstructure noise. I rescale the RVs so that they match the variance of a month.[7]

The constructed RV series is then used to compute the variance forecasts. Corsi (2009) proposes a Heterogeneous Autoregressive Realized Volatility (HAR-RV) model. The model assumes that the predicted value of volatility is linear in its autoregressive components – daily, weekly and monthly realized volatility. By distinguishing a long-run monthly component from the short-run daily component, the model performs well in capturing short-term variation in the volatility process together with the long memory feature of volatility. I use the variance version of the model to form a daily forecast of the variance and refer to it as the HAR-VAR Model. The variance version of the model suits this paper because the focus here is on variance. The results, however, do not depend on the particular model used, so using the volatility version does not change any of the findings of the paper, as shown in the robustness section.

The market variance on day $\tau + 1$ can be forecasted on day τ using past values of RVs by running the following regression:[8]

$$RV_{\tau'} = a_0 + a_d RV_{\tau'-1} + a_w(\sum_{j=1}^{5} RV_{\tau'-j}) + a_m(\sum_{j=1}^{22} RV_{\tau'-j}) + \phi_{1,\tau'} \tag{9}$$

[7]Adding or removing overnight returns has a small impact empirically.

[8]To avoid confusion, time τ indicates that a variable is in daily frequency. On the other hand, time t subscript implies that a variable is in monthly frequency.

where $\tau' = \tau, \tau-1, \ldots, \tau-L$. The one-day forecasts of the daily RV ($\widehat{RV}_{\tau}^{D}$) can then be constructed using the loadings on the daily, weekly and monthly components. I use six-month rolling regressions to account for the possibility that the forecast relation changes over time.

Similarly, the monthly total variance forecast can be constructed by running 22 separate regressions,

$$RV_{\tau'+k} = a_0 + a_d RV_{\tau'-1} + a_w(\sum_{j=1}^{5} RV_{\tau'-j}) + a_m(\sum_{j=1}^{22} RV_{\tau'-j}) + \phi_{1,\tau'+k}, \tag{10}$$

for k=0, ..., 21, and averaging them over the 22 forecasts.[9] Similar to the daily RV forecast, the monthly integrated variance forecast ($\widehat{RV}_{\tau}^{M}$) can be constructed from daily rolling regressions of 36 months. I use a longer estimation horizon to capture the long term movement of variance.

2. Estimation of the VRP

The VRP is measured by taking the difference between the VIX^2 and the monthly forecast of realized variance. Typically the variance forecast is implemented by using a variant of AR(1) model of the realized variances computed over the entire month. [10] However, this approach suffers from a timing mismatch, especially when variance is trending during the most recent month. While the VIX^2 reflects changes in market variance during a month, the above forecast of monthly variance does not account for possible changes in market variance during the most recent month. Even a small trend in market variance may have a significant impact on the VRP estimate as a return predictor of short-term market returns, such as for a single month.

Thus, I use the difference between the end-of-month values of the VIX^2 and the monthly total variance forecast as the primary measure of the VRP. Both the VIX and the RV forecasts are rescaled so that they match the one-month variance. If the month t consists of m(t) trading

[9] Note that this is equivalent to running one predictive regression with the sum of the 22 dependent variables, but the difference is what we can use in a single regression. On day τ, the method I use uses all past RVs up until day $\tau - 1$, while having the sum of 22 RVs can only use all data up to $\tau - 22$.

[10] See, for example, Bollerslev, Tauchen, and Zhou (2009), Bekaert and Hoerova (2014), and Gonzalez-Urteaga and Rubio (2016) among others.

days so that $t_{m(t)}$ refers to the last trading day of the month, the monthly forecast-based VRP, denoted by VRP_M of month t, is given as

$$VRP_{M,t} = \frac{VIX^2_{t_{m(t)}}}{12} - \widehat{RV}^M_{t_{m(t)}}.$$

A trending variance during the month does not bias this estimate, because the HAR-RVAR model does reflect changes in market variance during the final day and final week as well as over the month.

Bekaert and Hoerova (2014) argue that the predictability of returns by the VRP heavily depends on the variance forecast model used. To ensure that the empirical results are not dependent on a particular model, I supplement the primary measure with a non-parametric one. Denoted by VRP_N, the non-parametric VRP is the difference between the scaled VIX^2 and the daily RV, both averaged over the entire month. This non-parametric VRP is similar to the non-parametric measure of Bollerslev, Tauchen, and Zhou (2009), who take the difference between the end-of-month value of the implied variance and the monthly realized variance. Here, I use the monthly average for both the implied and the realized variance to avoid the possibility of volatility trends affecting the VRP.

Finally, a daily forecast-based VRP (VRP_D) is created directly from the HAR-VAR model. The daily forecast-based VRP is defined as

$$VRP_{D,t} = \frac{1}{12} VIX^2_{t_{m(t)}} - \widehat{RV}^D_{t_{m(t)}},$$

where $t_{m(t)}$ is the last trading day of month t. Similar to the monthly forecast-based measure, both the risk-neutral and the real-world components are measured at the end of the month. This measure also has some appeal because HAR-VAR type models as a predictor for the future variance have been extensively studied and shown to dominate other models in predicting variance(Andersen, Bollerslev, and Diebold 2007, Andersen, Bollerslev, and Huang 2011).

Admittedly, the last VRP is not a perfect one when the variance term-structure is not flat. While the VIX measures the expectation over the next one-month period, the second component is measured using a one-day forecast. A mismatch can arise, for example, when there are more than two components that determine the level of the volatility. To be specific, there is evidence that volatility can be decomposed into two parts, one which is highly persistent (long-run volatility) and the other that is less persistent (short-run volatility). Engle and Lee (1999) propose a component GARCH model that accounts for both a long-run and short-run component of volatility. For instance, when the variance is expected to stay high for a short time on a daily basis, but when there is strong mean-reversion on the monthly basis, the daily forecast may not be equivalent to the monthly forecast. However, showing that the empirical results even hold using this alternative daily forecast-based measure will ensure that findings are robust.

Figure 1 compares the three VRP measures used in this paper. The daily-based VRP has an outlier in 2003, possibly driven by a short estimation horizon. Consistent with previous studies, there is a negative spike during the Financial Crisis. This is because realized variance was unexpectedly high during these times and reverted to its original level swiftly. The VIX did not increases as much during that period because part of the spike was regarded as a jump in the index. Therefore, the forecast models did not appear to have captured the strength of the mean reversion that was observed.

3. Estimation of the Contemporaneous Betas and Correlations

The daily innovation of the variance is calculated by first computing the daily innovation in the realized variance, and rescaled to the value that matches the one-month interval. The monthly contemporaneous beta is estimated from the following regression, using only observations that belong to that particular month.

$$R_{m,\tau} = \beta_{0,t} + \beta_{v,t}(RV_\tau - \widehat{RV}^D_\tau) + \epsilon_\tau \tag{11}$$

The contemporaneous correlation (ρ_t) is the correlation between the left and right-hand side variables in the above equation – market returns and variance innovations – in the same period. Its absolute value is also equivalent to the square root of the coefficient of determination. Since each regression is based on observations from a single month, the result is a monthly series of betas and a monthly series of correlations estimated from non-overlapping samples.

4. Summary Statistics

Table I provides summary statistics for the key variables of interest. These variables include realized variance, option-implied variance, three measures of the VRP, the contemporaneous variance beta, and the contemporaneous correlation. The sample study period is from January of 1996 to December of 2015. There are a total of 240 months, 28 of which the NBER classifies as recession months. There are two recession periods, one in 2001 and the other from mid-2008 to mid-2009. The first two columns of the table summarize the means and standard deviations over the entire sample period.

The next four columns summarize the same statistics for several subsamples. I first divide the sample into two sub-periods, one in which the contemporaneous correlation is greater than and the other in which it is less than the median of the entire sample. The table only reports the statistics for the "greater" periods. If the correlations and the level of the VRP were related, this would indicate that what I show later captures a non-linear predictive relation of the VRP. However, as the summary statistics suggest, the level of the variance and the VRPs are similar regardless of whether the correlation is higher or lower than the median. The similarity of these results to the full sample tells us that both the correlations and the betas are unrelated to the VRP and the variance measures.

Recent studies suggest that market returns can be predicted better during recession periods. For example, Henkel, Martin, and Nardari (2011) and Dangl and Halling (2012) argue that the performance of traditional predictors, such as the dividend yield, is strong only during recessions. The set of predictors they consider does not include the VRP. To rule out the possibility that the contemporaneous correlation is not merely a proxy for business cycles, I ask whether the correlations are more negative during recession periods.

The next two columns of Table I provide descriptive statistics of variables during NBER recession periods. The statistics show that contemporaneous correlations are not apparently more negative during recessions. Two implications are worth noting. First, the statistics show that the findings of this paper are not implied by the work discussed above, in that I am not documenting a pattern in predictability that is related to business cycles.

The last column of the table summarizes the first-order serial correlations of these variables. Overall, the moderate level of these serial correlations suggests that these variables are stationary. The autocorrelation of the contemporaneous beta is slightly higher than 0.2, and that of the contemporaneous correlation is slightly above 0.1. Although the contemporaneous variables seem to be stationary, one may argue that the serial correlations of the variables are not high enough, suggesting that the estimates are mostly noise and do not provide much information about actual betas or correlations. To address this concern, I compute the transition matrices of the betas and correlations. Panel A and Panel B of Table II provides the transition matrices for the betas and correlations, respectively. I compute the first, second, and third quartile of the ex-post distribution of these contemporaneous estimates, where the first quartile contains values that are most negative. I then investigate the transition dynamics from one month to the next. Panel A shows that a highly negative beta this month is most likely to lead to a highly negative beta in the next month. Similarly, a beta that is positive or close to zero is also likely to remain so in the subsequent month. Testing for independence, using a Chi-square tests yields a p-value of 0.016. Conclusively, the beta of the current month does provide useful information about future betas.

However, whether the contemporaneous correlations are persistent is not as apparent from Panel B. Although a negative correlation does imply a highly negative correlation in the future, it seems that the opposite is not necessarily true. One of the reasons is related to the standard error of the estimates. The standard error of a correlation estimate increases as correlations converge to zero. Since each of the estimates uses approximately 22 observations, the 95% confidence interval is approximately $\pm$ 0.4 for estimates that are close to zero.[11] Testing for independence using a Chi-square test, I fail to reject the null hypothesis that there is an equal probability of transition to each quartile (p-value 0.12). However, despite the high amount of

[11]The standard error of the correlation is approximately $\frac{1-\rho^2}{\sqrt{n-1}}$. For details see, for example, Stuart and Ord (2010)

noise in these correlations, the transition matrix suggests that it is at least possible to identify periods when future correlations are likely to be highly negative. When removing quartile 4, where correlations are close to zero, the hypothesis of the transition probability being equal is rejected (p-value=0.03).

V. Empirical Results on Predicting Aggregate Stock Returns

When the market is more heavily exposed to market variance risk, a small change in the price of variance risk should affect the risk premium of the market by a higher amount. Moreover, the proportion of the market risk premium explainable by the VRP should depend on how much the two underlying risks correlate to each other. The first part of this section tests these hypotheses in the sample. The second part provides out-of-sample evidence.

1. In-sample Prediction Results

I first summarize the results of the classical predictive regression, replicating the results of Bollerslev, Tauchen, and Zhou (2009). Next, I examine properties of the predictive beta and discuss whether the predictive beta varies over time, and, if so, 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.

I analyze here whether we can predict one-month and three-month market returns. In particular, I focus on one-month returns because the VRP is the premium over a one-month horizon. Therefore, evaluating market returns at the same horizon is a natural choice, as suggested by the model. There are also other reasons for doing so. The contemporaneous relation between return and variance varies rapidly over time, as shown in Table I. Therefore, the contemporaneous relationship estimated based on past data may not be valid over a much longer horizon. Also, due to data availability, the intraday-based realized variances can only be computed starting in 1996. A longer sample would be required to evaluate longer horizon returns.

Bollerslev, Tauchen, and Zhou (2009) run a classical predictive regression of market excess returns on the VRP and find a positive and statistically significant predictive slope for horizons of two quarters or less. Table III replicates their findings at one-month and overlapping three-month horizons. Panel A provides the coefficients, t-statistics, and adjusted-R^2s for one-month returns, while Panel B provides the same statistics for three-month returns. Standard errors are adjusted for heteroscedasticity using the White (1980) correction for monthly returns, and for both heteroscedasticity and autocorrelation using the Newey and West (1987) correction with three lags for three-month returns.

The table shows that the predictive betas, across every measure of the VRP considered, are all positive and statistically significant. This is consistent with Bollerslev, Tauchen, and Zhou (2009). As reported in Table I, the average of the monthly contemporaneous betas is -0.042. The predictive beta in the table is 0.053 for the monthly variance forecast-based VRP measure, 0.047 for the daily variance forecast-based measure, and 0.051 for the non-parametric measure. Overall, the slopes of the predictive relations are very close to the average of the contemporaneous variance betas. For three-month returns, the predictive betas (in Panel B) are higher, in the range of 0.118 – 0.132, but these numbers are less than three times of the one-month predictive betas, implying that the predictability is stronger for shorter horizons. The in-sample adjusted-R^2 ranges from 3.0% to 4.9% for one-month returns and 7.3% - 9.9% for three-month overlapping returns. Overall, these numbers are comparable to those of Bollerslev, Tauchen, and Zhou (2009)

Most often, it is assumed that the predictive relation remains constant over time. However, in a dynamic economy, predictive betas may change over time. As noted before, for traditional predictors, previous work has found that the predictive power is higher for recessions. Similarly, predictive relations based on the VRP may also be time-dependent. The possibility of a structural break[12] makes using a constant coefficient predictive model even more problematic.

According to the model, we expect the predictive beta to be higher when the contemporaneous beta (β_v) is more negative. Since the contemporaneous relation is negative and becomes more negative as the market's exposure to variance increases, the predictive beta should get closer to zero as the exposure decreases. An interactive regression including the contemporaneous beta as an interactive variable can be used to test the relation. In the regression

$$R_{m,t+1} = \gamma_0 + \gamma_p VRP_t + \gamma_I VRP_t \times \hat{\beta}_{v,t} + \epsilon_{p,t+1}, \tag{12}$$

a negative and statistically significant interactive coefficient γ_I means that the predictive betas and the contemporaneous betas are negatively related, as hypothesized.

Panel B of Table III summarizes the regression coefficients, t-statistics, and the adjusted-R^2s of the interactive regressions. Across all measures of the VRP considered, the coefficient on the interactive variable is negative and statistically significant. We can therefore reject the hypothesis that there is no relation between the predictive and contemporaneous beta. On average, a single unit decrease in the contemporaneous beta corresponds to an approximately 0.7-unit increase in the predictive beta. However, I fail to reject the hypothesis that they are different from -1. For three-month returns, some of the coefficients are statistically insignificant, but the signs are all negative. A single unit decrease in the contemporaneous beta leads to a 0.8 to 1.2-unit increase in the predictive beta, approximately. In sum, this table suggests that the predictive betas are higher when the contemporaneous beta is more negative. This effect is stronger for one-month than for three-month market returns.

I 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 240 months in the full sample period of 1996-2015 is classified into one of three groups according to the monthly series of squared contemporaneous correlations between market returns and variance

[12] See, for example, Hamilton (1989), Ang and Bekaert (2002) and Dai, Singleton, and Yang (2007) among others

innovations. When the squared correlation during a particular month is lower than the first tercile of the distribution of past values, the month is classified as a low month. When the squared correlation is higher than the second tercile, it is classified as a high month. Otherwise, it is classified as a medium month. To avoid misclassification during the early years when there is not enough data for the benchmark distribution, I use the entire 1996-1999 year to classify the first 48 months. Note that a high month contains observations that have the most negative correlation since we are effectively classifying by the absolute correlation, which is monotonic in the R^2 of the contemporaneous regression, rather than the signed correlation.

Next, I run a constant coefficient predictive regression for each of these subsamples separately. Table IV reports the R^2s, coefficients, and t-statistics from each predictive regression. Panel A summarizes the results for one-month returns, and Panel B summarizes those for three-month returns. As in previous tables, standard errors are adjusted for heteroscedasticity in Panel A and for both heteroscedasticity and autocorrelation in Panel B. The R^2s are always high and statistically significant during high periods for both one- and three-month returns. On the contrary, during low periods, the R^2s are all insignificant for one-month predictive returns.

For one-month return predictions, the R^2s are between 7.0-18.0% during high periods, depending on how the VRP is measured, but they are only 0.4-0.6% during low periods. For three-month return predictions, the R^2s are in the range of 9.6-27.8% during high periods but only 0.3-4.6% during low periods. Although some of the R^2s are significantly different from zero for three-month returns during low periods, the R^2s are always much smaller than those of the high periods. Conclusively, this table suggests that there are times when predictability of the VRP is strong and other times when its predictability is non-existent. The predictive power depends on the relation between return and variance innovations.

In short, 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 absolute contemporaneous correlation between the market returns and the variance innovations becomes larger.

2. Out-of-sample Predictions

This section discusses the performance of out-of-sample (OOS) forecasts during the sample period of 2000-2015. The sample is shorter because the traditional approach needs at least several years of data to estimate the predictive betas. Several years of data is also required to estimate the variance forecast model used to compute the VRP and the variance innovations so that the results are truly out-of-sample.

As Goyal and Welch (2008) argue, significant in-sample prediction does not guarantee that predictions will perform well as a feasible trading strategy. To be a possible trading strategy, the predictor has to work out of sample. In examining an implementable strategy, this section documents two novel and striking findings. First, the beta that explains the predictive relationship is close to the negative of the contemporaneous beta. They are, in fact, so close that the contemporaneous beta can be directly used in place of the predictive beta. Second, the OOS predictions perform better when the contemporaneous correlation between market returns and variance innovations is highly negative.

The traditional approach to providing OOS forecasts of time T+1 returns consists of two stages. First, we run a predictive regression using the past k-months of historical data (from time $T-k+1$ up to time T) as

$$R_{m,t} = \beta_0 + \beta_p VRP_{t-1} + \epsilon_t. \tag{13}$$

We use the coefficient estimated at time T to forecast returns at time $T+1$. The one-step-ahead predicted value of the excess market returns ($\hat{R}_{m,T+1|T}$) is given as, $\hat{\beta}_{0,T} + \hat{\beta}_{p,T} VRP_T$.

The next step is to evaluate the OOS predictive performance, for example, using the OOS-R^2. To do so, Goyal and Welch (2008) and Campbell and Thompson (2008), among others, compute the OOS-R^2, defined as

$$1 - \frac{\sum_t (\widehat{R}_{m,t+1|t} - R_{m,t+1})^2}{\sum_t (\bar{R}_{m,t} - R_{m,t+1})^2},$$

where $\bar{R}_{m,t}$ is the historical average of the market returns up to time t. Finally, we compute a test statistic, for example, a Wald statistic, to test the significance of the predictor. Diebold and Mariano (1995) provide a formal test for such OOS prediction errors. Giacomini and White (2006) extend the OOS test and propose a Wald test that is valid for testing nested models. The Wald statistic is given as

$$W = n\left(n^{-1}\sum_t \Delta L_{t+1}\right)\hat{\Omega}^{-1}\left(n^{-1}\sum_t \Delta L_{t+1}\right),$$

where $\Delta L_{t+1} = \sqrt{(\bar{R}_{m,t} - R_{m,t+1})^2} - \sqrt{(\hat{R}_{m,t+1|t} - R_{m,t+1})^2}$ and $\hat{\Omega}^{-1} = \frac{1}{T}\sum_{t=1}^{T}(\Delta L_{t+1} - \overline{\Delta L})^2$. This Wald statistic follows a Chi-square distribution with degrees of freedom equal to the difference in the number of predictors.

Although this approach is reasonable, its performance may depend on which estimation period k the researcher chooses. The choice of k is particularly sensitive when predicting returns, since the R^2 of the in-sample predictive regression is low most of the time. For example, the R^2 of the single constant predictive regression reported in Table III is less than 7.6% for one-month returns. A low R^2 can be problematic when forming out-of-sample predictions since the standard errors of the regression will be high[13]. To have an accurate estimate, we need a higher estimation period (k). However, the horizon cannot be too long if the predictive relation is thought to change rapidly over time.

The approach I propose deviates in one critical dimension. First, the OOS forecast of month $T+1$ returns is formed by using the contemporaneous variance beta from month T in place of the predictive beta estimated over the past k periods. For the time being, I set the intercept of the predictive relation equal to zero. I term this the "contemporaneous beta" approach because it directly uses the contemporaneous betas ($\hat{\beta}_{c,T}$) estimated from regressions of returns and variance innovations. Recall that the market risk premium consists of two parts. The premium

[13]The standard deviation of the prediction error in a simple linear regression is,

$$MSPE = \sigma\sqrt{\frac{\sum(R_{m,t} - \hat{R}_{m,t})^2}{N}},$$

where N is the number of observations and $\sigma^2 = \mathrm{Var}(\epsilon_{p,t})$. The MSPE is equivalent to $\sigma\sqrt{(1-R^2)\mathrm{Var}(R_{m,t})}$.

that comes from the variance shock is equivalent to the product of the contemporaneous beta and the VRP. The premium that originates from the orthogonal shock may also be related to the VRP, but whether or how much it is related to the VRP is unknown. If the orthogonal premium is either unpriced or unrelated to the VRP, we expect the two premia to be related to the beta that measures the market's exposure to variance risk.

The first OOS forecast is formed by multiplying the VRP with the negative of the contemporaneous beta. The one-step-ahead predicted value of the excess returns is then,

$$\hat{R}_{m,T+1|T} = -\hat{\beta}_{v,T} VRP_T.$$

Since the betas are estimated from a single month of data, we only use the most up-to-date information on the market. There are several benefits of doing this. Above all, the R^2s of the contemporaneous regressions, estimated using daily data, are typically higher than those of the historical predictive regressions, estimated using monthly observations. This allows us to use a shorter time-period. Within a month, changes in economic conditions or structural breaks may only affect the coefficients by a marginal amount. Also, since the contemporaneous correlation is related to the strength of the predictive relationship, under the new approach we may only choose to use the information embedded in the VRP when the premium is likely to be more informative about the market risk premium.

It is worth noting that in the contemporaneous beta approach, the product of the negative variance risk exposure and the VRP predicts the excess market returns with a zero intercept. This is based on the assumption that the orthogonal component is unpriced. If the component is priced, forming OOS forecasts based only on the VRP could result in a poor forecast of the equity risk premium. It is even possible that the orthogonal premium is time-varying. The premium may be related to the VRP, as both the VRP and the orthogonal premium are prices of risk which will depend on aggregate risk aversion. This indicates that the relationship between the VRP and the market risk premium would change so that the contemporaneous beta provides a biased estimate of the predictive slope. There is also a possibility that the orthogonal premium can be explained by other well-known predictors of market returns, which would perform well for a longer horizon.

Therefore, I consider a third approach which is a combination of the contemporaneous beta and traditional approaches. I use predictors of market returns that are known to perform well. What I call the "hybrid approach" is designed such that the orthogonal premium is allowed to be a linear function of common predictors. To do so, after estimating the contemporaneous betas from a first-stage contemporaneous regression in each month, I run a second regression

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

to find estimates of $\hat{\delta}_0$ and $\hat{\delta}_1$ on a rolling basis. δ_0 and δ_1 are both intended to capture the orthogonal premium. Here, X_t can be any predictor of market returns, including the VRP. Under this approach, the OOS forecast at time T is

$$\hat{R}_{m,T+1|T} = -\hat{\beta}_{v,T} VRP_T + \hat{\delta}_0 + \hat{\delta}_1 \sqrt{1 - \hat{\rho}_T^2} X_T.$$

The forecasts of the hybrid approach are intended to incorporate the risk premium associated with the orthogonal risk component, in addition to the variance risk component. Since the variance risk component is well-captured by the product of the exposure and the price of variance risk, the role of this additional predictor is limited to explaining the orthogonal premium. The term in the square-root ensures that the predictors are weighted conditionally depending on the relative size of the orthogonal component to total market risk.

It is possible that the orthogonal premium is not well-captured by any of the predictors considered. There is also the possibility that the premium on orthogonal risk does not vary much over time or even remain constant. Hence, I consider a restricted case of the above approach, where δ_1 term is dropped. This particular case will be referred to as the contemporaneous beta approach with an intercept.

The OOS-R^2 is computed by comparing the performance of the one-step-ahead prediction using the model and the historical average. The predictive betas and the historical average of the market returns ($\bar{R}_{m,t}$) in the traditional approach may depend on how much data we use for the first-stage predictive regression. I use the entire sample available up to that point to compute the historical mean as a benchmark. The results do not change much if the estimation period for the historical mean is restricted to a shorter period, or when some of the pre-1996 sample is included. A seven-year rolling sample, ten-year rolling sample, or an expanding window is used to calculate the predictive betas in the traditional approach. An expanding window means that the predictions are made using all data up to that point.

Table V summarizes the OOS-R^2s and the Wald statistics, along with p-values, for the different methods discussed. When applying the traditional approach, one can observe that the VRP predicts market returns with a positive OOS-R^2. However, some of the R^2s are negative when only using seven years of data to estimate the coefficients, suggesting that 84 observations are simply not enough to form an accurate OOS forecast. When ten years of data is used to form a forecast, the OOS-R^2 increases to 4.4% for the main monthly forecast-based VRP. This number does not increase further by extending the estimation period beyond ten years. This is probably because the benefits of having a shorter horizon (time-varying coefficients) outweigh the benefits of additional data (smaller standard errors) beyond a certain point. For an estimation period of five years (not reported), the OOS-R^2s are highly negative, which confirms that at least 120 observations are needed to obtain a reliable estimate of the predictive beta. For other measures, the R^2s are also positive, although the magnitude is a little smaller. Importantly, however, despite having positive OOS-R^2s, none of these predictions is statistically significant, even at the 10% level. The Wald statistics are statistically insignificant, and we fail to reject the hypothesis that there is no predictability.

The numbers look much different when we use the contemporaneous beta approach. The second part of Table V provides the results for the new proposed approach. When we set the orthogonal premium to zero ("no intercept"), the OOS-R^2 for the monthly variance forecast-based VRP is much higher, at 10.4%. The numbers do not change much (9.6%) when the orthogonal premium is assumed to be a nonzero constant ("including intercept") These numbers are at least 5% higher than the best result of the traditional approach. Furthermore, the Wald statistics are now highly statistically significant. The prediction error is on average small, and it varies little over time. The results are similar for other VRP measures. These results suggest that using the new proposed approach, predicting one-month market returns in a statistically significant manner is possible, even out of sample.

The in-sample regressions in the previous section suggest that the R^2 of the predictive regression will be higher when the absolute value of the correlation between returns and variance innovations is high. This section evaluates whether the relations are preserved out of sample. The connection between the two correlations is tested by classifying the sample into one of three periods - low, medium or high - depending on whether the absolute value of the contemporaneous correlation is below or above its first and second terciles, as described in the previous section. Recall that the breakpoints or the terciles are formed on an out of sample basis for 2000-2015, which matches the OOS study period. Therefore, the classifications are made without any look-ahead bias. Table VI summarizes the OOS-R^2s for each of the subsamples. I provide only the OOS-R^2s and not the Wald statistics due to the short sample periods, so these results should be viewed somewhat informally.

Across every measure considered in this paper, and across all prediction methodologies examined, the OOS-R^2s are higher in the high periods than the low periods. That is, predictions can be made more accurately when volatility feedback is stronger, for both for the traditional and the contemporaneous beta approach. During the high periods, the R^2 is 17.3% when using the primary measure of the VRP in a traditional manner, while it decreases to -5.9% during the low periods. When using the contemporaneous beta approach without an intercept, the R^2 is 20.2% during high periods but 3.8% during low periods. These differences are substantial.

The traditional approach suggests that we should not use the VRP as a return predictor unless market returns are highly correlated with changes in its variance. During low correlation times, the traditional method of a running rolling predictive regression effectively overstates the role of the VRP by assuming a constant predictive coefficient. However, this is not the case for the contemporaneous beta approach. The new approach already encodes information about the relation between returns and market variance in the beta, so that we do not rely too much on the return predictor during these low correlation periods.

Thus, when market prices and variance move closely together, the VRP is a very powerful predictor of short-horizon market returns. On the other hand, when they move independently, it is hard to predict market returns using the VRP, since the market portfolio is less exposed to variance risk. Intuitively, during these times the price of orthogonal risk may play a greater role in determining the market risk premium. Hence, in the hybrid approach, I evaluate whether return predictions can be improved by allowing other known return predictors to determine the orthogonal premium. If a return predictor were to perform well in explaining variations in the orthogonal premium, then we might see increased predictive power by combining that predictor with the VRP. Moreover, the predictive power should be higher during medium or low correlation periods compared to the contemporaneous beta approach.

Other return predictors have the potential to complement the VRP for two reasons. First, the predictive power of the VRP is strong for monthly and quarterly returns. However, for other predictors, that power is higher for predictions of longer horizon returns (Poterba and Summers 1988, Fama and French 1988b). Second, the predictive strength of many common predictors tends to decrease for the post-1990 period (Goetzmann and Jorion 1993, Ang and Bekaert 2007, Goyal and Welch 2008). In contrast, the VRP has been demonstrated to be a predictor of market returns in the post-1990 period. Given these differences, I hypothesize that the other predictors may help explain the risk premium that arises from orthogonal risk.

I select several predictors that are well-known to predict market returns. These are: dividend yields (D/Y) (Campbell and Shiller 1988, Fama and French 1988a), the term (TERM) (Campbell 1987) and the default premium (DEF) (Keim and Stambaugh 1986), the short rate (Campbell 1987), the investment-to-capital ratio (I/K) (Cochrane 1991), *cay* (Lettau and Ludvigsen 2001), and NO/S (Jones and Tuzel 2013). All of these are known to perform well in predicting market returns over a long historical sample. For completeness, I also consider the left jump variation (LJV) from Bollerslev, Todorov, and Xu (2015) and the left tail risk from Kelly and Jiang (2014) as candidates that potentially affect the orthogonal premium. I also let the VRP itself explain variation in the orthogonal premium. This is important, because when the VRP may partly explain the market risk premium due to orthogonal risk, and if the orthogonal premium and the VRP are related, it will alter how the VRP relates to expected market returns. In this case, the contemporaneous beta would be a biased estimator of the predictive slope.

3. Explaining the Orthogonal Premium

Table VII provides the OOS predictive performance of the hybrid approach. The left three columns evaluate the performance for the post-2000 years. The first three rows repeat the results from Table V, where the orthogonal premium is assumed to be constant. These results will serve as a benchmark for determining whether the predictors help to explain the orthogonal premium. The right four columns summarize the R^2s for the high, medium, and low subsamples. The last column is the difference between the high and low periods. Only the R^2s for the monthly-based VRP is reported as the results for other measures are similar.

When the VRP is used additionally to explain the variation in the orthogonal premium, the R^2s increase slightly from the benchmark. However, the Wald statistics are lower, and the improvements are restricted to high absolute correlation periods. These results indicate that there is no strong evidence that the orthogonal premium is related to the VRP. The higher R^2 during high periods is primarily driven by the time during the Financial Crisis, where there was a negatively estimated VRP and a series of negative realized market returns.

Among the other ten predictors considered, only D/Y, TERM, I/K, and left tail risk will leave the OOS performance to be statistically significant. None of the R^2s are higher over the entire sample, although some predictors do appear to explain variation in market returns during low periods. When compared to the benchmark, DEF, the short rate, cay, I/K, NO/S, and left tail risk will help explain the market risk premium during medium periods and only D/Y and the left tail risk helps during low periods. It is possible that these variables contribute to predicting market returns due to an independent component that is orthogonal to variance risk.

4. Evaluating Economic Significance - A Trading Strategy

I also evaluate whether we can use the closeness between the two betas 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

$$w_T = \frac{\hat{R}_{m,T+1|T} - R_f}{\gamma \widehat{RV}_T^M}$$

where $\gamma = 3$ is assumed for the risk aversion coefficient. The remaining proportion $1 - w_T$ is invested in the risk-free asset. The weight in the market is capped at 200%. For all models, the same monthly forecast-based RV is used as a denominator of the portfolio weights. The certainty equivalent (CE) of the return is computed as

$$CE = \bar{R}_p - \frac{\gamma}{2} \mathrm{Var}(R_p),$$

where $\overline{R}_p$ and $\mathrm{Var}(R_p)$ are the sample mean and variance of the portfolio, respectively. I compare the Sharpe ratios and the certainty equivalent of various forecasts with the baseline, in which the historical mean is assumed to be the best predictor of future returns.

The previous tables on predictive performance show that predictions can be made more accurately when the absolute correlation between returns and variance is high. A concern is that the weights might rely too much on the VRP-based forecasts during periods when returns and variance innovations are unrelated. Therefore, I also 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

$$w_T = \frac{\hat{R}_{m,T+1|T} - R_f}{\gamma \widehat{RV}_T^M} \hat{\rho}_{c,T}^2 + \frac{\overline{R}_{m,T} - R_f}{\gamma \widehat{RV}_T^M} (1 - \hat{\rho}_{c,T}^2),$$

where $\hat{\rho}_{c,T}$ is the estimated contemporaneous correlation between index returns and variance innovations during month T. To distinguish this strategy formed on the conditional value of the correlation from the basic trading strategy, I call this the conditional trading strategy and the basic one as the unconditional trading strategy.

Table VIII summarizes the resulting gains/losses in the annualized Sharpe Ratios (SR) and percentage changes in the CEs. All numbers are all based on one-month returns. The left two columns compare the SR and CE of the unconditional trading strategy. The first row summarizes the values of the benchmark where the past average is used to form the weights on stocks and bonds. The SRs and the CEs are compared to the benchmark, and the relative gains or losses are reported. When predicting returns as in the traditional approach, there are no gains in the SR and CE. For the main measure of the VRP, there is a loss of 0.075 in the SR, and the CE decreases by 0.018. However, when using the contemporaneous beta approach, assuming that the orthogonal premium is zero, the SRs and the CEs all increase. The SR increases from 0.568 to 0.671 for the main measure of the VRP, and the CE increases by approximately 0.019. When the orthogonal premium is assumed to be a constant, both the SR and the CR increase but by only a small margin.

The right two columns provide results for the conditionally weighted portfolios. The same weights as in the unconditionally weighted portfolios, which is based on the historical average of market excess returns is used as a benchmark. As discussed, the weights are designed such that more information from the VRP is used as the size of the contemporaneous correlation increases. The first set of rows in the right two columns suggests that there is no gain in the SRs and the CRs for all but one specification, when making predictions in a traditional manner using rolling predictive regressions. However, when comparing the SRs and CEs to the unconditional weighting case, the conditional weighting outperforms the trading strategy based on the unconditional weighting. As summarized in the next set of rows, when using the contemporaneous beta approach, there is no big difference of forming the weights either conditionally or unconditionally. For the main measure, the gain in the SR is +0.103, is the same as the one based on unconditional weighting, as summarized in the left two columns. The gains in the CE are smaller for the conditional weighting portfolios. This is because the contemporaneous beta already incorporates the information about the contemporaneous correlations. When including an intercept, the conditional weighting actually helps increase the SR and the CE. This result suggests that the intercept is not accurately measuring the premium on the orthogonal part, but rather is only adding noise to the forecasts.

In conclusion, these results indicate that predictions under the traditional approach could be highly misleading during periods when returns and variance innovations are unrelated. During these times, investors appears to perceive variance risk to be unrelated to market risk. The VRP, therefore, provides little information about the market risk premium. On the other hand, 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 market's exposure to variance risk can replace and, in fact, is superior to the traditional predictive beta. The new proposed approach predicts the one-month market in both a statistically and economically significant manner.

VI. Cross-sectional Evidence from the Currency Market

It has been documented that the VRP predicts returns across different asset classes. For example, Gonzalez-Urteaga and Rubio (2016) study the VRP's influence on the cross-section of equity returns. Wang, Zhou, and Zhou (2013) document the predictability of credit default swap spreads using a variance risk premium. Londono and Zhou (2014) find that the VRP of stock returns as well as the variance risk premium of the currency returns predicts several currency returns, but not others. They claim that currencies whose returns are not predictable are ones that have a form of pegged exchange rates. In this section, I evaluate the predictive relation between the VRP and individual currency returns. The VRP may predict some currency returns because currencies of several countries are regarded as more risky, and are sensitive to time-varying aggregate risk aversion. However, it is also possible that there are currencies in which the level of risk is closely tied to the variance risk of the U.S. equity index. I study whether and how the currency's exposure to market variance risk is tied to the predictive relation.

In particular, I extend the study of Londono and Zhou and investigate 21 currency returns during the sample period of 1996-2015. These are: GBP (British Pound), AUD (Australian

Dollars), BRL (Brazilian Real), CAD (Canadian Dollar), DKK (Danish Krone), EUR (Euro), HKD (Hong Kong Dollar), INR (Indian Rupee), JPY (Japanese Yen), KRW (Korean Won), MYR (Malaysian Ringgit), MXN (Mexican Peso), NZD (New Zealand Dollar), NOK (Norwegian Krone), SGD (Singapore Dollar), ZAR (South African Rand), LKR (Sri Lankan Rupee), SEK (Swedish Krone), CHF (Swiss Franc), TWD (Taiwanese Dollar), and THB (Thai Bhat). Currencies were selected from those that were available on the FRED website at the Federal Reserve of St. Louis as of 2015. For the purpose of this study, currency returns are also attractive because, unlike the cross-section of individual stocks, currencies do not constitute the S&P 500 Index. Therefore, the results provided in this section cannot simply be mechanically driven by the predictability of index returns documented previously.

All currencies are quoted as the price of a unit of foreign currency per U.S. Dollar (indirect quotation). Under risk-neutrality, currencies are expected to depreciate in the future when their risk-free rate is relatively high. Otherwise, one could earn positive expected returns by borrowing from a country whose risk-free rate is lower, converting the currency, and investing in a riskless bond whose risk-free rate is relatively higher. This relation is known as uncovered interest rate parity (UIRP). In my analysis, currency returns are adjusted by the lagged difference in risk-free rates between two countries to account for the parity. This section evaluates whether the VRP explains these adjusted returns.

I perform the adjustment using the 3-month Treasury rate of each country, which I obtain from the central bank website of each county. Whenever the 3-month Treasury rate is not publicly available on these websites, I use the one-month interbank rate (EUR, HKD, MXN, MYR, SGD) instead. When neither of the rates is available, the central bank policy rate (BRL, JPY) is used as an alternative. Although these risk-free rates are not completely consistent, the effect of the adjustment is minimal since the marginal effect of risk-free rates is small for one-month intervals. All of the main results in this section remain intact even when returns are unadjusted.

A similar methodology to previous sections is applied to these adjusted currency returns. I first run a contemporaneous regression of the daily adjusted currency returns ($r_{i,\tau}$) on the variance innovations of the S&P 500 index returns.

$$r_{i,\tau} = \alpha_{i,t} + \beta_{i,v,t}(RV_\tau - E_{\tau-1}[RV_\tau]) + \epsilon_{i,\tau}$$

where $\tau = \{\tau_1^t, \ldots \tau_{m(t)}^t\}$ are trading days that belongs to a single month t. I run a regression for each month and obtain a monthly series of $\hat{\beta}_{i,v,t}$ and R^2s. Then, I compare these betas and the R^2s with the betas and R^2s of the predictive regressions. The predictive regression for each currency i is given as

$$r_{i,t+1} = \beta_{i,0} + \beta_{i,p}VRP_t + \epsilon_{i,t+1},$$

where VRP is the variance risk premium of the S&P 500 Index measured in one of the several ways discussed previously, and $r_{i,t+1}$ is the monthly return of currency i adjusted by UIRP. The predictive regression is estimated for each currency using the entire sample.

Table IX summarizes the results of the individual predictive regressions for the 21 currencies. Depending on how the VRP is measured, only 5-8 currencies are significantly predictable using the VRP. In particular, AUD, CAD, MXN, NZD, and ZAR can be predicted by every measure of the VRP considered. HKD, INR, JPY, KRW, MYR, SGD, LKR, CHF, and THB cannot be predicted using any of the measures. The next step is to compare these predictive betas with the average of the monthly contemporaneous variance betas. While there is a single predictive beta for each currency, the contemporaneous beta is a monthly time-series exist. I take an average of the monthly series of the contemporaneous betas to make it comparable to the predictive beta. Figure 2 plots the cross-sectional relation between the mean of the contemporaneous betas and the predictive betas. The plot shows that currencies that are more negatively exposed to market variance risk tend to have higher predictive betas. Across every VRP measure considered in this paper, there is a negative relation between contemporaneous and predictive betas.

This negative relation between the two betas can be tested by running a regression between the two betas. The solid line in the figure represents the regression line. It is well known that in OLS, measurement error in the independent variable creates a downward bias in the slope estimates. Since the contemporaneous beta can be estimated with higher accuracy, it is more appropriate to make predictive betas as dependent variable. The regression

$$\hat{\beta}_{i,p} = \lambda_0 + \lambda_1 \frac{1}{T} \sum_{t=1}^{T} \hat{\beta}_{i,v,t} + \xi_i$$

can be used to find the coefficients $\hat{\lambda} = [\hat{\lambda}_0 \;\; \hat{\lambda}_1]$. Although the interpretation of these estimates is simple and straightforward, it is clear that the OLS standard errors cannot be used. In the case of classical cross-sectional regressions in which we evaluate the relation between asset returns, it is likely that returns are cross-sectionally correlated, biasing the standard errors. Similarly, it is possible that there is cross-sectional correlation among the errors specifically because all currencies considered are quoted in terms of the U.S. dollar. For example, for some reason if the U.S. dollar depreciates in a high VRP period, all the predictive betas will be affected in the same direction. To control for possible cross-sectional error, the covariance matrix of the first stage predictive regression can be used to compute the covariance matrix of the cross-sectional regression. The formula is given as

$$\sigma^2(\hat{\lambda}) = \frac{1}{T}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\Sigma\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}, \tag{14}$$

where

$$\mathbf{B}' = \begin{bmatrix} 1 & \dots & 1 \\ \frac{1}{T}\sum_{t=1}^{T} \hat{\beta}_{1,c,t} & \dots & \frac{1}{T}\sum_{t=1}^{T} \hat{\beta}_{I,c,t} \end{bmatrix}$$

$$\Sigma_{i,j} = \frac{1}{T}(\mathbf{V}'\mathbf{V})^{-1}(\mathbf{V}'\mathbf{R}_i\mathbf{R}_j'\mathbf{V})(\mathbf{V}'\mathbf{V})^{-1},$$

$$\mathbf{V}' = \begin{bmatrix} 1 & \dots & 1 \\ VRP_1 & \dots & VRP_T \end{bmatrix}$$

and $\mathbf{R}_i$ is a T by 1 matrix of monthly currency returns[14].

[14]The result follows from the fact that $\Sigma_{i,j} = Cov(\hat{\beta}_{i,p}, \hat{\beta}_{j,p})$ is the robust covariance matrix of the cross-section of predictive betas. The betas are estimated from a regression of $\mathbf{R}$ on $\mathbf{V}$. Then (14) simply uses the covariance matrix of the predictive betas, which is the dependent variable of the cross-sectional regression, to derive the robust standard error.

Table X formally tests the relationship between the two betas. On average, currencies that have an average contemporaneous beta that is more negative by a single unit on average have a 0.6-0.7 units higher predictive beta. Across all VRP measures considered, t-statistics are highly statistically significant at the 1% level. This result shows that currencies that are related to contemporaneous market variance with a more negative beta can be forecasted using the VRP of the S&P 500 Index with a higher predictive beta.

I also compare the average R^2s of the contemporaneous regressions to the R^2s of the predictive regressions. Figure 3 illustrates the relation. The solid line represents the regression fit when the predictive R^2 is represented on the y-axis and the time-series average of the contemporaneous R^2 on the x-axis. The figure shows that the R^2s are also positively related. The equity VRP predicts currency returns that are more strongly correlated with contemporaneous changes in market variance.

Finally, I investigate whether the negative relation between the two betas also holds on an out-of-sample basis. The contemporaneous beta approach is used to form a one-step-ahead forecast. That is,

$$\hat{r}_{i,T+1|T} = -\hat{\beta}_{i,v,T} VRP_T,$$

where $\hat{\beta}_{i,c,T}$ is estimated from the contemporaneous regression of currency returns. OOS forecasts are typically evaluated relative to the historical average, as in the previous section. However, for currency returns, I compare the forecast with a theoretical value under risk-neutrality, in which the predicted value of the adjusted returns is zero. This is because currency returns are directly affected by monetary policies and comparing the UIRP adjusted returns with a zero based is much more reliable. For example, when the OOS-R^2s are computed using the historical mean of currency returns, the average of the OOS-R^2 of the contemporaneous beta approach is as high as 14%. This may seem as if there is a strong negative relation between the predictive and contemporaneous betas. However, rather than implying superior performance of the new approach, this result is primarily due to the poor performance of historical mean as a predictor. Comparing my approach to the predictive return of zero makes the findings I present a more conservative view of the strength of the contemporaneous beta method.

The upper panels (a and b) of Figure 4 compares the OOS-R^2s of the currency returns with the squared contemporaneous correlations between currency returns and changes in market variance over the 2000-2015 period. Using the traditional approach in (a), none of the currency returns can be predicted by the equity VRP out of sample. Furthermore, the relation between the R^2s and the squared contemporaneous correlations are only slightly positive. Using the contemporaneous beta approach in (b), more than half of the OOS-R^2s are positive. Moreover, returns that are predictable by the new proposed approach tend to have a high squared correlation between returns and changes in variance. The correlation between the OOS-R^2s and the squared contemporaneous (currency return-market variance innovations) correlations is 0.514 for the monthly-based measure of the VRP.

One might think that those currencies whose returns are more predictable are simply those that move more closely with the U.S. equity market. For example, if the rate of a particular currency closely tracks with the S&P 500 Index, it is natural that those returns would be predicted using the VRP since the VRP works well in predicting U.S. market returns. The lower panels of Figure 4 compare the relation between the OOS-R^2s and the square of the average correlation between currency and market returns. The column suggests that the equity VRP tends to predict well those currencies whose returns are highly correlated with market returns. However, if the contemporaneous variance beta is merely a proxy for exposure to the U.S. equity market, then the relation between R^2s should be stronger when the variance innovation is replaced by market returns (i.e. in Panels c and d rather than a and b) when computing the contemporaneous correlations. However, for both the traditional and the contemporaneous beta approach, the OOS-R^2s are more related to the R^2s of the “currency returns - variance innovations” relation than to the “currency returns - market returns” relation. This result would be hard to explain if the predictability of the currency returns were mainly driven by the close connection between currencies and the U.S. stock market.

VII. Robustness Tests

This section provides several robustness checks of the main results of the market return predictions. I show that the results of this paper are robust to using a predictive beta instead of the lagged realized beta. In fact, the key findings are even stronger. I also show that the results are robust to alternative variance forecast models and other measures of the VRP.

1. The Beta Forecast vs. Lagged Beta

This paper so far relies on the assumption that the level of variance risk in the market portfolio will persist, such that the lagged regression estimates are the best predictors of future relations. This section evaluates whether the key results of the paper are robust even when we fit a model to forecast the near-term betas using several state variables. I use an AR(2) specification to find a beta forecast. The predicted values are from a rolling predicted regression and the one-step-ahead forecast of this regression is denoted by $\hat{\beta}_{v,t|t-1}$.

Panels A and D of Table XI summarize the key results of the paper when $\hat{\beta}_{v,t|t-1}$ is used instead of the lagged beta estimates. The first two rows of Panel A provide the output of the in-sample interactive regressions, similar to those of Table III. The coefficients of the interactive variables are all negative and statistically significant. Compared to the main results, the coefficients are higher in magnitude. Panel D summarizes the OOS results, similar to those of Table V. The OOS-R^2s are higher than 5% across every measure considered, and all of the Wald statistics are highly statistically significant. Finally, the OOS-R^2s and the contemporaneous correlations between market returns and variance innovations are closely connected. When the absolute values of the correlations are high, the OOS-R^2s are also higher across every VRP measure considered.

2. Alternative Variance Forecast Models

The main variance forecast model (HAR-VAR model) used in this paper is a variant of the volatility forecast model of Corsi (2009). Concerned that the results are driven by a particular model, I ask whether the results hold when we use the original volatility forecast model of Corsi. The volatility forecast model is given as

$$\sqrt{RV}_{\tau+k} = a_0 + a_d\sqrt{RV}_{\tau} + a_w\sum_{j=0}^{4}\sqrt{RV}_{\tau-j} + a_m\sum_{j=0}^{21}\sqrt{RV}_{\tau-j} + \phi_{1,\tau+k}, \tag{15}$$

for k=1, ..., 22. Then, the contemporaneous variance betas, correlations, and the VRPs are constructed the same way as in previous sections.

Panels A, B, and C of Table XI summarize the key results of the paper for the alternative volatility forecast model. Overall, the panels indicate that the findings of this paper are not model specific. To be specific, the last two rows of Panel A provide results (comparable to Table III) of in-sample interactive regressions. Across all three measures considered, the coefficients are negative and statistically significant at 10% level. Panel B summarizes the in-sample R^2s for each subsample constructed from the contemporaneous correlations, which are comparable to Table IV. Across every measure considered, the in-sample R^2s are higher when correlations are more negative. Lastly, Panel C summarizes the performance of out-of-sample predictions. The contemporaneous beta approach outperforms the traditional approach across all three measures of the VRP. The Wald statistics are highly significant for the contemporaneous beta approach, but none of the three are significant for the traditional approach. The results again confirm that out-of-sample predictions can be made more accurately when correlations between return and variance innovations are more negative. Although not reported in the paper, the results are robust I use a model of the log of the variance instead of the variance itself. Also, the findings are robust when the Realized GARCH model of Hansen, Huang, and Shek (2012) is used instead.

3. Other Specifications of The Variance Risk Premium

This paper is motivated by the predictive regression of Bollerslev, Tauchen, and Zhou (2009), who report a positive relation between the VRP and future market returns. Their measure is constructed by taking the difference between the end-of-month value of implied variance and the monthly average realized variance. As discussed, this timing mismatch could result in biases when market variance has trends during a month. That said, Table XI indicates that the key findings remain intact even for the measure of Bollerslev et. al. This VRP measure is denoted by VRP_{BTZ}.

I also consider the measure of Bekaert and Hoerova (2014). The second component, the real-world expectation component of the VRP is measured using a forecast model of

$$RV_t = a_0 + a_1 RV_{t-1} + a_2 VIX_{t-1} + e_t$$

where RV is the monthly average of the daily RVs. This measure is often preferred (Gonzalez-Urteaga and Rubio 2016, Chen, Joslin, and Ni 2016) because the real-world expectation component is forward looking, which will reduce possible bias created by the trending variance. However, this measure will still create a bias when the variance is trending during the most recent month because the end-of-month value is often used for the implied variance, but the RV forecast models are based on monthly averaged observations.

The first column of Panel A summarizes results for the interactive regressions. Recall that the interactive coefficient measures how the predictive beta comoves with the contemporaneous variance beta. Results based on the main variance forecast model as well as for the volatility forecast model are provided. The panel suggests that the interactive coefficient is no longer statistically significant for the VRP_{BTZ}. However, the coefficient has a consistent sign with the hypothesis that the two betas are negatively related. When the measure of Bekaert and Hoerova (2014) is used, the coefficients are all negative and statistically significant. Overall, the panel suggests that the contemporaneous and the predictive betas are still closely related when the predictive variable is that from Bollerslev et al. The slightly weaker result could be due to the timing mismatch discussed above.

Panel B summarizes the results of in-sample predictive regressions after dividing the sample into subsamples. The first row of the panel shows that predictions can be made more accurately when the market and its variance move closely together. However, when they move independently, the VRP does not provide any information about the expected market returns.

Finally, the first two columns of Panels C and D provide the performance out-of-sample forecasts under the alternative VRP. The contemporaneous beta approach outperforms the traditional approach using both additional measures of VRP. Moreover, the size of the OOS-R^2s is closely connected to how the market and variance move together. They are higher when the absolute value of the correlation between returns and variance innovations is high.

VIII. Conclusion

It is well known that the market reacts negatively to unexpected shocks in market variance. This negative relation between market returns and variance innovation implies that the market is subject to variance risk. Fortunately, we can gauge the variance risk premium of the market relatively accurately using option prices and the realized variance of the market index. Moreover, previous research finds that this risk premium is useful in predicting short-horizon market returns.

This paper studies how the market's exposure to variance risk is related to the time-varying predictability of market returns by the VRP. First, 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. Second, the predictive power strongly depends on the contemporaneous correlation between returns and variance innovations. When correlations are highly negative, predictions can be made more accurately. This result holds both in sample and out of sample. 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, regardless of the strength of the asymmetry in the market.

The second part of the paper extends this result to the currency market. Across 21 currencies, this article shows that some of the returns can be predicted using the VRP of the equity index while others are unpredictable. Returns that are predictable are those that have a high correlation with the market variance, either positive or negative. Moreover, returns that react sensitively to changes in market variance (i.e. have a more negative beta) can be predicted by the VRP with a higher slope. These relations reinforce that it is the negative correlation between returns and market variance that is the primary determinant of VRP-based predictability.

References

Andersen, Torben G, and Tim Bollerslev, 1997, Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns, *Journal of Finance* 52, 975–1005.

Andersen, Torben G, Tim Bollerslev, and Francis X Diebold, 2007, Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility, *Review of Economics and Statistics* 89, 701–720.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2001, The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association* 96, 42–55.

Andersen, Torben G, Tim Bollerslev, Francis X Diebold, and Paul Labys, 2003, Modeling and Forecasting Realized Volatility, *Econometrica* 71, 579–625.

Andersen, Torben G., Tim Bollerslev, and Xin Huang, 2011, A Reduced Form Framework for Modeling Volatility of Speculative Prices based on Realized Variation Measures, *Journal of Econometrics* 160, 176–189.

Ang, Andrew, and Geert Bekaert, 2002, Regime Switches in Interest Rates, *Journal of Business and Economic Statistics* 20, 163–182.

Ang, Andrew, and Geert Bekaert, 2007, Stock Return Predictability: Is it There?, *Review of Financial Studies* 20, 651–707.

Bandi, F.M., and R. Reno, 2016, Price and Volatility Co-jumps, *Journal of Financial Economics* 119, 107–146.

Bekaert, Geert, and Marie Hoerova, 2014, The VIX, the Variance Premium and Stock Market Volatility, *Journal of Econometrics* 183, 181–192.

Bekaert, Geert, Marie Hoerova, and Marco LoDuca, 2013, Risk, Uncertainty and Monetary Policy, *Journal of Monetary Economics* 60, 771–788.

Black, Fischer, 1976, Studies of Stock Price Volatility Changes, in *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section,* pp. 177–181.

Bollerslev, Tim, Michael Gibson, and Hao Zhou, 2011, Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities, *Journal of Econometrics* 160, 235–245.

Bollerslev, Tim, Julia Litvinova, and George Tauchen, 2006, Leverage and Volatility Feedback Effects in High-Frequency Data, *Journal of Financial Econometrics* 4, 353–384.

Bollerslev, Tim, James Marrone, Lai Xu, and Hao Zhou, 2014, Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence, *Journal of Financial and Quantitative Analysis* 49, 633–661.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.

Bollerslev, Tim, Viktor Todorov, and Lai Xu, 2015, Tail Risk Premia and Return Predictability, *Journal of Financial Economics* 118, 113–134.

Bollerslev, Tim, Lai Xu, and Hao Zhou, 2015, Stock Return and Cash Flow Predictability: The Role of Volatility Risk, *Journal of Econometrics* 187, 458–471.

Busch, Thomas, Bent Jesper Christensen, and Morten rregaard Nielsen, 2011, The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets, *Journal of Econometrics* 160, 48–57.

Campbell, John Y, 1987, Stock Returns and the Term Structure, *Journal of Financial Economics* 18, 373–399.

Campbell, John Y, and Robert J Shiller, 1988, Stock Prices, Earnings, and Expected Dividends, *Journal of Finance* 43, 661–676.

Campbell, John Y, and Samuel B Thompson, 2008, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies* 21, 1509–1531.

Carr, Peter, and Liuren Wu, 2009, Variance Risk Premiums, *Review of Financial Studies* 22, 1311–1341.

Carr, Peter, and Liuren Wu, 2016, Analyzing Volatility Risk and Risk Premium in Option Contracts: A New Theory, *Journal of Financial Economics* 120, 1–20.

Chen, Hui, Scott Joslin, and Sophie X Ni, 2016, Demand for Crash Insurance, Intermediary Constraints, and Risk Premia in Financial Markets, *Working Paper*.

Chen, Long, 2009, On the Reversal of Return and Dividend Growth Predictability: A Tale of Two Periods, *Journal of Financial Economics* 92, 128–151.

Chen, Xilong, and Eric Ghysels, 2011, News -Good or Bad?- And Its Impact on Volatility Predictions over Multiple Horizons, *Review of Financial Studies* 24, 46–81.

Choi, Hoyong, Philippe Mueller, and Andrea Vedolin, 2015, Bond Variance Risk Premiums, *Working Paper*.

Christie, Andrew A., 1982, The Stochastic Behavior of Common Stock Variances, *Journal of Financial Economics* 10, 407–432.

Cochrane, John H., 1991, Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations, *Journal of Finance* 46, 209–237.

Corsi, Fulvio, 2009, A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics* 7, 174–196.

Cujean, Julien, and Michael Hasler, 2015, Why Does Return Predictability Concentrate in Bad Times?, *Working Paper*.

Dai, Qiang, Kenneth J. Singleton, and Wei Yang, 2007, Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields, *Review of Financial Studies* 20, 1669–1706.

d'Amico, Stefania, Don H Kim, and Min Wei, 2014, Tips from TIPS: The Informational Content of Treasury Inflation-Protected Security Prices, *Working Paper*.

Dangl, Thomas, and Michael Halling, 2012, Predictive Regressions with Time-varying Coefficients, *Journal of Financial Economics* 106, 157–181.

Diebold, Francis X., and Roberto S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business & Economic Statistics* 13, 253–263.

Drechsler, Itamar, and Amir Yaron, 2011, What's Vol Got to Do with It, *Review of Financial Studies* 24, 1–45.

Engle, R. F., and G. G. J. Lee, 1999, A Permanent and Transitory Component Model of Stock Return Volatility, in *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, ed. by R.F. Engle, and H. White.

Fama, Eugene F., and Kenneth R. French, 1988a, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1988b, Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96, 246–273.

Feunou, Bruno, Mohammad R Jahan-Parvar, and Cedric Okou, 2015, Downside Variance Risk Premium, *Working Paper*.

French, Kenneth R, G William Schwert, and Robert F Stambaugh, 1987, Expected Stock Returns and Volatility, *Journal of Financial Economics* 19, 3–29.

Garcia, Diego, 2013, Sentiment during Recessions, *Journal of Finance* 68, 1267–1300.

Gerd, Herber, Asger Lunde, Neil Shephard, and Kevin K. Sheppard, 2009, Oxford-Man Institute's Realized Library, *Oxford-Man Institute, University of Oxford*.

Giacomini, Raffaella, and Halbert White, 2006, Tests of Conditional Predictive Ability, *Econometrica* 74, 1545–1578.

Goetzmann, William Nelson, and Philippe Jorion, 1993, Testing the Predictive Power of Dividend Yields, *Journal of Finance* 48, 663–679.

Gonzalez-Urteaga, Ana, and Gonzalo Rubio, 2016, The Cross-sectional Variation of Volatility Risk Premia, *Journal of Financial Economics* 119, 353–370.

Goyal, Amit, and Ivo Welch, 2008, A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455–1508.

Hamilton, James D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica* 57, 357–384.

Hansen, Peter Reinhard, Zhuo Huang, and Howard Howan Shek, 2012, Realized GARCH: a Joint Model for Returns and Realized Measures of Volatility, *Journal of Applied Econometrics* 27, 877–906.

Hansen, Peter R, and Asger Lunde, 2006, Realized Variance and Market Microstructure Noise, *Journal of Business & Economic Statistics* 24, 127–161.

Henkel, Sam James, J. Spencer Martin, and Federico Nardari, 2011, Time-varying Short-horizon Predictability, *Journal of Financial Economics* 99, 560–580.

Hodrick, Robert J., 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, *Review of Financial Studies* 5, 357–386.

Johannes, Michael, Arthur Korteweg, and Nicholas Polson, 2014, Sequential Learning, Predictability, and Optimal Portfolio Returns, *Journal of Finance* 69, 611–644.

Jones, Christopher S., and Selale Tuzel, 2013, New Orders and Asset Prices, *Review of Financial Studies* 26, 115–157.

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting Returns in the Stock and Bond Markets, *Journal of Financial Economics* 17, 357–390.

Kelly, Bryan, and Hao Jiang, 2014, Tail Risk and Asset Prices, *Review of Financial Studies* 27, 2841–2871.

Lettau, Martin, and Sydney C. Ludvigsen, 2001, Consumption, Aggregate Wealth, and Expected Stock Returns, *Journal of Finance* 56, 815–849.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence, *Review of Financial Studies* 21, 1607–1652.

Londono, Juan M, 2014, The Variance Risk Premium around the World, *Working Paper.*

Londono, Juan M, and Hao Zhou, 2014, Variance Risk Premiums and the Forward Premium Puzzle, *Working Paper.*

Lustig, Hanno, Nikolai Roussanov, and Adrien Verdelhan, 2014, Countercyclical Currency Risk Premia, *Journal of Financial Economics* 111, 527–553.

Martin, Ian, and Christian Wagner, 2016, What is the Expected Return on the Market?, *Quarterly Journal of Economics* Forthcoming.

Merton, Robert C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica* 41, 867–87.

Newey, Whitney K, and Kenneth D. West, 1987, A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Pastor, Lubos, and Robert Stambaugh, 2009, Predictive Systems: Living with Imperfect Predictors, *Journal of Finance* 64, 1583–1628.

Pindyck, Robert S., 1984, Risk, Inflation, and the Stock Market, *American Economic Review* 74, 335–351.

Poterba, James M, and Lawrence H Summers, 1988, Mean Reversion in Stock Prices: Evidence and Implications, *Journal of Financial Economics* 22, 27–59.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, *Review of Financial Studies* 23, 821–862.

Stambaugh, Robert F., 1999, Predictive Regressions, *Journal of Financial Economics* 54, 375–421.

Stuart, Alan, and Keith Ord, 2010, *Kendall's Advanced Theory of Statistics.* (Wiley).

Todorov, Viktor, 2010, Variance Risk-premium Dynamics: The Role of Jumps, *Review of Financial Studies* 23, 345–383.

Wang, Hao, Hao Zhou, and Yi Zhou, 2013, Credit Default Swap Spreads and Variance Risk Premia, *Journal of Banking & Finance* 37, 3733–3746.

White, Halbert, 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica* 48, 817–838.

Zhang, Lan, P. Mykland, and Yacine Ait-Sahalia, 2005, A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data, *Journal of the American Statistical Association* 100, 1394–1411.

A. Figures and Tables

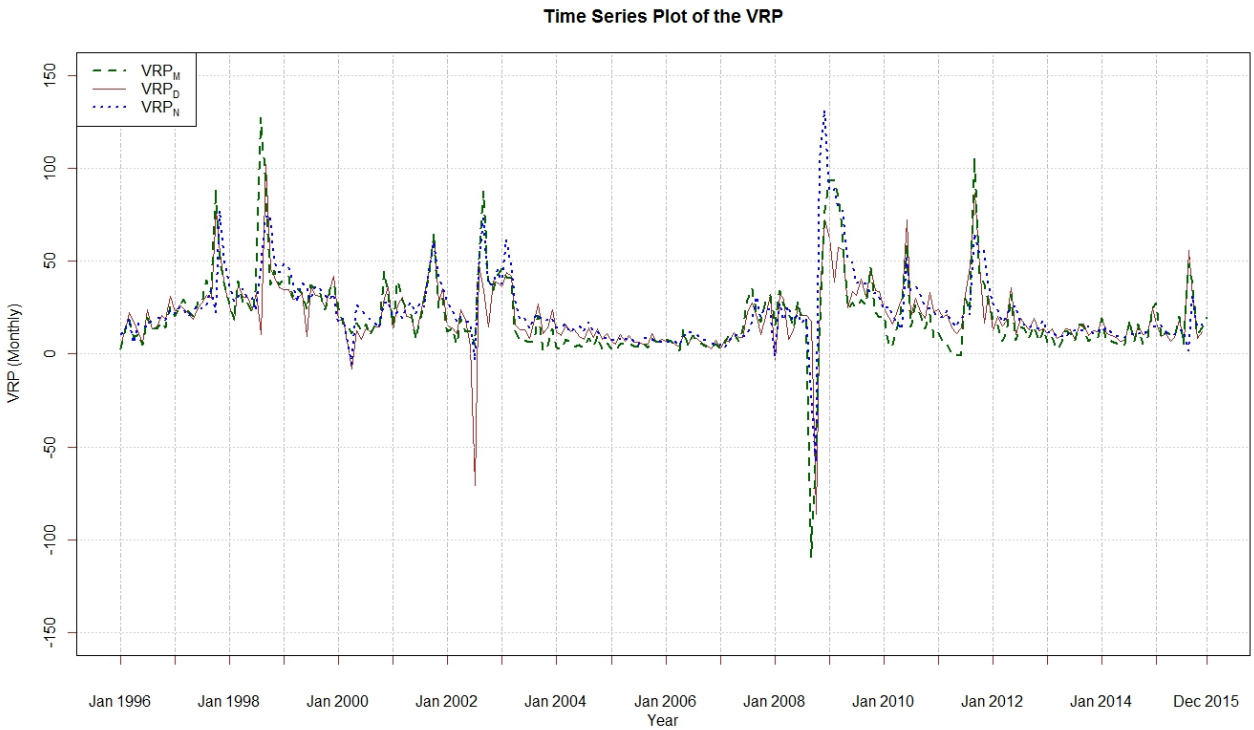

Figure 1. Time Series of the Variance Risk Premium

This figure compares the three VRP measures considered in this paper. VRP$_N$ is a nonparametric measure, defined as the monthly average of VIX^2/12 and the monthly realized variance. VRP$_D$ and VRP$_M$ are model-based measures estimated at the end of each month. The daily variance forecast is used for VRP_D and the cumulative sum of 22-day forecasts is used for VRP_M. A variant of Corsi's (2009) model is used to predict the market variance.

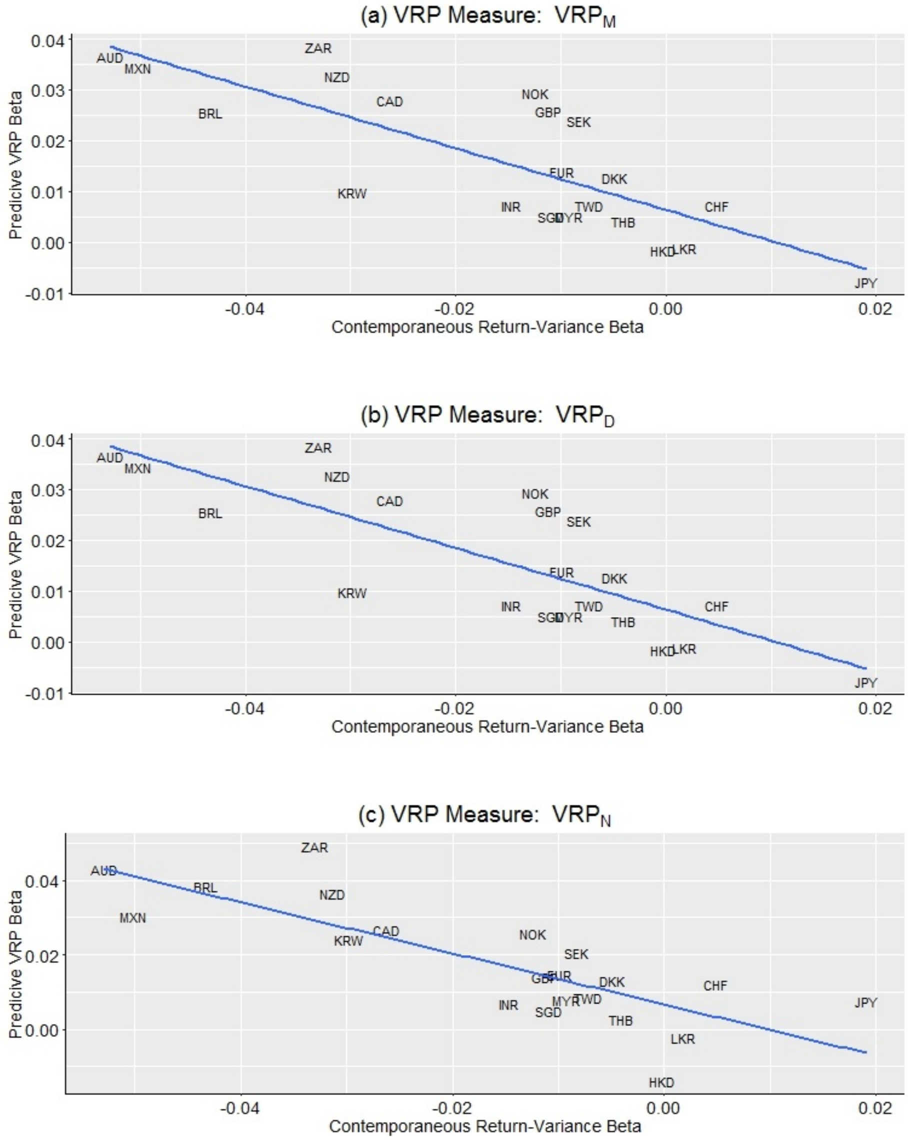

Figure 2. Cross-sectional Relation between the Contemporaneous and Predictive Betas

This figure illustrates the relation between the average contemporaneous variance beta and the predictive VRP beta of currency returns. Each figure represents the relation for a different VRP measure, and the solid line represents the OLS fit between these two betas. The contemporaneous beta is from the daily regression

$$r_{i,\tau} = \alpha_{i,t} + \beta_{i,v,t}(RV_\tau - E_{\tau-1}[RV_\tau]) + \epsilon_{i,\tau}.$$

where r_i is the UIRP-adjusted currency return. RV is the realized variance of the equity index. The regressions are estimated every month using daily observations that belongs to that month. Then, the time-series averages of these monthly betas are computed for each currency. The predictive beta is from a single regression,

$$r_{i,t+1} = \beta_{i,0} + \beta_{i,p}VRP_t + \epsilon_{i,t+1}$$

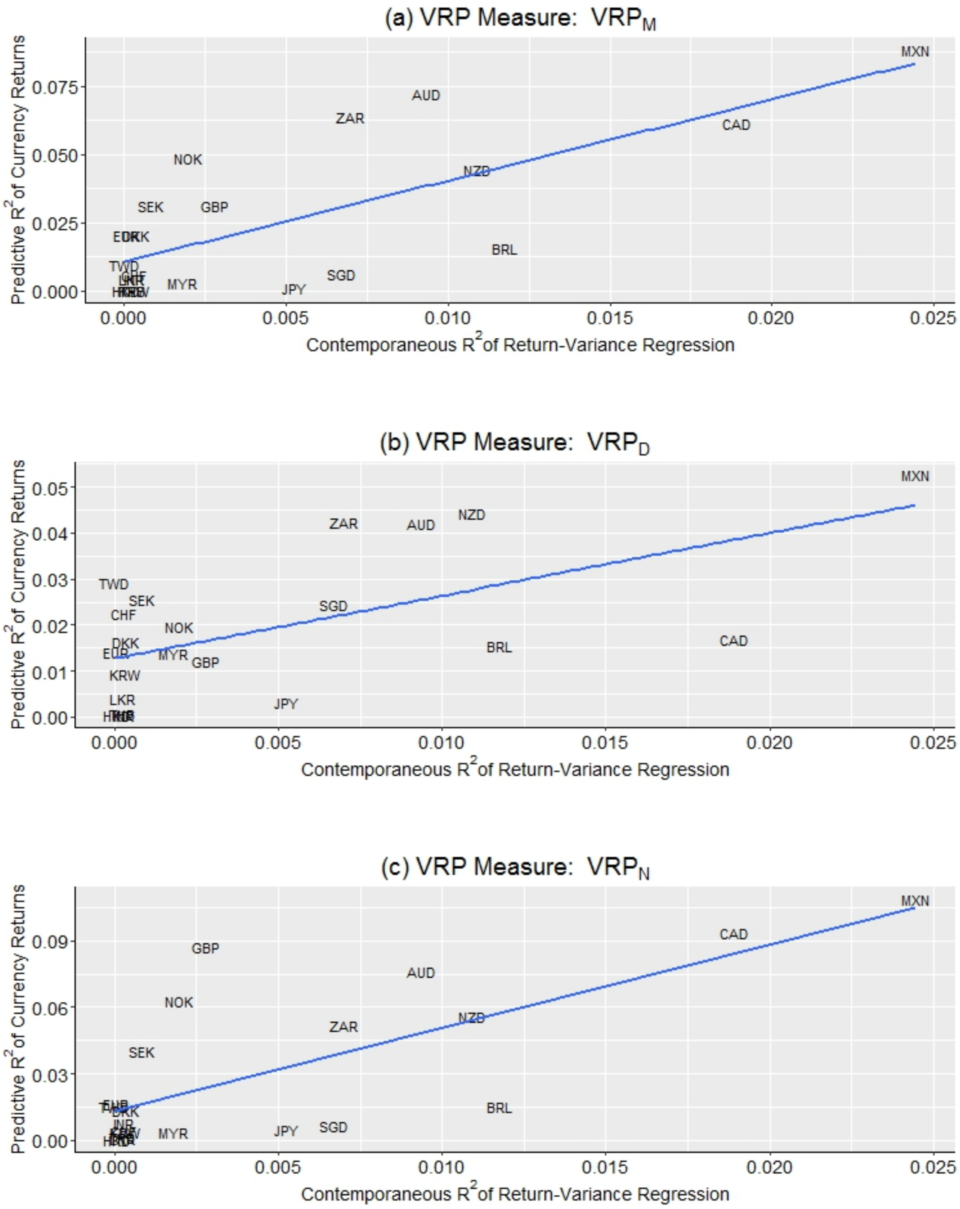

Figure 3. Cross-sectional Relation between the Contemporaneous and Predictive R^2s

This figure illustrates the relation between the average contemporaneous return-variance R^2s and the predictive return- VRP R^2s of currency returns. Each figure represents the relation for a different VRP measure, and the solid line represents the OLS fit between the two R^2s. The contemporaneous R^2 is from the daily regression

$$r_{i,\tau} = \alpha_i + \beta_{i,v}(RV_\tau - E_{\tau-1}[RV_\tau]) + \epsilon_{i,\tau}.$$

where r_i is the UIRP-adjusted currency return. RV is the realized variance of the equity index. The regressions are estimated using daily observation for the entire sample. The predictive R^2 is from a single regression,

$$r_{i,t+1} = \beta_{i,0} + \beta_{i,p}VRP_t + \epsilon_{i,t+1}.$$

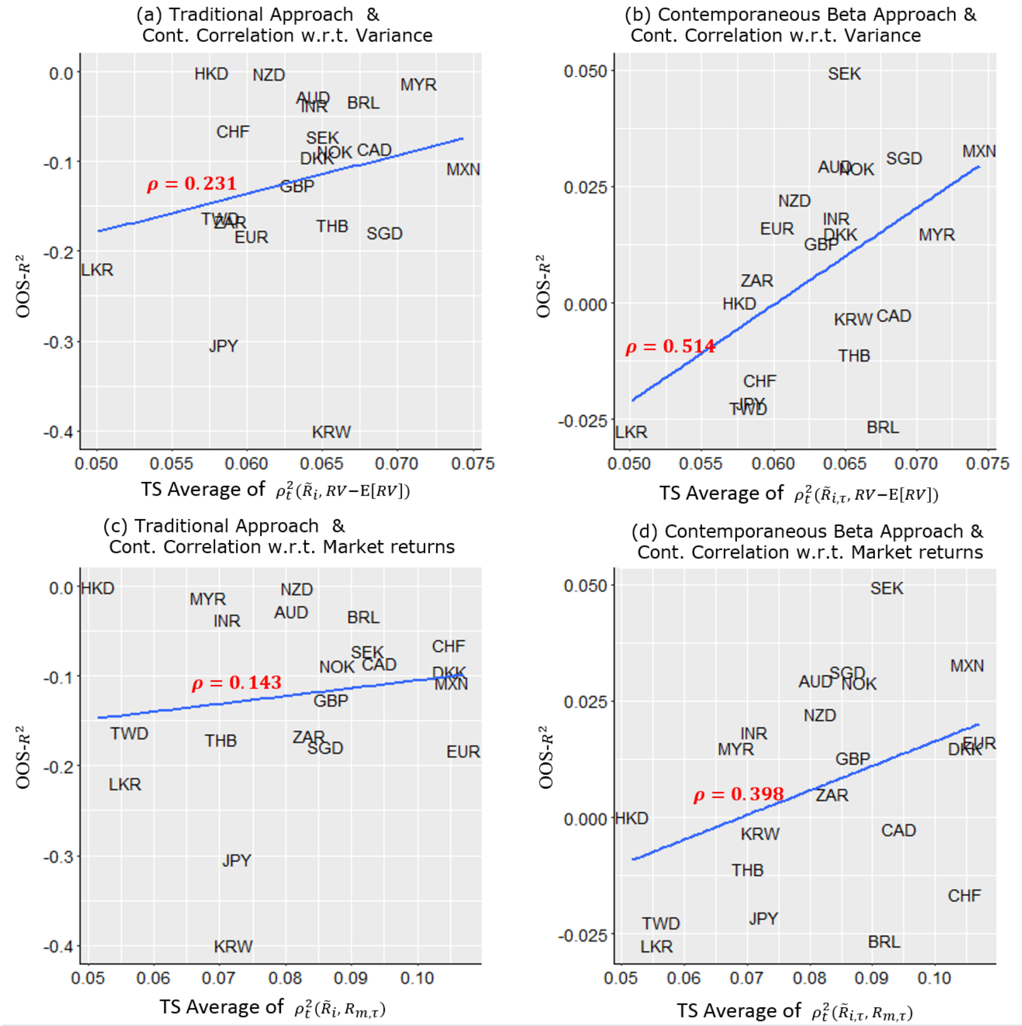

Figure 4. Out-of-Sample Predictive Performance of Currency Returns and the Contemporaneous Correlation

This figure compares the OOS-R^2s with the time-series average contemporaneous square of correlations. The contemporaneous correlations are between adjusted currency returns and changes in U.S. equity variance for (a) and (b), while the correlations are between adjusted currency returns and stock returns in (c) and (d). (a) and (c) compares the relation when the $OOS - R^2$ are computed from the traditional approach while (b) and (d) are from the contemporaneous beta approach.

Table I
Summary Statistics

This table summarizes the means, standard deviations, and the first-order autocorrelations for the variables of interest during the sample period of 1996-2015. The realized variance (RV) is the sum of the square of five-minute market returns from the first to the last day of the month, averaged over five one-minute sub-samples. The implied variance (IV) is the monthly mean of the daily volatility index (VIX) squared divided by 12. The variance risk premium is estimated using three different methods. VRP_M is the difference between the end-of-month IV and the 22-day cumulative variance forecast from the HAR-VAR model. VRP_D is the difference between the end-of-month IV and the one-day-ahead RV forecast. VRP_N is the monthly average of the difference between the implied and the realized variance. The contemporaneous beta of month t $\hat{\beta}_{c,t}$ is the slope of the regression

$$R_{m,\tau} - R_f = \beta_{0,t} + \beta_{v,t}(RV_\tau - E_{\tau-1}[RV_\tau]) + \epsilon_\tau,$$

where $R_{m,\tau}$ is the market return of day τ, and $E_{\tau-1}[RV_\tau]$ is the one-day-ahead forecast of the daily realized variance made on day $\tau - 1$ from the HAR-VAR model. These regressions are estimated each month using data from the first to the last day of the month. The monthly contemporaneous correlation ($\hat{\rho}_t$) of month t is the correlation between daily market returns and variance innovations during the same estimation window.

			$\hat{\rho}_c \leq$ median		NBER Recession		
(1996-2015)	Mean	St Dev	Mean	St Dev	Mean	St Dev	Autocorrelation
Realized Variance	20.19	33.98	18.21	8.48	59.56	81.52	0.630
Implied Variance	42.25	36.96	40.26	38.89	88.25	68.9	0.790
VRP_M	20.89	20.02	19.53	19.91	34.09	30.29	0.613
VRP_D	20.57	17.92	19.42	16.32	27.22	28.51	0.311
VRP_N	23.41	19.38	20.80	17.55	38.74	38.28	0.675
$\hat{\beta}_v$	-0.042	0.052	-0.077	0.040	-0.016	0.034	0.213
$\hat{\rho}$	-0.268	0.295	-0.502	0.143	-0.147	0.233	0.111
Number of Months	240		120		28		

Table II
Transition Matrix of the Contemporaneous Correlations and Betas

This table provides a transition matrix of the contemporaneous betas and correlations. Betas are estimated from the regression

$$R_{m,\tau} - R_f = \beta_{0,t} + \beta_{v,t}(RV_\tau - E_{\tau-1}[RV_\tau]) + \epsilon_\tau,$$

where variables are defined as in Table 1. The contemporaneous correlation is the correlation between daily returns and variance innovations. These regressions are estimated every month using observations from the first to the last day of that month. Then, the entire sample is divided into quartiles based on either the beta or the correlation. The table shows transition probabilities from one quartile to the other.

Panel A. Contemporaneous Beta

From \ To	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Total
Quartile 1 (Most Negative)	26	16	11	7	60
	43.33%	26.67%	18.33%	11.67%	
Quartile 2	11	15	15	18	59
	18.64%	25.42%	25.42%	30.51%	
Quartile 3	11	14	20	15	60
	18.33%	23.33%	33.33%	25.00%	
Quartile 4 (Least Negative)	12	14	14	20	60
	20.00%	23.33%	23.33%	33.33%	
Total	60	59	60	60	

Panel B. Contemporaneous Correlation

From \ To	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Total
Quartile 1 (Most Negative)	20	13	15	12	60
	33.33%	21.67%	25.00%	20.00%	
Quartile 2	10	24	12	14	60
	16.67%	40.00%	20.00%	23.33%	
Quartile 3	15	9	17	18	59
	25.42%	15.25%	28.81%	30.51%	
Quartile 4 (Least Negative)	15	13	16	16	60
	25.00%	21.67%	26.67%	26.67%	
Total	60	59	60	60	

Table III
Predictive Regressions of Market Returns

This table summarizes the coefficients and the p-values of the predictive regressions of the one-month and three-month market returns ($R_{m,t+1}$). Panel A summarizes the regression with constant coefficients,

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

where VRP_t is the monthly variance risk premium estimated in one of several ways, as described in Table I. Panel B summarizes the results of the interactive predictive regression

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

where $\hat{\beta}_{c,t}$ is the contemporaneous beta of the daily regression,

$$R_{m,\tau} - R_f = \beta_{0,t} + \beta_{v,t}(RV_\tau - E_{\tau-1}[RV_\tau]) + \epsilon_\tau.$$

The betas are estimated each month using observations from the first to the last trading day of month $t.E_{\tau-1}[RV_\tau]$ is the one-day forecast of the realized variance from the HAR-VAR model.
The t-statistics are reported in parentheses using heteroscedasticity-consistent standard errors (one-month returns) or using the Newey and West (1987) standard errors with three lags (three-month returns).

Panel A. Constant Coefficient Predictive Regression

Dependent Var	*One-month Returns*			*Three-Month Returns*		
	VRP_M	VRP_D	VRP_N	VRP_M	VRP_D	VRP_N
	(1)	(2)	(3)	(4)	(5)	(6)
VRP_t	0.053**	0.047**	0.051**	0.132***	0.129***	0.118***
	(2.49)	(2.51)	(2.08)	(4.86)	(5.34)	(2.67)
Adj-R^2	0.049	0.030	0.043	0.099	0.074	0.073

Panel B. Interactive Regressions

Dependent Var	*One-month Returns* $R_{m,t+1}$			*Three-Month Returns* $R_{m,t+1}$		
	VRP_M	VRP_D	VRP_N	VRP_M	VRP_D	VRP_N
	(1)	(2)	(3)	(4)	(5)	(6)
VRP_t	0.039*	0.037*	0.045**	0.110***	0.118***	0.108**
	(2.04)	(1.93)	(1.97)	(3.09)	(4.17)	(2.19)
$\text{VRP}_t \times \hat{\beta}_{c,t}$	−0.689*	−0.706**	−0.735**	−1.147*	−0.936*	−1.231**
	(1.86)	(2.22)	(2.40)	(1.89)	(1.92)	(2.04)
Adj-R^2	0.067	0.051	0.071	0.115	0.086	0.097

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

Table IV
In-sample Predictive Performance and the Contemporaneous Correlations

This table summarizes the in-sample R^2s and t-statistics of monthly predictive regressions of one-month (Panel A) and three-month (Panel B) market returns ($R_{m,t+1}$)

$$R_{m,t+1} = \beta_{0p} + \beta_p VRP_t + \epsilon_{t+1}$$

for split samples, where VRP_t is the monthly series of the variance risk premium estimated in one of several ways, as described in Table I.

Each month in the sample (1996-2015) is classified as either a high, medium or low month depending on whether the square of the contemporaneous correlation is greater or less than the historical terciles evaluated from data available up to that point. The contemporaneous correlation is defined as the correlation between daily market returns ($R_{m,\tau}$) and variance innovations ($RV_\tau - E_{\tau-1}[RV_\tau]$). 'High' contains samples in which daily market returns and variance innovations move closely together.

The t-statistics, reported in parentheses, are adjuste for heteroscedasticity (one-month returns) or for both heteroscedasticity and autocorrelation using the approach of Newey-West with three lags (three-month returns).

Panel A. One-Month Market Return Prediction

		Samples classified by $\hat{\rho}^2$			
		High	Medium	Low	High-Low
	Number of months	81	69	81	
VRP$_M$	In-sample R^2	0.180***	0.012	0.006	0.174
	Predictive beta ($\hat{\beta}_p$)	0.095	0.022	0.022	
		(5.04)	(0.98)	(0.06)	
VRP$_D$	In-sample R^2	0.070***	0.024	0.005	0.065
	Predictive beta ($\hat{\beta}_p$)	0.096	0.035	0.013	
		(2.63)	(1.42)	(0.66)	
VRP$_N$	In-sample R^2	0.145***	0.017	0.004	0.141
	Predictive beta ($\hat{\beta}_p$)	0.110	0.026	0.012	
		(3.95)	(1.18)	(0.62)	

Panel B. Three-Month Market Return Prediction

		Samples classified by $\hat{\rho}^2$			
		High	Medium	Low	High-Low
	Number of months	81	69	81	
VRP$_M$	In-sample R^2	0.278***	0.047*	0.043*	0.234
	Predictive beta ($\hat{\beta}_p$)	0.184	0.116	0.084	
		(5.44)	(1.81)	(1.86)	
VRP$_D$	In-sample R^2	0.096***	0.059***	0.046**	0.024
	Predictive beta ($\hat{\beta}_p$)	0.146	0.102	0.084	
		(2.36)	(2.66)	(2.11)	
VRP$_N$	In-sample R^2	0.222***	0.058**	0.003	0.219
	Predictive beta ($\hat{\beta}_p$)	0.208	0.091	0.023	
		(5.13)	(2.21)	(0.54)	

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

Table V
Out-of-Sample Performance Evaluation (2000-2015)

This table summarizes the out-of-sample predictive performance for one-month market returns over 2000-2015 using the variance risk premium (VRP) as a predictor. VRP_t is the monthly series of the VRP estimated in one of several ways, as described in Table I.
The contemporaneous betas are the slopes from the monthly regressions of market returns on variance innovations, estimated using daily data from the first to the last day of each month. The traditional approach uses either seven years, ten years, or all data available up to that point to run a predictive regression of monthly market returns on the lagged VRP. A one-step-ahead forecast is then formed from the coefficients and compared with the realized returns.
The contemporaneous beta approach of out-of-sample prediction sets the forecast equal to the negative of the contemporaneous beta multiplied by the VRP. The risk-free rate is added to make it comparable to actual returns. As an alternative approach, the out-of-sample forecasts are obtained by adding a constant δ_0, which is estimated by taking the historical mean of $R_{m,t} + \beta_{v,t-1} VRP_{t-1}$. The out-of-sample R^2s and Wald statistics along the p-values of the statistics are provided for each of these methods.

	Sample	VRP Measure		
	Period	VRP$_M$	VRP$_D$	VRP$_N$
Traditional Approach				
7 Years	OOS-R^2	0.006	−0.005	−0.010
Rolling Sample	Wald	0.021	0.020	0.037
	p-value	(0.886)	(0.888)	(0.848)
10 Years	OOS-R^2	0.044	0.016	0.031
Rolling Sample	Wald	0.793	0.002	0.278
	p-value	(0.373)	(0.965)	(0.598)
Expanding Window	OOS-R^2	0.039	0.020	0.025
	Wald	0.569	0.221	0.137
	p-value	(0.451)	(0.638)	(0.711)
Contemporaneous Beta Approach				
No Intercept	OOS-R^2	0.104	0.084	0.080
	Wald	8.473	5.619	4.610
	p-value	(0.004)	(0.018)	(0.032)
Including Intercept	OOS-R^2	0.096	0.076	0.075
	Wald	6.930	4.874	4.112
	p-value	(0.008)	(0.027)	(0.043)

Table VI
Subsample Analysis - Out-of-Sample R^2

This table compares the out-of-sample performance of predictions for one-month market returns over the 2000-2015, with the variance risk premium (VRP) as a predictor. The entire sample is classified based on a method discussed in Table IV.
The contemporaneous betas are the slopes from the monthly regressions of market returns on variance innovations estimated using daily data from the first to the last day of each month. The traditional approach uses either seven years, ten years, or all data available up to that point to run a predictive regression of monthly market returns on the lagged VRP. A one-step-ahead forecast is then formed from the coefficients and compared with the realized returns.
The contemporaneous beta approach uses the negative of the contemporaneous beta multiplied by the VRP. The risk-free rate is added to make it comparable to actual returns. The contemporaneous beta approach with an intercept is obtained by running the following regression for a rolling ten years

$$R_{m,t+1} = -\hat{\beta}_{v,t} VRP_t + \delta_0 + \eta_{t+1}$$

to find $\hat{\delta}_0$. The out-of-sample forecast is formed using the estimate $\hat{\delta}_0$. The out-of-sample R^2s are provided for each of these methods and subsamples.

	Sample	VRP Measure		
	Period	VRP_M	VRP_D	VRP_N
Traditional Approach				
7 Years	High	0.107	0.067	0.076
Rolling Window	Medium	−0.039	0.004	−0.066
	Low	−0.088	−0.115	−0.070
	High-Low	0.195	0.182	0.146
10 Years	High	0.183	0.090	0.147
Rolling Window	Medium	−0.052	0.004	−0.065
	Low	−0.075	−0.052	−0.031
	High-Low	0.257	0.142	0.178
Expanding Window	High	0.173	0.094	0.139
	Medium	−0.043	0.017	−0.067
	Low	−0.059	−0.080	−0.039
	High-Low	0.232	0.173	0.178
Contemporaneous Beta Approach				
No Intercept	High	0.202	0.147	0.138
	Medium	0.034	0.032	0.032
	Low	0.038	0.048	0.046
	High-Low	0.164	0.099	0.092
Including Intercept	High	0.179	0.123	0.118
	Medium	0.040	0.050	0.045
	Low	0.003	0.004	0.006
	High-Low	0.177	0.118	0.111

Table VII
An Out-of-Sample Analysis of the Orthogonal Premium

This table provides out-of-sample performance of the hybrid approach, in which additional predictors are used to complement the contemporaneous beta approach. Forecasts are formed based on the regression

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

where X_t is some predictor. The out-of-sample R^2s and Wald statistics along the p-values of the statistics, are provided for each set of forecasts. Also, the R^2s of these approach for each of the subsample classification is provided in right four columns.

Additional		2000-2015			Subsamples			
					High	Medium	Low	H-L
Variable (X_t)		VRP$_M$	VRP$_D$	VRP$_N$	VRP$_M$			
Intercept	OOS-R^2	0.096	0.076	0.075	0.192	0.041	0.027	0.165
(Benchmark)	Wald	6.930	4.874	4.112				
	p-value	(0.008)	(0.027)	(0.043)				
VRP	OOS-R^2	0.110	0.102	0.102	0.276	0.020	-0.065	0.340
	Wald	4.075	3.391	3.317				
	p-value	(0.044)	(0.066)	(0.069)				
D/Y	OOS-R^2	0.086	0.071	0.069	0.139	0.012	0.079	0.060
	Wald	4.036	3.315	2.782				
	p-value	(0.045)	(0.069)	(0.095)				
TERM	OOS-R^2	0.076	0.058	0.056	0.187	0.006	-0.060	0.248
	Wald	2.901	1.940	1.452				
	p-value	(0.089)	(0.164)	(0.228)				
DEF	OOS-R^2	0.045	0.035	0.035	0.123	0.045	-0.094	0.217
	Wald	0.747	0.441	0.455				
	p-value	(0.387)	(0.507)	(0.500)				
Short rate	R^2	0.059	0.048	0.045	0.131	0.067	-0.059	0.190
	Wald	1.760	1.385	1.033				
	p-value	(0.185)	(0.239)	(0.309)				
cay	OOS-R^2	0.073	0.058	0.053	0.182	0.082	-0.119	0.301
	Wald	2.485	1.747	1.335				
	p-value	(0.115)	(0.186)	(0.248)				
I/K	OOS-R^2	0.082	0.064	0.062	0.160	0.074	-0.060	0.220
	Wald	3.107	2.307	1.980				
	p-value	(0.078)	(0.129)	(0.159)				
NO/S	OOS-R^2	0.056	0.035	0.037	0.091	0.044	0.010	0.081
	Wald	1.395	0.519	0.592				
	p-value	(0.238)	(0.471)	(0.442)				
Left Tail	R^2	0.090	0.074	0.069	0.133	0.056	0.057	0.076
	Wald	5.351	4.868	3.105				
	p-value	(0.021)	(0.027)	(0.078)				
LJV	OOS-R^2	0.003	−0.018	−0.003	0.169	-0.215	-0.024	0.193
	Wald	0.001	0.033	0.001				
	p-value	(0.974)	(0.856)	(0.971)				

Table VIII
Evaluation of the Economic Significance of Out-of-Sample Trading Strategies

This table summarizes the out-of-sample investment performance in terms of the annual Sharpe Ratios and the Certainty Equivalents, assuming a risk-averse investor with a risk-aversion coefficient of 3. The weight on the market portfolio is determined by the out-of-sample forecasts described previously, and capped at 200%. The rest $1 - w_T$ is invested in a risk-free asset. "Unconditional weights" are formed in a way that does not rely on the contemporaneous correlations. The weights are

$$w_T = \frac{\hat{R}_{m,T+1|T} - R_f}{\gamma \widehat{RV}_T^M}.$$

The 'conditional' weights depend on the size of the estimated contemporaneous correlations ($\hat{\rho}_c$), which equals to

$$w_T = \frac{\hat{R}_{m,T+1|T} - R_f}{\gamma \widehat{RV}_T^M} \hat{\rho}_T^2 + \frac{\overline{R}_{m,T} - R_f}{\gamma \widehat{RV}_T^M}(1 - \hat{\rho}_T^2).$$

The benchmark uses that the historical average as a forecast for the future. Gains/Losses relative to the benchmark is reported.

	Unconditional Weights		Conditional Weights	
Historical Average (Benchmark)				
	Sharpe Ratio	Certainty Equivalent		
	0.568	0.053		
	Δ Sharpe Ratio	Δ Certainty Equivalent	Δ Sharpe Ratio	Δ Certainty Equivalent
Traditional Approach				
VRP$_M$	−0.075	−0.018	+0.032	+0.005
VRP$_N$	−0.279	−0.068	−0.032	−0.006
VRP$_D$	−0.136	−0.032	+0.026	+0.004
Contemporaneous Beta Approach				
No Intercept				
VRP$_M$	+0.103	+0.021	+0.103	+0.019
VRP$_N$	+0.126	+0.026	+0.097	+0.018
VRP$_D$	+0.106	+0.022	+0.081	+0.015
Including Intercept				
VRP$_M$	+0.019	+0.002	+0.093	+0.017
VRP$_N$	+0.007	−0.005	+0.088	+0.016
VRP$_D$	+0.017	+0.001	+0.067	+0.012

Table IX
Predictions of Currency Returns using the VRP of S&P 500

This table summarizes the results of predictive regressions,

$$r_{i,t+1} = \beta_{i,0} + \beta_{i,p} VRP_t + \epsilon_{i,t+1},$$

where r_i is the monthly adjusted currency return of currency i, and where the VRP is measured as in Table I. The table summarizes the slopes $\hat{\beta}_{i,p}$, t-statistics, and the R^2s of the regression. Standard errors are adjusted for heteroscedasticity.

	$\mathbf{VRP}_M$	R^2	$\mathbf{VRP}_D$	R^2	$\mathbf{VRP}_N$	R^2
GBP	0.019***	0.031	0.026***	0.087	0.014*	0.011
	(2.772)		(4.755)		(1.655)	
AUD	0.044***	0.072	0.036***	0.076	0.043***	0.048
	(4.312)		(4.430)		(3.458)	
BRL	0.032*	0.016	0.026*	0.015	0.038*	0.016
	(1.951)		(1.929)		(1.953)	
CAD	0.028***	0.062	0.028***	0.094	0.027***	0.038
	(3.954)		(4.957)		(3.080)	
DKK	0.019**	0.020	0.013*	0.014	0.013	0.007
	(2.225)		(1.823)		(1.252)	
EUR	0.021**	0.020	0.014*	0.016	0.015	0.008
	(2.050)		(1.844)		(1.295)	
HKD	0.000	0.000	−0.002	0.000	−0.014	0.000
	(0.011)		(0.052)		(0.298)	
INR	0.007	0.004	0.007	0.008	0.007	0.003
	(1.033)		(1.372)		(0.877)	
JPY	−0.004	0.001	−0.008	0.005	0.007	0.002
	(0.486)		(1.092)		(0.687)	
KRW	0.000	0.000	0.010	0.004	0.024	0.010
	(0.030)		(0.949)		(1.558)	
MYR	0.006	0.003	0.005	0.003	0.008	0.004
	(0.805)		(0.910)		(0.951)	
MXN	0.039***	0.088	0.034***	0.109	0.030***	0.039
	(4.799)		(5.383)		(3.088)	
NZD	0.036***	0.045	0.033***	0.056	0.036***	0.031
	(3.333)		(3.747)		(2.780)	
NOK	0.032***	0.049	0.029***	0.063	0.026**	0.022
	(3.497)		(3.998)		(2.310)	
SGD	0.006	0.006	0.005	0.006	0.005	0.003
	(1.205)		(1.236)		(0.828)	
ZAR	0.053***	0.064	0.039***	0.052	0.049***	0.038
	(4.022)		(3.608)		(3.082)	
LKR	−0.004	0.005	−0.001	0.001	−0.002	0.001
	(1.040)		(0.367)		(0.482)	
SEK	0.026***	0.031	0.024***	0.040	0.021*	0.014
	(2.761)		(3.169)		(1.831)	
CHF	0.011	0.006	0.007	0.004	0.012	0.005
	(1.177)		(0.979)		(1.097)	
TWD	0.007	0.009	0.007*	0.015	0.008	0.009
	(1.508)		(1.920)		(1.489)	
THB	0.002	0.000	0.004	0.001	0.003	0.000
	(0.201)		(0.508)		(0.233)	

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

Table X
Cross-sectional Regressions of Currency Returns

This table summarizes the coefficients of the cross-sectional regression of the predictive betas on the contemporaneous betas

$$\hat{\beta}_{i,p} = \gamma_0 + \gamma_1 \frac{1}{T} \sum_{t=1}^{T} \hat{\beta}_{i,v,t} + \epsilon_{i,t},$$

where $\hat{\beta}_{i,p}$ is the predictive beta of currency i's adjusted returns. The predictive regression is

$$r_{i,t+1} = \beta_{i,0} + \beta_{i,p} VRP_t + \epsilon_{i,t+1},$$

where VRP is measured as in Table I. Similarly, $\hat{\beta}_{i,v,t}$ is the contemporaneous beta of currency i's returns regressed on variance innovations of the market index using daily observations of month t. Standard errors are from equation (14).

	Dependent Var: $\hat{\beta}_p$		
VRP used	VRP$_M$	VRP$_D$	VRP$_N$
Intercept	0.007**	0.006	0.007
CSR T-stat	(2.281)	(1.462)	(1.047)
$\hat{\beta}_c$	-0.686^{***}	-0.608^{***}	-0.685^{***}
CSR T-stat	(6.804)	(4.471)	(5.702)
Adj-R^2	0.540	0.625	0.621

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

Table XI
Robustness Tests

This table summarizes several robustness tests. $\hat{\beta}_{t+1|t}$ is the predicted contemporaneous beta, forecasted using an AR(2) specification. The betas are formed out of the sample on a rolling basis. VRP$_{BTZ}$ is from Bollerslev, Tauchen, and Zhou (2009), VRP$_{BH}$ is from Bekaert and Hoerova (2014). $\hat{\beta}_{v,t}(vol)$ indicates that the variance innovations are estimated using a forecast model of market volatility.

Panel A. Interactive Predictive Regressions

Dependent Variable		$R_{m,t+1}$				
		VRP$_{BTZ}$	VRP$_{BH}$	VRP_M	VRP_D	VRP_N
VRP$_t$ ×	$\hat{\beta}_{t+1\|t}$	−1.183	−2.314**	−2.447***	−2.119**	−2.412***
		(−1.23)	(2.20)	(2.35)	(1.98)	(2.46)
	$\hat{\beta}_{c,t}$	−0.419*	−1.804***			
		(1.67)	(1.86)			
	$\hat{\beta}_{c,t}(vol)$	−0.377	−0.754***	−0.582*	−0.476*	−0.629**
		(1.42)	(2.45)	(1.95)	(1.65)	(2.12)

Panel B. Predictive Power by Subsamples Classified on $\hat{\rho}^2$

	Forecast Model	Subsamples		
		High	Medium	Low
VRP$_{BTZ}$	HAR-RV	0.102	0.097	0.011
	HAR-VAR	0.212	0.114	0.002
VRP$_{BH}$	HAR-RV	0.048	0.043	0.009
	HAR-VAR	0.080	0.056	0.003
VRP_M	HAR-RV	0.149	0.049	0.000
VRP_D	HAR-RV	0.240	0.056	0.000
VRP_N	HAR-RV	0.148	0.143	0.001

Panel C. Out-of-Sample Performance Using a Volatility Forecast Model

		VRP$_{BTZ}$	VRP$_{BH}$	VRP_M	VRP_D	VRP_N
Traditional Approach						
OOS-R^2		0.069	0.015	0.036	0.090	0.035
Wald		2.484	0.211	0.602	1.308	2.487
p-value		(0.115)	(0.646)	(0.438)	(0.253)	(0.115)
	High	0.142	0.044	0.155	0.278	0.138
OOS-R^2	Medium	0.062	0.012	0.009	0.025	0.085
	Low	0.011	−0.025	0.013	0.020	0.027
Contemporaneous Beta Approach						
OOS-R^2		0.074	0.089	0.086	0.090	0.072
Wald		2.795	7.896	4.030	3.810	3.348
p-value		(0.095)	(0.005)	(0.045)	(0.051)	(0.067)
	High	0.162	0.178	0.194	0.210	0.140
OOS-R^2	Medium	0.008	0.026	0.013	0.000	0.026
	Low	0.013	0.032	0.013	0.020	0.021

Panel D. Out-of-Sample Performance using the Predicted Beta ($\hat{\beta}_{t+1|t}$)

		VRP$_{BTZ}$	VRP$_{BH}$	VRP_M	VRP_D	VRP_N
Contemporaneous Beta Approach						
OOS-R^2		0.080	0.056	0.112	0.086	0.083
Wald		3.305	3.070	7.887	5.206	4.395
p-value		(0.069)	(0.080)	(0.005)	(0.023)	(0.036)
	High	0.154	0.126	0.222	0.147	0.138
OOS-R^2	Medium	0.023	0.030	0.036	0.040	0.042
	Low	0.037	−0.013	0.039	0.051	0.044