Working paper

2020-11

Statistics and Econometrics

ISSN 2387-0303

Adaptative predictability of stock market returns

Isabel Casas, Xiuping Mao, Helena Veiga

uc3m | Universidad Carlos III de Madrid

Adaptive predictability of stock market returns

Isabel Casas[a], Xiuping Mao[b], Helena Veiga[c]

[a] *Department of Finance and Economics, Universidad de Deusto, Sàom Corresponding author. E-mail: icasas@deusto.es*
[b] *School of Finance, Zhongnan University of Economics and Law*
[c] *Department of Statistics and Instituto Flores de Lemus, Universidad Carlos III de Madrid and BRU-IUL, Instituto Universitário de Lisboa*

Abstract

We revisit the stock market return predictability using the variance risk premium and conditional variance as predictors of classical predictive regressions and time-varying coefficient predictive regressions. Also, we propose three new models to forecast the conditional variance and estimate the variance risk premium. Our empirical results show, first, that the flexibility provided by time-varying coefficient regressions often improve the ability of the variance risk premium, the conditional variance, and other control variables to predict stock market returns. Second, the conditional variance and variance risk premium obtained from varying coefficient models perform consistently well at predicting stock market returns. Finally, the time-varying coefficient predictive regressions show that the variance risk premium is a predictor of stock market excess returns before the global financial crisis of 2007, but its predictability decreases in the post global financial crisis period at the 3-month horizon. At the 12-month horizon, both the variance risk premium and conditional variance are predictors of stock excess returns during most of 2000-2015.

JEL-Classifications: C22; C51; C52; C53; G1

Keywords: Nonparametric methods, Predictability, Realized variance, Time-varying coefficient HAR-type models, Time-varying coefficient predictive regressions, Variance risk premium

1. Introduction

A recurring question in the financial literature is whether stock returns are predictable. No clear evidence of this predictability exists and the debate is divided into studies looking for the best predictors of stock returns and studies confirming their unpredictability. Bossaerts and Hillion (1999) observe that out-of-sample prediction of stock returns is not possible after running a set of regressions of 1-month excess returns on the dividend yield, short-term yield, and bond yield for 14 countries. Welch and Goyal (2008) conclude that standard predictive variables are not statistically significant, neither in-sample nor out-of-sample, after running similar regressions with the dividend price ratio, dividend yield, earnings-price ratio, dividend-earnings (payout) ratio, various interest rates and spreads, inflation rates, book-to-market ratio, and volatility as predictors of excess returns. In the same line, Ang and Bekaert (2007) express that long-run predictability is non-significant by studying whether various dividend yields can predict stock excess returns. In a more positive light, Campbell and Shiller (1988), Wilcox (2007), Lettau and Van Nieuwerburgh (2008) and references therein state that returns are predictable by the dividend- and earning-price, among others. Campbell et al. (2008) determine that the out-of-sample predictability of future stock excess returns is weak but meaningful by analyzing several subsample periods and several variables, including both accounting and equity market variables. Recently, Bollerslev et al. (2009, 2011, 2012, 2014), Drechsler and Yaron (2011), Galaix (2012), Bekaert and Hoerova (2014) and Kelly and Jiang (2014) indicate that the predictability of stock returns is stronger when the variance risk premium (VRP), alone or together with a set of potential predictors, is accounted for. Following this last line of research, we explore the predictive power of the VRP and conditional variance (CV) to predict stock market returns. We obtain the VRP by decomposing the squared CBOE volatility index (VIX) into the equity VRP and CV of the stock market measured over the next month, (see Bollerslev et al., 2009; Bekaert and Hoerova, 2014). The 1-month-ahead forecast of the realized variance (RV) is our proxy of the CV of the stock market measured over the next month. Accurate forecasts of the RV are, therefore, important for calculating the VRP.

Extensions of the heterogeneous autoregressive (HAR) models (Corsi, 2009) have recently been proposed to improve the forecasts of RV. Worthy of mention are the semivariance heterogeneous autoregressive (SHAR) model of Patton and Sheppard (2015) and the semivariance heterogeneous autoregressive-Q (HARQ and SHARQ) models of Bollerslev et al. (2016). The HAR and SHAR consider a constant relationship between the RV and its immediate past, whereas the HARQ and SHARQ assume that this relationship is a linear function of the past realized quarticity. Bianchi et al. (2017) analyze the U.S. stock returns with multiple macroeconomic risk factors whose coefficients vary with time. Following this idea, we propose to extend the HAR, SHAR, HARQ and SHARQ models by allowing their coefficients to vary depending on time or on the past realized quarticity. They are named in the paper as the TVHAR, TVSHAR, TVHARQ, and TVSHARQ and are estimated nonparametrically. Note that the TVHAR model has previously been proposed by Chen et al. (2018), but to the best of our knowledge, the other three models have not been used in the related literature. Altogether, we forecast the CV, and thereupon we estimate the VRP with the eight aforementioned models to use them as predictors of future stock market excess returns.

In the spirit of Bekaert and Hoerova (2014) and Bollerslev et al. (2014), we predict stock excess returns using classical regressions, but also time-varying coefficient regressions as in Dangl and Halling (2012) and Johannes et al. (2014). Structural changes in the parameters that relate future stock excess returns to a wide set of explanatory variables can occur, specially if the out-of sample period is long, and it could invalidate any inference based on full sample predictors (see, for instance Viceira, 1996; Paye and Timmermann, 2006; Pettenuzzo and Timmermann, 2011). In addition to the CV and VRP, we include other potential predictors as control variables in the predictive regressions, such as the real 3-month rate, the logarithm of the dividend yield, the credit spread, and the term spread. The predictive power of these variables is measured with the adjusted R^2s of the predictive regressions, root mean square errors (RMSE) and mean absolute errors (MAE).

Recently, other nonparametric methods, such as machine learning, are becoming popular for measuring asset risk premia. Gu et al. (2020b) show large economic gains to investors using machine learning forecasts. Also related, Gu et al. (2020a) propose a latent factor conditional asset pricing model estimated using machine leaning methods that generates out-of-sample pricing errors substantially smaller to those obtained with conventional factor models. This literature shows very promising results, but it is out of the scope of our work.

This paper makes both theoretical and empirical contributions to the literature. First, we propose a set of varying coefficient HAR-type models to forecast the RV and consequently, to estimate the VRP. Second, we use time-varying coefficient regressions to predict future stock excess returns and show that they improve the prediction ability of the VRP and CV. Finally, we observe that the power of the VRP and CV to predict stock market excess returns changes after the Global Financial Crisis (GFC). For moderate prediction horizons these variables lose completely their ability to predict stock market excess returns, while for long horizons some years after the GFC these variables recover their prediction ability.

The remainder of this paper is organized as follows: Section 2 defines the relationship between the variance risk premium and the realized variance. Section 3 presents the HAR-type models, the semiparametric varying coefficient HAR-type models, and reports the estimation results. Section 4 compares all models performance in predicting stock market returns. Finally, Section 5 summarizes our results.

2. The relationship between VRP and RV

For as long as high-frequency data have been available, RV has been a major focus of the research in financial econometrics. Among numerous applications, RV is considered a proxy of economic uncertainty (Bekaert and Hoerova, 2014) and has a critical role in estimating the VRP, which is a measure of risk aversion to uncertainty. RV of day t is defined by $RV_t = \sum_{i=1}^{M} r_{t,i}^2$, where $r_{t,i}$ is the price return at time i of day t and M is the total number of intraday time values.

The VRP is a risk compensation measure commonly used to determine the investor's degree of risk aversion to uncertainty and it is not observed. Generally, there are two approaches that can be used to estimate the VRP: the model-free and the parametric methodology. The first compares the risk-neutral probability implied by options prices to the forecast of the conditional variance of stock returns, and the second *"...consists of specifying the time-series (objective) and pricing (risk-neutral) dynamics of variance risk entering the pricing of asset returns and derivatives"* (see Li and Zinna, 2018). We use the first approach and represent the VRP as the difference between the squared VIX and the 1-month-ahead forecast of the RV, which is obtained with some known parametric models and new semiparametric proposals. (see Bekaert and Hoerova, 2014).[1]

The HAR model of Corsi (2009) is often used to forecast the RV, using past values of the daily RV aggregated at different time frequencies. This model does not account for the asymmetric response of volatility to positive and negative shocks (see Christie, 1982; Campbell and Hentschel, 1992; Bollerslev et al., 2006, for the concept of leverage effect).[2] Yet, negative shocks might have a stronger effect on volatility and, consequently, on the forecasts of the RV and on the estimates of the VRP, than positive shocks of the same magnitude. Investors dislike the uncertainty embedded in negative shocks because the likelihood of high losses increases. Thus, it is plausible to argue *a priori* that RV models that consider volatility asymmetry will increase the prediction ability of the VRP to predict future stock returns. Hence, we summarize the existing asymmetric HAR-type models and propose their related varying coefficient semiparametric models in the next section.

[1] VIX^2 is expressed in monthly percentages squared, that is, $VIX^2/12$. Note that VIX is the quoted VIX index level in annualized percent.

[2] Although asymmetry and leverage are not exactly the same, we use them interchangeably hereafter. Leverage is considered a special case of asymmetry (see, for example, McAleer, 2014).

3. HAR-type models

Although HAR models do not formally belong to the family of long memory models, it has been empirically observed that they are able to capture high persistence and to mimic the long memory feature of some financial time series. This makes them suitable for modeling the RV.

3.1. Parametric HAR models

Patton and Sheppard (2015) recently propose the semivariance heterogeneous autoregressive (SHAR) model, which is specified by

$$RV_t = \beta_0 + \beta_1^+ RV_{t-1}^+ + \beta_1^- RV_{t-1}^- + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t, \tag{1}$$

where $RV_t^+ = \sum_{i=1}^M r_{t,i}^2 I(r_{t,i} \geq 0)$ and $RV_t^- = \sum_{i=1}^M r_{t,i}^2 I(r_{t,i} \leq 0)$ (see Barndorff-Nielsen et al., 2010, for further details on the realized semivariance). Note that $r_{t,i}$ is the return of day t at moment i and $I(\cdot)$ is an indicator function that takes the value of one if the argument is true and zero otherwise. Moreover, $RV_{t-j|t-h} = \frac{1}{h+1-j}\sum_{i=j}^h RV_{t-i}$ with $j \leq h$. Therefore, RV_{t-1}, $RV_{t-1|t-5}$, and $RV_{t-1|t-22}$ correspond to the daily, weekly, and monthly lag RVs. The expression of the HAR model of Corsi (2009) is similar to (1), but $\beta_1^+ RV_{t-1}^+ + \beta_1^- RV_{t-1}^-$ is replaced by $\beta_1 RV_{t-1}$.

Bollerslev et al. (2016) show that although the RV is a consistent estimator of the integrated volatility under certain conditions, it is often affected by measurement errors in finite samples. This can cause biases on the parameter estimators of the HAR model. To account for these biases, Bollerslev et al. (2016) propose a new family of models named HARQ, which are basically HAR models whose coefficients are linear functions of the realized quarticity (RQ), that is:

$$\begin{aligned} RV_t =& \beta_0 + \left(\beta_1^+ + \beta_{1Q}^+ RQ_{t-1}^{1/2}\right) RV_{t-1}^+ + \left(\beta_1^- + \beta_{1Q}^- RQ_{t-1}^{1/2}\right) RV_{t-1}^- + \\ & \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t, \end{aligned} \tag{2}$$

where $RQ_t \equiv \frac{M}{3}\sum_{i=1}^M r_{t,i}^4$. In particular, this model is named SHARQ since it includes the two components of the RV. Again, replacing $\left(\beta_1^+ + \beta_{1Q}^+ RQ_{t-1}^{1/2}\right) RV_{t-1}^+ + \left(\beta_1^- + \beta_{1Q}^- RQ_{t-1}^{1/2}\right) RV_{t-1}^-$ by $\left(\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}\right) RV_{t-1}$ in equation (2) leads to the HARQ model. Bollerslev et al. (2016) show empirically that the HARQ and SHARQ models outperform the HAR model and its extensions, and in some cases even the SHAR model, in forecasting the RV.

In this paper, we do not consider the presence of jumps. Patton and Sheppard (2015) report that specifications that include jumps (for example, the HAR-J) show poor performance in comparison with the SHAR model.

3.2. Semiparametric HAR models

Recently, Chen et al. (2018) study a flexible model to estimate and forecast the RV of the S&P 500 returns by adding time-varying coefficients to the HAR model. This model, named TVHAR in this paper, is mathematically expressed by

$$RV_t = \beta_0(\tau) + \beta_1(\tau)RV_{t-1} + \beta_2(\tau)RV_{t-1|t-5} + \beta_3(\tau)RV_{t-1|t-22} + u_t, \tag{3}$$

for $t = 1, 2, \ldots, T$, $\tau = t/T \in [0, 1]$. Coefficients $\beta_i(\tau)$ are unknown functions of τ allowing for a lot of flexibility in the relationships between the current RV and its past values. Robinson (1989) introduced time-varying coefficients in multivariate regression, which has been generalized into various contexts of linear models over the years: (i) with stationary regressors (see Orbe et al., 2005; Cai, 2007; Gao, 2007; Kristensen, 2012; Phillips et al., 2017; Casas et al., 2019a, among many others) and with non-stationary regressors (see Gao and Phillips, 2013b,a; Casas et al., 2019b,a). It is a flexible approach but its forecast is inconsistent because no information from the dependent variable exists at time $T + h$, $h \geq 1$.

Bollerslev et al. (2016) indicate that the measurement error in the RV in finite samples is stronger for the daily lagged RV, although it still exists for the weekly and monthly lagged RVs, and they model the coefficients of the HARQ and SHARQ models as linear functions of the square root of the lagged realized quarticity. Instead, we extend this innovative idea allowing for linear and nonlinear relationships as required. The varying coefficient HARQ model, named TVHARQ, is given by

$$\begin{aligned} RV_t =& \beta_0(RQ_{t-1}^{1/2}) + \beta_1(RQ_{t-1}^{1/2})RV_{t-1} + \beta_2(RQ_{t-1}^{1/2})RV_{t-1|t-5} + \\ & \beta_3(RQ_{t-1}^{1/2})RV_{t-1|t-22} + u_t, \end{aligned} \tag{4}$$

where the expression of $\beta_i(RQ_{t-1}^{1/2})$ represents a coefficient that is an unknown function of the lagged realized quarticity and may vary for different values of it. Theoretical results for the nonparametric estimators of these coefficients are given in Cai et al. (2009) for stationary regressors, and stationary and nonstationary smoothing variables. The extension to the case of nonstationary regressors and smoothing variables is shown in Gao and Phillips (2013b,a) and Sun et al. (2013). On the other hand, Das (2005), Xiao (2009), Henderson et al. (2015) use this approach in the context of instrumental variables, cointegration and seemingly unrelated regressors, respectively. Whereas Sun et al. (2009); Feng et al. (2017); Feng et al. (2019); Dong et al. (2019) use it for panel data models.

Nonparametric methods are not convenient for models with numerous regressors because their rate of convergence decreases as the number of regressors increases. Semiparametric methods such as models (3)-(4), also called dynamic generalized linear models as in Hastie and Tibshirani (1993) and references therein, are good alternatives because they maintain a high level of flexibility and a satisfactory rate of convergence. The estimations and forecasts of models (3)-(4) are performed using the local constant (LC) or Nadaraya-Watson nonparametric estimator in this paper. This estimator requires that functions $\beta_i(\cdot)$ have a first derivative either at τ or in a region where $RQ_{t-1}^{1/2}$ is defined. The LC estimator is sometimes negative for finite samples, even when the dependent variable is always positive such as in our application. Asymptotically, this is not the case, and the number of negative values decreases as T increases. One way of avoiding this problem is to use the $\log RV_t$ series, as in Chen et al. (2018). However, we use the RV_t series following the example of Bekaert and Hoerova (2014) and interpolate the possible few negative estimates.

Generalizing, we propose a varying coefficient model given by

$$RV_t = M_t \beta(z_t) + u_t. \tag{5}$$

The regressors' vector at time t is $M_t = (1, RV_{t-1}, RV_{t-1|t-5}, RV_{t-1|t-22})'$ for the TVHAR and TVHARQ models and $M_t = (1, RV_{t-1}^{+}, RV_{t-1}^{-}, RV_{t-1|t-5}, RV_{t-1|t-22})'$ for the TVSHAR and TVSHARQ models. The coefficient estimates minimize the mean squared error of Equation (5) and are calculated by

$$\hat{\beta}_t = S_T(z_t)^{-1} T_T(z_t),$$

where $S_T(z_t) = T^{-1} \sum_{t=1}^{T} M_t M_t' K_b(Z - z_t)$ and $T_T = \sum_{t=1}^{T} M_t K_b(Z - z_t) RV_t$. The conditional variable for the TVHAR and TVSHAR models is $z_t = t/T$, whereas is $z_t = RQ_{t-1}^{1/2}$ for the TVHARQ and TVSHARQ models. The kernel function, $K_b(u) = K(u/b)/b$, is a symmetric continuous function with compact support and bandwidth b. The bandwidth should reach zero at a slower rate than T approaches infinity. In this paper, the bandwidth is selected by leave-k-out cross-validation (Chu and Marron, 1991) with $k = \lfloor T/3 \rfloor$ to ensure the independence of the subsamples.

3.3. Estimation results

Our empirical results are based on the daily S&P 500 index. High-frequency prices for the index are obtained from 1-minute close prices.[3] Our complete sample ranges from January 2, 1990 until December 31, 2019 with 7557 daily observations of the RV and its asymmetric measures.

Figure 1 displays the daily realized semivariance. We observe that the RV^- shows larger values of volatility than the RV^+. This is expected since negative returns affect the volatility more than positive returns of the same magnitude. The highest spikes occur on October 10, 2008 in both series when the Asian and European stock markets closed with large losses (for example, the "Yamato Life" insurance company filed for bankruptcy). October 10, 2008 is also preceded and followed by periods of high volatility that correspond to the 2008 global financial crisis.

Figure 1: Daily S&P 500 realized semivariance. Negative (left panel) and positive (right panel) measures ranging from January 2, 1990 untill December 31, 2019.

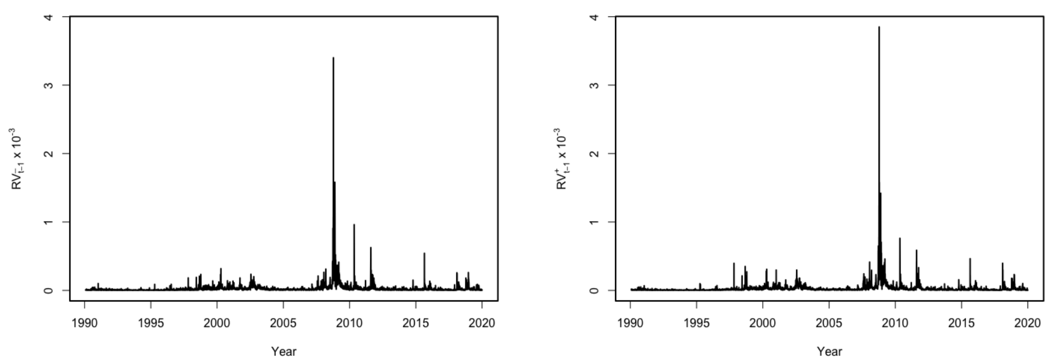


Table 1 lists the coefficient estimates of the HAR, SHAR, HARQ and SHARQ models, their standard errors in brackets, p-values, adjusted R^2, RMSE, and mean values of the QLIKE loss function. Similarly, Table 2 lists the pseudo-R^2, RMSE, and QLIKE of the semiparametric specifications.[4] Estimation is done for three different samples: the full sample from January 2, 1990 until December 31, 2019; the pre-GFC sample from January 2, 1990 until December 31, 2006; and the post-GFC sample

[3] Detailed information can be found on `http://download-stock-data.webs.com/`.

[4] The pseudo-R^2 is calculated with the classical equation, $R^2 = 1 - \frac{\sum_{t=1}^{T}(y_t - \hat{y})^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$, where y_t is the dependent variable, $\bar{y}$ is its mean and $\hat{y}_t$ are the fitted values. The QLIKE loss function of Patton (2011) is given by $QLIKE_t \equiv \frac{RV_t}{E(RV_t)} - \log \frac{RV_t}{E(RV_t)} - 1$, where $E(RV_t)$ corresponds to the estimated realized variance at day t obtained from the selected models.

from January 2, 2007 until December 31, 2019. The largest differences among samples are found in the $\widehat{\beta_1}^-$, $\widehat{\beta_1}^+$ and $\widehat{\beta_{1Q}}^+$. The two former estimates show an important increment in magnitude during the post-GFC due to the market instability. In general, the β_1^- estimates are larger in absolute value than the β_1^+ estimates for all samples, and are always statistically significant. Whereas the β_1^+ estimate is non-significant during the pre-GFC. This result agrees with Patton and Sheppard (2015) who obtain a similar evidence for the SHAR model. We also observe that the sum of $\hat{\beta}_2$ and $\hat{\beta}_3$ is smaller for the HARQ and SHARQ than for the other models, suggesting that the importance of the RV components corresponding to one week and one month lags decreases when the RQ plays a role in the model. In fact, $\widehat{\beta_3}$ is positive and non-significant for the HARQ and SHARQ during the pre-GFC, but negative and statistically significant during the post-GFC. In addition, β_{1Q} and β_{1Q}^- are statistically significant in all periods, but β_{1Q}^+ is only significant in the post-GFC period.

Figure 2 displays the estimates of the HARQ and TVHARQ that explain the relationship between RV_t and RV_{t-1}, i.e., $\hat{\beta}_1(RQ_{t-1}^{1/2})$ in the TVHARQ and $\hat{\beta}_1 + \hat{\beta}_{1Q}RQ_{t-1}^{1/2}$ in the HARQ. There are obvious differences, not only between the two models, but also between results at different periods. The HARQ estimates (dashed line) are mostly greater than the TVHARQ estimates (continuous line) for all RQ values and periods. The 95% confidence intervals obtained from the two models overlap for most of the pre-GFC period. Note that estimates during the pre-GFC period are less volatile than during the post-GFC. Furthermore, there is a clear decreasing trend in the TVHARQ estimates as $RQ_{t-1}^{1/2}$ increases during the post-GFC and during the full sample. In all periods, the 95% confidence intervals of the TVHARQ coefficients include the zero for large values of $RQ_{t-1}^{1/2}$, suggesting that the relationship between RV_t and RV_{t-1} is not significant for days with very large realized quarticity. Regarding the HARQ model, the relationship between RV_t and RV_{t-1} is mostly positive and significant, but it also diminishes in significance and sign as RQ increases.

Table 1: Parametric models' estimation results (period 1990-2019)

The table reports estimated parameters of parametric HAR-type models for the daily frequency, their p-values in parentheses, and adjusted R^2s of the estimated models together with RMSE and QLIKE as measures of goodness of fit. The value of the RMSE is multiplied by 1000. *,**,*** means that the variable is statistically significant at 10%, 5% and 1%, respectively.

	S&P 500											
	Full sample				**pre-GFC**				**post-GFC**			
	HAR	SHAR	HARQ	SHARQ	HAR	SHAR	HARQ	SHARQ	HAR	SHAR	HARQ	SHARQ
β_0	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000*	0.000	0.000**	0.000*	0.000	0.000**
	(0.000)	(0.004)	(0.000)	(0.002)	(0.000)	(0.005)	(0.093)	(0.182)	(0.013)	(0.052)	(0.136)	(0.029)
β_1	0.207***		0.784***		0.393***		0.600***		0.199***		0.832***	
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
β_1^+		−1.068***		−0.347***		−0.123***		0.053		−1.324***		−0.810***
		(0.000)		(0.000)		(0.000)		(0.287)		(0.000)		(0.000)
β_1^-		1.593***		1.936***		0.954***		1.065***		1.853***		2.555***
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
β_2	0.494***	0.588***	0.354***	0.346***	0.285***	0.302***	0.209***	0.251***	0.530***	0.497***	0.349***	0.338***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_3	0.183***	0.223***	−0.028	−0.021	0.263***	0.240***	0.196***	0.195***	0.182***	0.177***	−0.051*	−0.049*
	(0.000)	(0.000)	(0.148)	(0.267)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.088)	(0.094)
β_{1Q}			−83.338***				−196.49***				−88.156***	
			(0.000)				(0.000)				(0.000)	
β_{1Q}^+				−29.897				−18.310				166.2**
				(0.472)				(0.671)				(0.019)
β_{1Q}^-				−131.52***				−310.68***				−361.44***
				(0.005)				(0.000)				(0.000)
Adj. R^2	0.561	0.588	0.629	0.647	0.649	0.671	0.665	0.678	0.546	0.578	0.622	0.642
RMSE	0.103	0.100	0.095	0.092	0.028	0.027	0.027	0.026	0.154	0.148	0.140	0.136
QLIKE	0.206	0.224	0.176	0.266	0.134	0.130	0.127	0.125	0.285	0.274	0.248	0.481

Figure 2: Estimates of the daily relationship between RV_t and RV_{t-1} from the HARQ (dashed) and the TVHARQ (continuous). The bands correspond to the 95% confidence intervals.

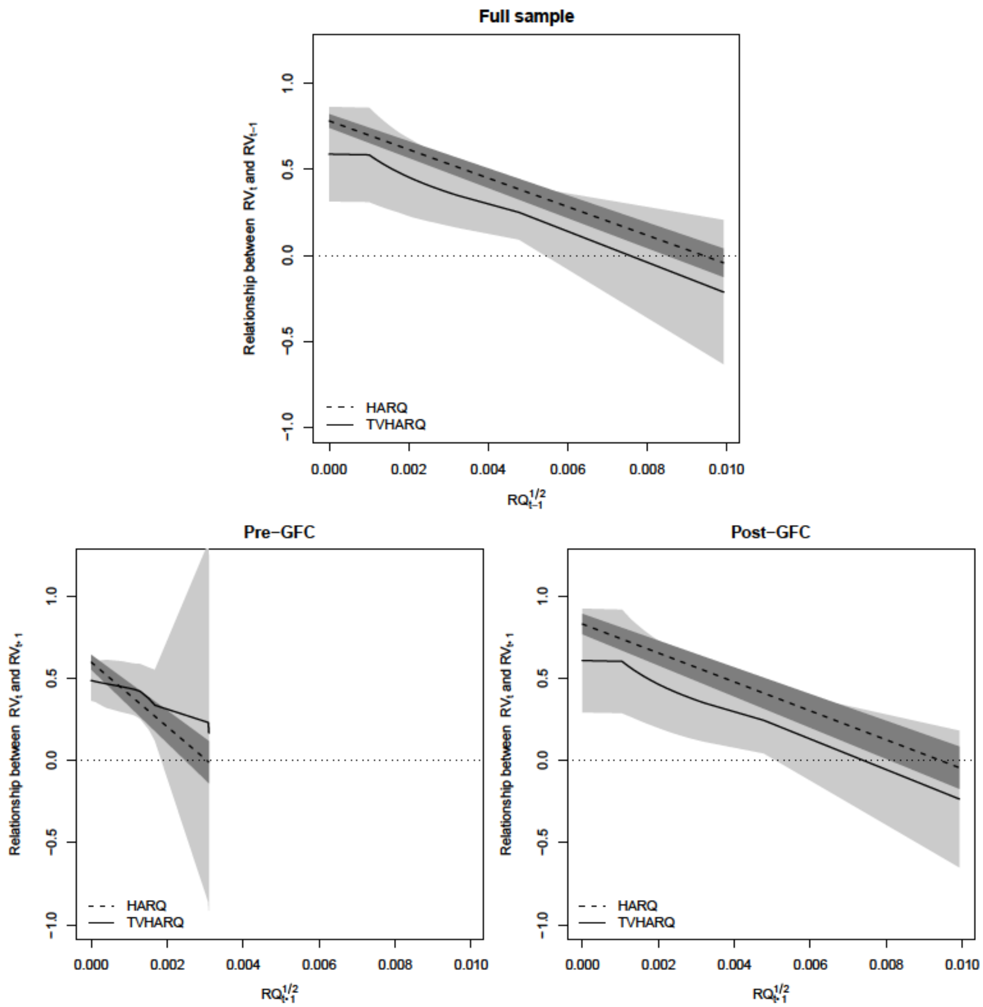


Table 2: Semiparametric models' goodness-of-fit (period 1990-2019)

The table provides summaries of coefficient distributions and in-sample measures of goodness-of-fit for the semiparametric models. The value of the RMSE is multiplied by 1000. The frequency is daily.

S&P 500						
	TVHAR	TVSHAR		TVHARQ	TVSHARQ	
	$\beta_1(t)$	$\beta_1^+(t)$	$\beta_1^-(t)$	$\beta_1(RQ_t)$	$\beta_1^+(RQ_t)$	$\beta_1^-(RQ_t)$
Full sample						
Min.	0.207	-1.072	1.597	-0.214	-1.423	1.220
Q1	0.207	-1.072	1.597	0.589	-0.450	1.658
Median	0.207	-1.072	1.598	0.589	-0.450	1.658
Mean	0.207	-1.072	1.598	0.588	-0.451	1.658
Q3	0.207	-1.072	1.598	0.589	-0.450	1.659
Max.	0.207	-1.072	1.598	0.589	-0.450	1.659
pseudo-R^2	0.561	0.589		0.625	0.642	
RMSE	0.103	0.100		0.095	0.093	
QLIKE	0.146	0.198		0.132	0.156	
pre-GFC						
Min.	0.393	-0.179	0.909	0.170	-0.130	0.831
Q1	0.393	-0.146	0.965	0.485	-0.045	0.965
Median	0.393	-0.138	0.976	0.486	-0.044	0.965
Mean	0.393	-0.141	0.972	0.485	-0.044	0.965
Q3	0.393	-0.136	0.985	0.488	-0.043	0.965
Max.	0.393	-0.108	0.992	0.488	-0.043	0.966
pseudo-R^2	0.649	0.672		0.660	0.676	
RMSE	0.028	0.027		0.027	0.026	
QLIKE	0.113	0.114		0.112	0.113	
post-GFC						
Min.	0.144	-1.338	1.509	-0.233	-1.234	0.963
Q1	0.190	-1.330	1.848	0.609	-0.594	1.839
Median	0.475	-1.316	1.852	0.610	-0.594	1.840
Mean	0.394	-1.261	1.834	0.608	-0.596	1.839
Q3	0.546	-1.284	1.853	0.610	-0.594	1.841
Max.	0.631	-0.440	1.853	0.610	-0.587	1.841
pseudo-R^2	0.549	0.579		0.616	0.634	
RMSE	0.153	0.148		0.141	0.138	
QLIKE	0.168	0.182		0.152	0.160	

We observe in Table 1 that the SHARQ is the parametric model with the largest adjusted R^2 and the smallest RMSE in all periods, while the HARQ is the model with the smallest QLIKE values in the full sample and post-GFC periods. Similarly and looking at Table 2, the TVSHARQ reports the largest pseudo-R^2 and the smallest RMSE in all sample periods, while the TVHARQ reports the smallest QLIKE in all sample periods. Like in Bollerslev et al. (2016), we observe that adding the realized quarticity to HAR-type models improves the goodness-of-fit of the models, in particular, if we allow for varying coefficients. In the next section, we study whether a better estimation of the RV also results in a more accurate prediction of stock market returns.

4. Stock market return predictability

Bollerslev et al. (2009), Bekaert and Hoerova (2014), Bollerslev et al. (2014), and Bollerslev et al. (2015) adopt a linear model for stock market return prediction. They observe that the VRP is a significant predictor in their regressions. Here in, the VRP is used as a proxy of risk aversion and the RV is used as a proxy of economic uncertainty (see Bekaert and Hoerova, 2014). The monthly VRP is calculated using the monthly squared VIX and the 1-month-ahead forecast of RV.[5]

To avoid the misspecification of univariate regressions, Bekaert and Hoerova (2014) consider regressions in which the VRP and CV are simultaneously predictors of future stock excess returns in addition to: (1) the real 3-month rate (3MTB), which is the 3-month T-bill minus CPI inflation; (2) the logarithm of the dividend yield (log(DY)); (3) the credit spread (CS) obtained using the difference between Moody's BAA and AAA bond yield indexes; and (4) the term spread (TS) calculated using the difference between the 10-year and 3-month treasury yields.

[5] In order to obtain the 1-month ahead monthly forecast of the RV, we use the aforementioned models to estimate the $RV_t^{(22)}$. For example, using the HARQ model we find the estimate of $RV_t^{(22)}$ with

$$RV_t^{(22)} = \beta_0 + \beta_1^{+} RV_{t-22}^{+(1)} + \beta_1^{-} RV_{t-22}^{-(1)} + \beta_2 RV_{t-22}^{(5)} + \left(\beta_3 + \beta_{3Q}\sqrt{RQ_{t-22}^{(22)}}\right) RV_{t-22}^{(22)} + u_t,$$

where $RV_t^{(22)} = \sum_{j=1}^{22} RV_{t-j+1}$, $RV_t^{\pm(h)} = \frac{22}{h}\sum_{j=1}^{h} RV_{t-j+1}^{\pm}$, $RQ_t^{(h)} = \frac{22}{h}\sum_{j=1}^{h} RQ_{t-j+1}$, and $RV_t^{(h)} = \frac{22}{h}\sum_{j=1}^{h} RV_{t-j+1}$. Bollerslev et al. (2016) show that when forecasting the monthly RV, the monthly lag becomes relatively more important than the daily and weekly lags. Therefore, we correct the coefficient of $RV_{t-22}^{(22)}$ using the monthly RQ instead. Similar transformations apply to the other models.

Conforming with the aforementioned articles, we run the following regressions to predict stock market excess returns for horizons $h = 1, 3$ and 12 months

$$h^{-1} \sum_{j=1}^{h} r_{t,t+j} = a^h + b^h VRP_t + c^h CV_t + d^h X_t + u_{t,t+h} \quad t = 1, \ldots, T, \tag{6}$$

where $r_{t,t+j}$ denotes the $j = 1, \ldots, h$-month excess returns on the S&P 500. The control variables listed above are expressed in annualized percentages and are represented by the X_t vector. Note that t, hereafter, corresponds to the observation of month t. Bekaert and Hoerova (2014) state that the overlapping in the monthly data can generate serial correlation and biased standard errors. We use robust Newey-West standard errors to take care of this issue. Even though, high adjusted-R^2s are expected for long horizons since they tend to increase with the horizon when the predictors are persistent (see, for example, Fama and French, 1988; Campbell et al., 1997; Campbell, 2001).

Empirically studies like Paye and Timmermann (2006), Lettau and Van Nieuwerburgh (2008), Henkel et al. (2011), Dangl and Halling (2012) and Johannes et al. (2014) show that the relationship between the returns and some predictors varies with time. Our time-varying coefficient regressions are

$$h^{-1} \sum_{j=1}^{h} r_{t,t+j} = a^h(\tau) + b^h(\tau) VRP_t + c^h(\tau) CV_t + d^h(\tau) X_t + u_{t,t+h}. \tag{7}$$

The main difference with Equation (6) is that these coefficients may be a constant, a linear function of $\tau = (0, 1/T, \ldots, T)$ or a non-linear function of τ, depending on the information in the data. The coefficients in Equation (7) are estimated using the local linear kernel estimator (Fan and Gijbels, 1996) for horizons 1-, 3- and 12-months. The bandwidths are selected using leave-k-out cross-validation (Chu and Marron, 1991) for k equal to 2 months of data. In practice, the estimates and plots are obtained using the R package `tvReg` (Casas and Fernandez-Casal, 2019).

We analyze the performance of the proposed estimators of the VRP and CV in predicting excess returns from April 1994 until December 2019. Table 3 reports the adjusted-R^2s, the pseudo-R^2s, the RMSEs and the mean absolute errors (MAE) obtained from the classical predictive regressions (PR) and the time-varying coefficient predictive regressions (TVPR). Their statistical difference in predictive accuracy is tested with the model confident set procedure (MCS) proposed by Hansen et al. (2011). The procedure consists of a sequence of statistic tests to construct the "Superior Set of Models" (SSM) at a given confidence level, $100(1-\alpha)\%$. Models in the

SSM have statistically the same predictive accuracy, measured by a loss function, at significant level α. We use the root mean squared error and mean absolute error loss functions.

Table 4 displays the 80% SSM obtained from comparing all our regressions (see Granger, 1996, that considers confidence levels of 50% and 80% adequate *to provide "warning" signals that the model is breaking down*). The capability to adapt to changes in the coefficients over time of the TVPR appear to be crucial in the prediction of stock market returns for 12-month horizon. Thus, the last row of Table 4 shows that the only specification in the 80% SSM for both loss functions is the TVPR where the VRP and CV are obtained using the TVHAR model. This adaptability of the coefficients is not as important for shorter horizons. Regarding the 1-month horizon, most of the VRP and CV obtained with the HAR-type and TVHAR-type models perform equally well at predicting future excess returns, at the confidence level 80%. Nevertheless, the combination of TVHAR predictors with TVPR achieves the smallest RMSE and MAE values; see Table 3. At the 3-month horizon, all TVPR have the same predictive ability, but only the PR whose VRP and CV are obtained with the RV, TVHARQ and TVSHARQ are in the 80% SSM for the RMSE loss function. Furthermore, the combination of SHAR predictors with TVPR provides the smallest RMSE. Finally, within the class of varying coefficient models, the TVHAR is the model with the smallest RMSE and MAE for three out of six cases; see Table 3.

Figure 3 plots the VRP and CV coefficient estimates of our predictive regressions (PR and TVPR) with the smallest RMSE for the three horizons. The figure shows, first, that the TVPR coefficient estimates are linear for the 1-month horizon and non-linear for the other horizons. Second, the width of the 95% confidence intervals of all estimates decreases as the horizon increases, meaning that there is less uncertainty in the prediction of future excess returns for longer horizons. Third, the sign of the estimates associated with the VRP and CV predictors are always positive in the PR; however, their sign vary over time and over the prediction horizon in the TVPR. Finally, the VRP and CV are predictors of excess returns in the PR at all horizons. Yet, in the TVPR and based on the 95% confidence intervals, we can observe that the VRP is a predictor until the end of 2005 at the 3-month horizon; and both the VRP and CV are predictors of excess returns during most of 2000-2015 at the 12-month horizon, with the exception of the period between November, 2009 and April, 2012 when the 95% confidence interval of the VRP coefficient estimates includes the zero and the period between September, 2010 and November, 2011 for the CV coefficient estimates.

Table 3: Stock excess return regressions: April 1994–December 2019

Predictive regressions (monthly observations) with variance risk premium, conditional variance, real 3-month rate, the logarithm of dividend yield, credit spread, and term spread. RMSE corresponds to root mean square error and MAE to mean absolute error. PR stands for parametric predictive regressions and TVPR corresponds to time-varying coefficient predictive regressions. Cells in gray correspond to the best values of the measures.

	RV	HAR	SHAR	HARQ	SHARQ	TVHARQ	TVSHARQ	TVHAR	TVSHAR
Panel A: 1-month horizon									
PR									
Adj. R^2	0.230	0.242	0.239	0.239	0.236	0.227	0.227	0.255	0.241
RMSE	6.495	6.470	6.477	6.477	6.482	6.500	6.500	6.442	6.472
MAE	31.065	30.970	31.032	31.037	31.047	31.223	31.188	30.878	31.075
TVPR									
Pseudo R^2	0.295	0.306	0.304	0.303	0.302	0.290	0.296	0.319	0.314
RMSE	6.386	6.359	6.364	6.366	6.370	6.395	6.383	6.330	6.341
MAE	30.681	30.147	30.183	30.334	30.333	30.551	30.366	30.002	30.114
Panel B: 3-month horizon									
PR									
Adj. R^2	0.518	0.495	0.495	0.504	0.504	0.512	0.525	0.496	0.496
RMSE	4.755	4.811	4.812	4.789	4.791	4.770	4.739	4.809	4.809
MAE	16.392	16.915	16.947	16.681	16.701	16.863	16.809	16.949	16.898
TVPR									
Pseudo R^2	0.586	0.601	0.601	0.576	0.576	0.581	0.588	0.574	0.575
RMSE	4.602	4.559	4.558	4.629	4.629	4.616	4.594	4.634	4.632
MAE	15.768	15.531	15.548	16.056	16.067	16.223	16.203	16.126	16.097
Panel C: 12-month horizon									
PR									
Adj. R^2	0.817	0.815	0.815	0.815	0.815	0.816	0.816	0.815	0.815
RMSE	3.233	3.240	3.240	3.241	3.241	3.237	3.236	3.241	3.241
MAE	7.834	7.985	7.987	7.958	7.960	7.951	7.957	7.957	7.962
TVPR									
Pseudo R^2	0.875	0.875	0.875	0.875	0.875	0.879	0.880	0.899	0.860
RMSE	2.954	2.950	2.952	2.953	2.954	2.926	2.924	2.798	3.038
MAE	6.704	6.696	6.716	6.708	6.717	6.576	6.577	5.721	7.318

All in all, we conclude that: (i) the VRP and CV are not predictors of excess returns in the TVPR at 1-month horizon; (ii) only the VRP is a predictor of excess returns at the 3-month horizon and only before the GFC; and (iii) the VRP and CV are both predictors of future excess returns at the 12-month horizon from the first half of 2000 until 2010 and from the end of 2011 until the end of 2014.

Table 4: P-values of the MCS procedure.

P-values of the MCS procedure for the predictive regressions in Equation (6) and the time-varying predictive regressions in Equation (7). Results are from the MCS with two different loss functions, the root mean squared error (RMSE) and the mean absolute error (MAE), and 5000 bootstrap samples. PR stands for predictive regressions and TVPR for time-varying predictive regressions. The bottom part of the table summarizes the models that belong to the Superior Set of Models at the 80% confidence level.

		RMSE			MAE		
		1-month	3-month	12-month	1-month	3-month	12-month
PR	RV	0.084	0.349	0.013	0.494	0.564	0.053
	HAR	0.443	0.124	0.006	0.690	0.209	0.034
	SHAR	0.364	0.126	0.006	0.605	0.186	0.032
	HARQ	0.363	0.186	0.007	0.560	0.424	0.046
	SHARQ	0.236	0.173	0.007	0.533	0.370	0.043
	TVHARQ	0.091	0.242	0.010	0.301	0.242	0.042
	TVSHARQ	0.084	0.349	0.009	0.372	0.314	0.039
	TVHAR	0.684	0.150	0.008	0.690	0.208	0.049
	TVSHAR	0.443	0.173	0.009	0.605	0.267	0.042
TVPR	RV	0.753	0.627	0.048	0.690	0.564	0.053
	HAR	0.865	0.775	0.049	0.966	1.000	0.053
	SHAR	0.753	1.000	0.048	0.910	0.564	0.053
	HARQ	0.753	0.349	0.048	0.691	0.564	0.053
	SHARQ	0.753	0.349	0.048	0.690	0.564	0.053
	TVHARQ	0.684	0.349	0.049	0.690	0.424	0.053
	TVSHARQ	0.753	0.627	0.049	0.691	0.564	0.053
	TVHAR	1.000	0.349	1.000	1.000	0.478	1.000
	TVSHAR	0.865	0.349	0.025	0.966	0.564	0.049

	SSM at 80% confidence level	
	SE	AE
1-month	All in **TVPR** + HAR,SHAR, HARQ, SHARQ, TVHAR, and TVSHAR in **PR**	All in **TVPR** + All in **PR**
3-month	All in **TVPR** + RV, TVHARQ, TVSHARQ in **PR**	All in **TVPR** + All **except** SHAR in **PR**
12-month	TVHAR in **TVPR**	TVHAR in **TVPR**

Figure 3: VRP annd CV coefficient estimates obtained from either classical and time-varying predictive regressions. The VRP and CV are obtained from models: TVHAR (1-month horizon), SHAR (3-month horizon) and TVHAR (12-month horizon).

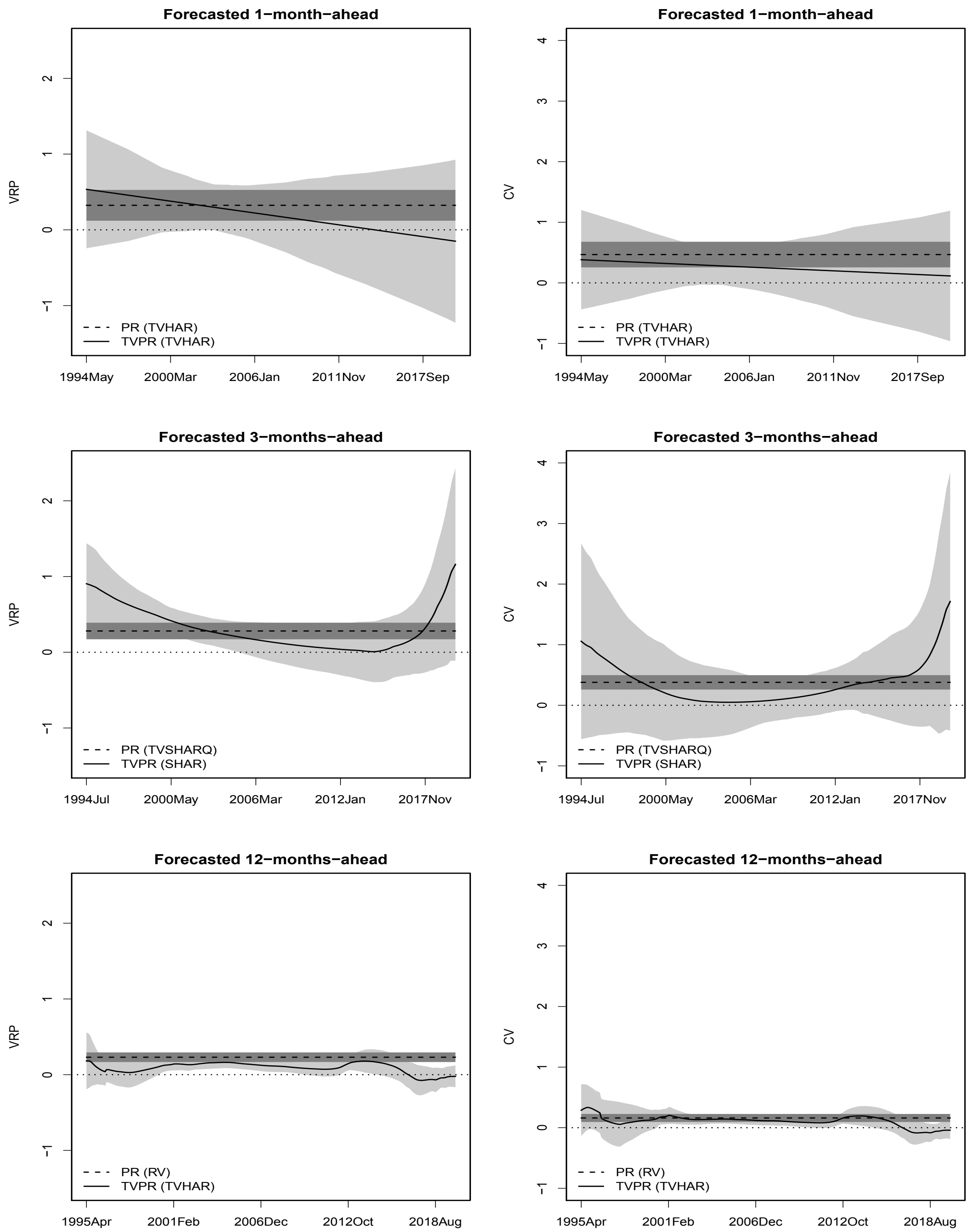
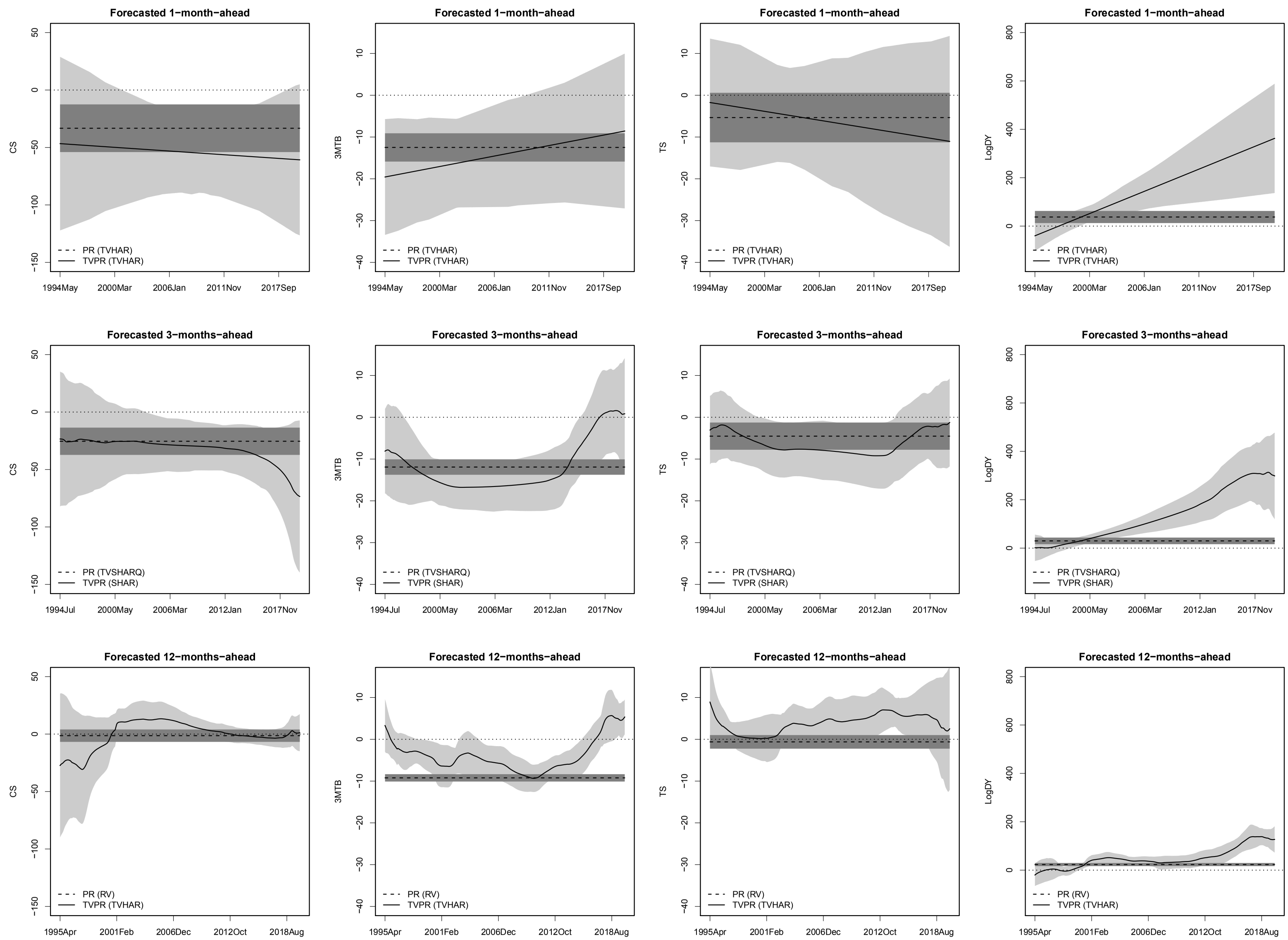


Figure 4: Coefficient estimates of extra potential predictors obtained from either classical or time-varying predictive regressions. The estimates are obtained from models: TVHAR (1-month horizon), SHAR (3-month horizon) and TVHAR (12-month horizon).



Regarding the set of extra potential predictors, Figure 4 plots the coefficient estimates of the credit spread, the real 3-month rate, the term spread and the logarithm of the dividend yield. We notice that the sign of the coefficient estimates of most predictors are the same in the PR and TVPR. Focusing on the results from the TVPR, the CS predicts future returns with a negative sign at the 1- and 3-month horizons from early 2000s until the end of 2019, but it loses its predictive power at the 12-month horizon. The 3MTB is also a predictor of excess returns at all horizons, but its predictability is not homogeneous over the out-of-sample period. For instance, at the 1-month horizon, the 3MTB is a significant predictor from late 1990s until around year 2015, and the sign of its coefficient is negative during most of the sample. On the other hand, the TS is not a predictor of excess returns at the 1-month horizon, but it is at the 3-month horizon from early 2000s till around 2015, with a negative sign. Finally, the logarithm of the DY is a predictor with a very steep increasing trend in its coefficient estimates during the whole period and all horizons. These are positive and statistically significant from year 2000.

5. Conclusion

The purpose of this paper is to find the best methodologies and models to obtain an accurate monthly forecast of the realized variance and an accurate monthly estimate of the variance risk premium, and consequently, to increase the predictability power of these variables in predicting future stock excess returns. We compare a total of eight forecasters of the realized variance. To the best of our knowledge, three of them have never been used before in the related literature. Four of those forecasters are HAR-type models and the other four are their extensions with varying coefficients, which may vary over time or react to variations in the past realized quarticity. The function form of these coefficients are estimated by kernel smoothing from the data, allowing for very flexible relationships.

We start by comparing the performance of these models to estimate the daily realized variance in-sample for three time periods: The period before the global financial crisis from January 1990 until December 2006; the period posterior to this crisis from January 2007 until December 2019; and the union of the two previous sample periods, the full sample. We find that the best strategy to estimate the daily realized variance is often to use semiparametric models whose coefficients vary as a function of the realized quarticity. Including the asymmetric response of the volatility to positive and negative returns also improves the goodness-of-fit of both parametric and semiparametric models. Then, we proceed to obtain the monthly forecast of the realized variance, and the variance risk premium which is used to predict the excess returns with both classical and time-varying coefficient predictive regressions.

The out-of-sample results are clear: first, time-varying coefficient predictive regressions improve the predictive ability of the variance risk premium and conditional variance to predict future excess returns. Second, and most importantly, some parametric HAR models have similar predictive ability to varying coefficient HAR models in predicting excess returns, except if the forecasting horizon is long. In this case, we need both varying coefficient models together with time-varying coefficient predictive regressions. Finally, given that time-varying coefficient predictive regressions allow the coefficients of the predictors to change over time, we can detect the periods in which the variance risk premium and conditional variance are predictors of the future excess returns and the sign of this prediction. We observe that for the majority of the horizons the sign of the prediction for both the variance risk premium and conditional variance is positive which supports the risk compensation theory, and that the variance risk premium is a predictor of the future excess returns at the 3-month horizon before the last global financial crisis, and at the 12-month horizon before the great financial crisis and during most of 2000-2015. The conditional variance is only a predictor of the excess returns at the 12-month horizon.

Acknowledgements

We thank Kevin Berg Kartaszewicz-Grell for an interesting conversation and participants of the NIPE seminar for the all the comments. Veiga acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities, research project PGC2018-096977-B-l00, from the Agencia Estatal de Investigación PID2019-108079GB-C21/AIE/10.13039/501100011033 and from Fundação para a Ciência e a Tecnologia, grant UIDB/00315/2020.

References

Ang, A., Bekaert, G., 2007. Stock return predictability: Is it there? The Review of Financial Studies 20 (3), 651–707.

Barndorff-Nielsen, O. E., Kinnebrock, S., Shephard, N., 2010. Measuring downside risk? Realised semivariance. In: Bollerslev, T., Russell, J., Watson, M. (Eds.), Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle. Oxford University Press, New York.

Bekaert, G., Hoerova, M., 2014. The VIX, the variance premium and stock market volatility. Journal of Econometrics 183 (2), 181–192.

Bianchi, D., Guidolin, M., Ravazzolo, F., 2017. Macroeconomic factors strike back: A Bayesian change-point model of time-varying risk exposures and premia in the U.S. cross-section. Journal of Business & Economic Statistics 35 (1), 110–129.

Bollerslev, T., Gibson, M., Zhou, H., 2011. Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. Journal of Econometrics 160 (1), 235–245.

Bollerslev, T., Litvinova, J., Tauchen, G., 2006. Leverage and volatility feedback effects in high-frequency data. Journal of Financial Econometrics 4 (3), 353–384.

Bollerslev, T., Marrone, J., Xu, L., Zhou, H., 2014. Stock return predictability and variance risk premia: Statistical inference and international evidence. Journal of Financial and Quantitative Analysis 49 (3), 633–661.

Bollerslev, T., Patton, A. J., Quaedvlieg, R., 2016. Exploiting the errors: A simple approach for improved volatility forecasting. Journal of Econometrics 192 (1), 1–18.

Bollerslev, T., Sizova, N., Tauchen, G., 2012. Volatility in equilibrium: Asymmetries and dynamic dependencies. Review of Finance 16 (1), 31–80.

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. The Review of Financial Studies 22 (11), 44–63.

Bollerslev, T., Xu, L., Zhou, H., 2015. Stock return and cash flow predictability: The role of volatility risk. Journal of Econometrics 187 (2), 458–471.

Bossaerts, P., Hillion, P., 1999. Implementing statistical criteria to select return forecasting models: What do we learn? The Review of Financial Studies 12 (2), 405–428.

Cai, Z., 2007. Trending time-varying coefficient time series with serially correlated errors. Journal of Econometrics 136, 163–188.

Cai, Z., Li, Q., Park, J. Y., 2009. Functional-coefficient models for nonstationary time series data. Journal of Econometrics 148, 101–113.

Campbell, J. Y., 2001. Why long horizons? A study of power against persistent alternatives. Journal of Empirical Finance 8, 459–491.

Campbell, J. Y., Hentschel, L., 1992. No news is good news: A asymmetric model of changing volatility in stock returns. Journal of Financial Economics 31, 281–318.

Campbell, J. Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. The Journal of Finance 63 (6), 2899–2939.

Campbell, J. Y., Lo, A. W., MacKinlay, A. C., Adamek, P., Viceira, L. M., 1997. The Econometrics of Financial Markets. Princeton University Press.

Campbell, J. Y., Shiller, R. J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies 1 (3), 195–228.

Casas, I., Fernandez-Casal, R., 2019. tvreg: Time-varying coefficients linear regression for single and multi-equations in r. Tech. rep.
URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3363526

Casas, I., Ferreira, E., Orbe, S., 03 2019a. Time-varying coefficient estimation in SURE models. application to portfolio management. Journal of Financial Econometrics, 1–39.

Casas, I., Gao, J., Peng, B., Xie, S., 2019b. Time-varying income elasticities of healthcare expenditure for the OECD and Eurozone. Monash Econometrics and Business Statistics Working Papers 28/19, Monash University, Department of Econometrics and Business Statistics.
URL https://ideas.repec.org/p/msh/ebswps/2019-28.html

Chen, X. B., Gao, J., Li, D., Silvapulle, P., 2018. Nonparametric estimation and forecasting for time-varying coefficient realized volatility models. Journal of Business & Economic Statistics 36 (1), 88–100.

Christie, A., 1982. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. Journal of Financial Economics 10, 407–432.

Chu, C.-K., Marron, J. S., 1991. Comparison of two bandwidth selectors with dependent errors. The Annals of Statistics 19, 1906–1918.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7 (2), 174–196.

Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. Journal of Financial Economics 106 (1), 157–181.

Das, M., 2005. Instrumental variables estimators of nonparametric models with discrete endogenous regressors. Journal of Econometrics 124, 335–361.

Dong, C., Gao, J., Peng, B., 2019. Estimation in a semiparametric panel data model with nonstationarity. Econometric Reviews 38 (8), 961–977.

Drechsler, I., Yaron, A., 2011. What's vol got to do with it. The Review of Financial Studies 24 (1), 1–45.

Fama, E. F., French, K., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.

Fan, J., Gijbels, I., 1996. Local Polynomial Modeling and Its Applications. Hapman and Hall, London.

Feng, G., Gao, J., Peng, B., 2019. An integrated panel data approach to modelling economic growth. arXiv e-prints, arXiv:1903.07948.

Feng, G., Gao, J., Peng, B., Zhang, X., 2017. A varying-coefficient panel data model with fixed effects: Theory and an application to us commercial banks. Journal of Econometrics 196 (1), 68 – 82.

Galaix, X., 2012. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. Quartely Journal of Economics 127, 645–700.

Gao, J., 2007. Nonlinear Time Series: Semiparametric and Nonparametric Methods. Chapman & Hall/CRC, London.

Gao, J., Phillips, P., 2013a. Functional coefficient nonstationary regression. Cowles Foundation Discussion Papers 1911, Cowles Foundation for Research in Economics, Yale University.
URL `https://ideas.repec.org/p/cwl/cwldpp/1911.html`

Gao, J., Phillips, P., 2013b. Semiparametric estimation in triangular system equations with nonstationarity. Journal of Econometrics 176 (1), 59–79.

Granger, C. W. J., 1996. Can we improve the perceived quality of economic forecasts? Journal of Applied Econometrics 11 (5), 455–473.

Gu, S., Kelly, B., Xiu, D., 2020a. Autoencoder asset pricing models. Journal of Econometrics, forthcoming.

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

Hansen, P. R., Lunde, A., Nason, J. M., 2011. The model confidence set. Econometrica 79, 453–497.

Hastie, T., Tibshirani, R., 1993. Varying-coefficient models. Journal of the Royal Statistical Society. Series B (Methodological) 55, 757–796.

Henderson, D. J., Kumbhakar, S. C., Li, Q., Parmeter, C. F., 2015. Smooth coefficient estimation of a seemingly unrelated regression. Journal of Econometrics 189, 148–162.

Henkel, S. J., Martin, J. S., Nardari, F., 2011. Time-varying short-horizon predictability. Journal of Financial Economics 99, 560–580.

Johannes, M., Korteweg, A., Polson, N., 2014. Sequential learning, predictability, and optimal portfolio returns. The Journal of Finance 69 (2), 611–644.

Kelly, B., Jiang, H., 2014. Tail risk and asset prices. The Review of Financial Studies 27 (10), 2841–2871.

Kristensen, D., 2012. Non-parametric detection and estimation of structural change. Econometrics Journal 15, 420–461.

Lettau, M., Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. The Review of Financial Studies 21 (4), 1607–1652.

Li, J., Zinna, G., 2018. The variance risk premium: Components, term structures, and stock return predictability. Journal of Business & Economic Statistics 36 (3), 411–425.

McAleer, M., 2014. Oil shocks and the macroeconomy: The role of price variability. Econometrics 2 (3), 145–150.

Orbe, S., Ferreira, E., Rodriguez-Poo, J., 2005. Nonparametric estimation of time varying parameters under shape restrictions. Journal of Econometrics 126, 53–77.

Patton, A. J., 2011. Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics 160 (1), 246–256.

Patton, A. J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility full access. Review of Economics and Statistics 97 (3), 683–697.

Paye, B. S., Timmermann, A., 2006. Instability of return prediction models. Journal of Empirical Finance 13 (3), 274–315.

Pettenuzzo, D., Timmermann, A., 2011. Predictability of stock returns and asset allocation under structural breaks. Journal of Econometrics 164 (1), 60–78.

Phillips, P. C. B., Li, D., Gao, J., 2017. Estimating smooth structural change in cointegration models. Journal of Econometrics 196, 180–195.

Robinson, P., 1989. Nonparametric estimation of time-varying parameters. In: Hackl, P. (Ed.), Statistical Analysis and Forecasting of Economic Structural Change. Springer, Berlin.

Sun, Y., Cai, Z., Li, Q., 2013. Semiparametric functional coefficient models with integrated covariates. Econometric Theory 29, 659–672.

Sun, Y., Carroll, R., Li, D., 12 2009. Semiparametric estimation of fixed-effects panel data varying coefficient models. Advances in Econometrics 25, 101–129.

Viceira, L. M., 1996. Testing for structural change in the predictability of asset returns. Working Paper.

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

Wilcox, S. E., 2007. The adjusted earnings yield. Financial Analysts Journal 63, 54–68.

Xiao, Z., 2009. Functional-coefficient cointegration models. Journal of Econometrics 152, 81–92.