Journal of Econometrics 217 (2020) 312–334

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Dynamics of variance risk premia: A new model for disentangling the price of risk☆

Jeroen V.K. Rombouts [a,*], Lars Stentoft [b,c], Francesco Violante [d]

[a] ESSEC Business School, 1 Avenue Bernard Hirsch, Cergy Pontoise, 95021, France
[b] Department of Economics, University of Western Ontario, Social Science Centre, London, ON, N6A 5C2, Canada
[c] Department of Statistical and Actuarial Sciences, University of Western Ontario, Western Science Centre, London, ON, N6A 5B7, Canada
[d] CREST-ENSAE-ParisTech, 5, Av. Henry Le Chatelier, TSA 96642, 91764 Palaiseau CEDEX, France

ARTICLE INFO

Article history:
Available online 28 December 2019

JEL classification:
G12
G13
C58

Keywords:
Variance risk premium
Return predictability
Sentiment indicators

ABSTRACT

This paper formulates a new dynamic model for the variance risk premium based on a state space representation of a bivariate system for the observable ex-post realized variance and the ex-ante option implied variance expectation. A regime switching structure accommodates for periods of unusually high volatility, heterogeneous dynamics and changes in the dependence between the latent states. The model allows separating the continuous component of the variance risk premium from the impact of jumps on option implied variance expectations. Using options and high frequency returns for the S&P500 index, we explain what is generating return predictability by disentangling the part of the variance risk premium associated with normal sized price fluctuations from that associated with tail events. The latter component predicts to a significant extent, and asymmetrically with respect to their sign, future market return variations.

1. Introduction

Understanding the dynamics and economic drivers of volatility, or financial assets' price variability, is of great interest to academics and practitioners, see Bauwens et al. (2006), Francq and Zakoian (2010) and Bauwens et al. (2012) for thorough introductions on the topic. The recent literature focuses on the variance risk premium (VRP) that investors require for the well known fact that volatility is stochastic. The VRP is defined as the difference between the risk neutral and physical expectations of an asset's total return variation. While clear conceptually, the estimation of the VRP requires multiple sources of financial data as well as assumptions on the latent volatility processes, rendering its dynamic properties difficult to pinpoint. The vast majority of the literature estimates the VRP period-by-period using ex-post realizations ignoring its latent nature and temporal dependence. Such practice leads to a VRP that is too volatile and systematically becomes negative, rather than largely positive, in periods of market distress.

☆ We would like to thank Torben Andersen, Tim Bollerslev, Nicola Fusari, Elise Gourier, Kris Jacobs, Peter Phillips, Roberto Reno, Paul Schneider, Romeo Tedongap, Jules Tinang, Viktor Todorov, Fabio Trojani, Andrea Vedolin, Jun Yu, Hao Zhou. Jeroen Rombouts acknowledges support from Institut Louis Bachelier and the French National Research Agency (ANR), under grant ANR-17-CE26-0001 (project BREAKRISK). Lars Stentoft acknowledges support from CREATES, funded by the Danish National Research Foundation (DNRF78). Francesco Violante acknowledges support by the grant of the French National Research Agency (ANR), Investissements d'Avenir (LabEx Ecodec/ANR-11-LABX-0047), and of the Danish Council for Independent Research (1329-00011A).

* Corresponding author.
E-mail addresses: rombouts@essec.edu (J.V.K. Rombouts), lars.stentoft@uwo.ca (L. Stentoft), francesco.violante@ensae.fr (F. Violante).

https://doi.org/10.1016/j.jeconom.2019.12.006
0304-4076/

This paper formulates a dynamic model for the latent VRP in a joint specification for the variance and its option implied risk neutral expectation. We separate the component of the VRP associated with normal market activity from the contribution of unusual and extreme episodes of market volatility. The normal component of the VRP is characterized by level and variance shifts, strong persistence, is systematically positive and is higher in periods of market turmoil. By exploiting the temporal causality between realization of shocks on the spot market and expectations in the option market, we gauge the investors' reaction to extreme market volatility events and build risk factors representing their confidence towards future market states. The model, flexible yet easy to estimate, delivers precise and realistic estimates of the market price of volatility risk that together with the confidence factors allow to explain the puzzling return predictability found in the existing literature.

Carr and Wu (2009) proxy the VRP as the ex-post payoff of a variance swap contract rather than a premium that investors require ex-ante. To be in line with the proper definition of a risk premium, Bollerslev et al. (2009) compute the VRP from one period to the next as the difference between the implied variance computed from options and the expected realized variance computed with high frequency historical returns and filtered with a particular choice of dynamic model. While the above approaches are relatively simple to implement, their drawback stands in a resulting VRP time series extremely noisy and violating the positivity constraint too often to be a genuine risk premia.

We construct a joint dynamic model for the historical variance, its physical and its risk neutral expectations. We leave the dynamics of the underlying price process unspecified, allowing the direct use of realized and implied variances computed respectively from 5-min returns and option data on the underlying index, see also Bollerslev et al. (2011), Wu (2011) and Gruber et al. (2015). An alternative strategy is the specification of a full parametric model, as introduced for example in Andersen et al. (2015c), Ait-Sahalia et al. (2018), and Bardgett et al. (2019). However, in these models, the tight parameterization often constrains the dynamics of the VRP to mirror that of the variance itself.

Our model can be written in state-space form bringing the following advantages. First, it allows to account for possible errors in the measurement of the quadratic variation and option implied risk neutral variance expectations. Second, it allows to embed the expectations under the physical measure directly into the model. This avoids relying on a multistep estimation approach which is inefficient because subject to the compounding of estimation errors. Third, the model remains simple and its parameters can be estimated by maximum likelihood using the Kalman–Hamilton filter, see for example Kim (1994).

Identification of proper risk premia has important economic applications. Bollerslev et al. (2009), shows that market index returns become more predictable when including the VRP as a regressor in a future return model. Their result relies on random walk dynamics for the volatility. The random walk assumption, though, is at odds with the recent literature on volatility modelling, see e.g. Engle and Gallo (2006), Corsi (2009), Bauwens et al. (2013), Wang et al. (2013) and Augustyniak et al. (2019). Assuming more realistic dynamics, Drechsler and Yaron (2011) and Bekaert and Hoerova (2014) reveal a puzzle showing that predictions made by more sophisticated models fall short in terms of offering predictability. Since the latter models generally produce smoother VRP time series than the random walk implied VRP, which is completely exposed to extreme variance realizations, this suggests that most of the return predictability is due to tail effects, see Bollerslev and Todorov (2011) and Bollerslev et al. (2015).

Our joint model for the realized variance, the implied variance and the VRP reconciles the conflicting results found in the literature. The model relies on regimes that switch during events of unusual high volatility. This allows to explicitly quantify agents' fear, uncertainty deriving from heteroskedasticity, changes in the dynamics and dependence between the latent states. We identify the occurrence and measure the intensity of extreme variance shocks, and exploiting the causal effect between their realization and next period variance expectations, we define sentiment indicators based on agents' reaction to such shocks. We show that a substantial part of the return predictability ascribed in the literature to the variance risk premium stems from the part of the premium related to how agents gauge extreme variance events, their prediction and compensation. The documented predictive power is distinct from that of alternative well known financial predictors.

We estimate our model on data for the S&P 500 index. Our specification is strongly supported by the data, the filtered variance predictions and the risk neutral variance expectations match the level and the dynamics of their observable counterparts along the entire sample. The filtered variance risk premium satisfies naturally the positivity constraint, is highly persistent and exhibits changes in level and variability. We show that there is no substantial return predictability from the smooth component of the variance risk premium. In fact, the predictability attains a maximum value of 1.3% at the five month horizon. Including our sentiment indicators, the R^2s in the predictive regressions increase dramatically at every horizon and surpass 10% at the four month horizon. This predictive power is distinct from that contained in several well-known economic predictors, such as price–earning ratio, dividend yield, term spread, consumption–wealth ratio and output gap among others. Extending our empirical application to seven other indices, we find that the filtered VRP exhibits similar characteristics across the different markets. Return predictability due to the smooth component emerges in some cases and in combination with the sentiment indicators often dominates that of benchmark VRP measures for all markets.

The rest of the paper is organized as follows. Section 2 introduces the modelling framework. Section 3 reports results for the S&P 500 index and illustrates the dynamic properties of the VRP. Section 4 addresses the return predictability puzzle. Section 5 provides international evidence. Section 6 concludes.

2. A dynamic model for variance risk premia

2.1. Theoretical framework

We denote by S_t the spot price at time t of an asset defined on a probability space $(\Omega, \mathcal{F}, m)$, $m = P, Q$, with continuous time dynamics described by the stochastic differential equation

$$dS_t = S_{t-}(r_t^f + \lambda_t 1_{\{m=P\}})dt + S_{t-}\sigma_t dW_t^m + \int_{\mathbb{R}} S_{t-}(e^x - 1)\left(\mu(dt, dx) - \nu_t^m(x)dtdx\right), \tag{1}$$

where P and Q represent, respectively, the physical and risk neutral measures, r_t^f is the instantaneous risk-free rate, λ_t is the equity risk premium, σ_t is the instantaneous volatility, W_t^m is a standard Brownian motion, $\left(\mu(dt, dx) - \nu_t^m(x)dtdx\right)$ is a compensated counting process with $\mu(dt, dx)$ a measure which takes nonzero values when jumps occur, and $\nu_t^m(x)$ a jump compensator which gives the arrival rate of jumps of size x.

The quadratic variation of the continuously compounded return, i.e. $d\log S_t$, in the unit interval $[t, t+1]$ is defined as

$$QV_{t,t+1} = \int_t^{t+1} \sigma_s^2 ds + \int_t^{t+1}\int_{\mathbb{R}} x^2\mu(ds, dx), \tag{2}$$

where the first integral represents the portion of quadratic variation due to the diffuse price increments and is denoted by $\Sigma_{t,t+1}$. The second integral is associated with the jump component controlled by the counting measure μ and is denoted by $J_{t,t+1}$. Note that since μ is a martingale centred at $\nu_t^m(dx)dt$, then $E_t^m(\int_t^{t+1}\int_{\mathbb{R}} x^2\mu(ds, dx)) = E_t^m(\int_t^{t+1}\int_{\mathbb{R}} x^2\nu_s^m(dx)ds)$ for $m = P, Q$.

Carr and Wu (2009), decomposing $E_t^Q(QV_{t,t+1})$ as the sum of $E_t^P(QV_{t,t+1})$ and the conditional covariance between the normalized pricing kernel and the quadratic variation, define the VRP at time t as

$$\Pi_{t,t+1} = E_t^Q[QV_{t,t+1}] - E_t^P[QV_{t,t+1}]. \tag{3}$$

This is the definition also adopted by Bollerslev et al. (2009), Drechsler and Yaron (2011) and Ait-Sahalia et al. (2018) among others, and represents a special case of Kozhan et al. (2013). Indeed, (3) relates the VRP to the expected profit to the short side of a synthetic variance swap contract, which is entered into at time t and held until maturity at time $t+1$. The floating leg of a variance swap contract pays at $t+1$ an estimate of $QV_{t,t+1}$, computed as the annualized sum of daily squared log-returns over the time horizon $[t, t+1]$. The fixed leg pays an amount fixed at time t, defined as the variance swap rate. No arbitrage implies that the variance swap rate equals $E_t^Q(QV_{t,t+1})$ which is calculated from a portfolio of out of the money options, see Section 3.1. At maturity, the short position in a variance swap contract pays the difference between the realized variance and the variance swap rate which is then multiplied by a fixed notional amount to convert the payoff to dollar terms.

We do not impose sign constraints to the VRP in (3). This implies that we do not make prior assumptions on agents' risk preferences or exclude the possibility for them to change over time. However, the literature generally agrees on the positivity of the VRP. Drechsler and Yaron (2011), building on Bansal and Yaron (2004), associate positivity to risk averse agents and preference for early resolution of uncertainty. See Kozhan et al. (2013), and references therein, for empirical evidence.

Given that QV in (2) can be decomposed in a diffusive and jump part, the VRP in (3) can be written as

$$\begin{aligned}\Pi_{t,t+1} &= \left(E_t^Q[\Sigma_{t,t+1}] - E_t^P[\Sigma_{t,t+1}]\right) + \left(E_t^Q[J_{t,t+1}] - E_t^P[J_{t,t+1}]\right)\\ &= \Pi_{t,t+1}^C + \left(E_t^Q[J_{t,t+1}] - E_t^P[J_{t,t+1}]\right).\end{aligned} \tag{4}$$

The variable $\Pi_{t,t+1}^C$ represents the smooth component of the VRP, i.e. the part of the variance risk premium attributable to normal sized price fluctuations. The upperscript C, short for continuous, stresses the absence of discontinuities in the variance process. The term in brackets in (4) refers to the jump risk premium. As formulated in (1), the decomposition in (4) finds justification in the fact that, while the no arbitrage condition restricts the smooth component of the volatility process $\Sigma_{t-1,t}$ to be the same under the physical and the risk neutral measures, it puts no restrictions on jump dynamics.

2.2. Smooth VRP dynamics

In this section, we focus on the smooth component of the VRP

$$\Pi_{t,t+1}^C = E_t^Q[\Sigma_{t,t+1}] - E_t^P[\Sigma_{t,t+1}]. \tag{5}$$

We denote realized variance $RV_{t-1,t}$ an unbiased measurement of the latent $\Sigma_{t-1,t}$ in absence of discontinuities, such that

$$RV_{t-1,t} = \Sigma_{t-1,t} + \varepsilon_t^{RV}, \tag{6}$$

where ε_t^{RV} is a Gaussian measurement error due to discrete sampling of the returns, independently and identically distributed (i.i.d.) with zero mean and with $\mathrm{Var}(\varepsilon_t^{RV}) = \sigma_{RV}^2$. The Gaussianity assumption for the realized variance finds justification in Barndorff-Nielsen and Shephard (2002). The variance σ_{RV}^2 relates to the integrated quarticity which can be estimated using return data. Measuring directly the integrated quarticity has the advantage of allowing explicitly for time variation in the measurement error but at the cost of adding an extra latent state to the model.

Let us define the observable variance swap rate $SW_{t,t+1}$, a measure of $E_t^Q[\Sigma_{t,t+1}]$, such that $SW_{t,t+1} = E_t^Q[\Sigma_{t,t+1}]+\varepsilon_t^{SW}$. The measurement error ε_t^{SW} is also an i.i.d. zero mean Gaussian noise with $\mathrm{Var}(\varepsilon_t^{SW}) = \sigma_{SW}^2$. For example, SW can be estimated using the VIX, see CBOE (2015), such that the measurement error originates from the use of a limited set of available options at a given date. Then by using (5) we have that

$$SW_{t,t+1} = \Pi_{t,t+1}^C + E_t^P[\Sigma_{t,t+1}] + \varepsilon_t^{SW}, \tag{7}$$

where the stochastic component $\Pi_{t,t+1}^C$ acts as a wedge between the physical and risk neutral expectations of the variance. The model can be naturally interpreted as a structural model for risk neutral expectations. Given current and past information on the physical market, agents construct expectations about the risk and they price it. The decomposition of $SW_{t,t+1}$ in a stochastic latent factor and measurement error can also be justified using the arguments of Andersen et al. (2015a).

We now specify suitable dynamics for the latent states. First, we assume that $\Sigma_{t-1,t}$ is a stationary AR(1) process, i.e. $\Sigma_{t-1,t} = b + \phi^\Sigma(\Sigma_{t-2,t-1} - b) + \sigma_\Sigma v_t^\Sigma$ with $v_t^\Sigma \sim i.i.d.\, N(0, 1)$. An AR(1) specification is commonly used to model dynamics of the realized variance, see e.g. Dew-Becker et al. (2017), although a more general AR(p) with exogenous regressors can also be used, see e.g. Bekaert and Hoerova (2014). The autoregressive dynamics of $\Sigma_{t-1,t}$ allow us to structurally embed the variance expectation under the physical measure into the system by substituting $E_t^P[\Sigma_{t,t+1}] = b + \phi^\Sigma(\Sigma_{t-1,t} - b)$ into (7). Second, we specify the latent state $\Pi_{t,t+1}^C$ as

$$\Pi_{t,t+1}^C = F_{t,t+1} + d(\Sigma_{t-1,t} - b), \tag{8}$$

i.e., the sum of a dynamic factor $F_{t,t+1}$, idiosyncratic to $E_t^Q[\Sigma_{t,t+1}]$, and a term that captures the contemporaneous relationship between $\Pi_{t,t+1}^C$ and $\Sigma_{t-1,t}$. Central to the estimation of the VRP, the latent factor $F_{t,t+1}$ represents the idiosyncratic component of its smooth dynamics.

Note that the common latent factor $\Sigma_{t-1,t}$ enters the dynamics of the risk neutral expectation of the variance in two ways: (i) directly through the expectation under the physical measure and (ii) indirectly through the VRP allowing, as advocated by Bollerslev et al. (2011), for correlation between the variance premium and the variance itself. In our empirical application, we assume the idiosyncratic factor $F_{t,t+1}$ to be a stationary AR(1) process, i.e. $F_{t,t+1} = a+\phi^F(F_{t-1,t}-a)+\sigma_F v_t^F$ with $v_t^F \sim i.i.d.\, N(0, 1)$. As $E_t^Q[\Sigma_{t,t+1}]$ results from the contemporaneous aggregation of two AR(1) processes, it inherits a longer persistence than $\Sigma_{t-1,t}$, mimicking the behaviour documented in Andersen et al. (2007) and Andersen et al. (2015b).

Concentrating the parameters in the measurement system (6)–(7) and expressing the transition equations system in terms of the latent states $\Sigma_{t-1,t}$ and $F_{t,t+1}$, we can write the model in the following state-space formulation

$$\begin{bmatrix} RV_{t-1,t} \\ SW_{t,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ h^{(12)} & 1 \end{bmatrix} \begin{bmatrix} \Sigma_{t-1,t} \\ F^*_{t,t+1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^{RV} \\ \varepsilon_t^{SW} \end{bmatrix}, \tag{9}$$

with

$$\begin{bmatrix} \Sigma_{t-1,t} \\ F^*_{t,t+1} \end{bmatrix} = \begin{bmatrix} k^{(1)} \\ k^{(2)} \end{bmatrix} + \begin{bmatrix} \phi^\Sigma & 0 \\ 0 & \phi^F \end{bmatrix} \left(\begin{bmatrix} \Sigma_{t-2,t-1} \\ F^*_{t-1,t} \end{bmatrix} - \begin{bmatrix} k^{(1)} \\ k^{(2)} \end{bmatrix} \right) + \begin{bmatrix} \sigma_\Sigma & 0 \\ 0 & \sigma_F \end{bmatrix} \begin{bmatrix} v_t^\Sigma \\ v_t^F \end{bmatrix}, \tag{10}$$

where $F^*_{t,t+1} = F_{t,t+1} + b(1 - \phi^\Sigma - d) = F_{t,t+1} + k^{(2)} - a$, $h^{(12)} = \phi^\Sigma + d$, $k^{(1)} = b$, and $k^{(2)} = a + b(1 - \phi^\Sigma - d)$.

In general terms, this state-space formulation represents a latent common factor model with $\Sigma_{t-1,t}$ being the factor linking the variance dynamics and its risk neutral expectation. Note that the substitution of $E_t^P[\Sigma_{t,t+1}]$ into (7) is justified by the fact that for a linear state-space system the Kalman filter delivers the best linear predictions of the state vector, conditionally on the observations. Furthermore, in the particular case of Gaussian innovations in both the state and measurement equations, the one-step-ahead prediction of the state vector coincides with the conditional expectations. The model in (9)–(10) has the drawback that the assumption of normality may hold only approximately and allows for potentially negative realized variance and swap rates. In the next section, we introduce a regime switching model which reconciles the Gaussian linear state space form with the theory of realized variance. This is achieved by ensuring positivity of conditional expectations of the variance latent state for all t and a variance of the variance strictly decreasing with its level.

2.3. Heteroskedasticity and extreme variance events

To be realistic, the model above should be enriched by accounting for volatility level changes, heteroskedasticity and discontinuities in the variance dynamics, features empirically present in the data. We incorporate these features by means of a regime switching setup. As we want to accommodate separately for regime dependence both in the level and in the variance of the variance as well as extreme variance episodes, we work with a (N+1)-regime model with the following characteristics:

-*Regime* 0 *to* $N-1$: Increasing variance level, increasing variance of variance (both from very low to very high), decreasing measurement accuracy. The error vector $(\varepsilon_t^{RV}, \varepsilon_t^{SW})$ represents a pure measurement error. We advocate that market turmoil hampers the measurement accuracy of both the variance and its risk neutral expectation, i.e. the measurement error variance increases. We assume that higher levels of $\Sigma_{t-1,t}$ and $\Pi_{t,t+1}^C$ correspond to higher variance of the latent state errors. The change in the variance of the latent state errors of $\Sigma_{t-1,t}$ and $\Pi_{t,t+1}^C$ accounts for the heteroskedastic behaviour of the variance.

-*Regime* N: Extreme variance events (extreme variance level/extreme variance of variance/rare and short lasting episodes). The error vector $(\varepsilon_t^{RV}, \varepsilon_t^{SW})$ is a combination of measurement error and variance of the discontinuous component. The average size of the extremes is reflected in the marginal increase in the variance level from regime $N-1$. The measurement error variance of the observables and the Nth regime variance cannot be individually identified. Therefore, we assume that the latter component dominates and impose the restriction that the measurement error variance of the observables remains the same between regime $N-1$ and regime N. We define a regime indicator $S_t = \{0, 1, \dots, N\}$ and assume that it is a first order Markov Chain with transition probability matrix P with elements p_{ij} $i, j = 0, 1, \dots, N$.

We generalize the model accounting for extreme variance realizations by including the second term in (4) which relates to the jump risk premium. Modelling explicitly the jump risk premium requires the expectation of the jump process under the physical measure, which can be attained only by imposing tightly parameterized jump dynamics, see for example Bardgett et al. (2019). However, as stressed by Andersen et al. (2016), jumps are rare and heterogeneous in size leading to imprecise inference. To avoid this problem, we instead model discontinuities in the total variance process by considering them separately under the physical measure and under the risk neutral expectation. The measurement system in (6)–(7) can be written as

$$RV_{t-1,t} = \Sigma_{t-1,t} + J_{t-1,t} + \varepsilon_t^{RV}, \text{ and} \tag{11}$$

$$SW_{t,t+1} = \Pi_{t,t+1}^C + E_t^P[\Sigma_{t,t+1}] + E_t^Q[J_{t,t+1}] + \varepsilon_t^{SW}. \tag{12}$$

The system in (11)–(12) allows us to exploit the temporal causality that links the occurrence of extreme shocks in the spot market to the agents' reactions in the option market and thus infer agents' attitude towards risk and their sentiment about future market conditions. We elaborate more on this in Section 4.1.

The dynamics of the smooth component of the variance takes the form

$$\Sigma_{t-1,t} = b_{S_t} + \phi^{\Sigma}(\Sigma_{t-2,t-1} - b_{S_{t-1}}) + \sigma_{\Sigma,S_t} v_t^{\Sigma}, \tag{13}$$

with $v_t^{\Sigma} \sim i.i.d. N(0, 1)$. To model the level changes and heteroskedasticity in $\Sigma_{t-1,t}$, captured by movements between the regimes $S_t = 0, 1, \dots, N-1$, we impose the identifying restrictions $b_0 \le b_1 \le \dots \le b_{N-1} = b_N$ and $\sigma_{\Sigma,0} < \sigma_{\Sigma,1} < \dots < \sigma_{\Sigma,N-1} = \sigma_{\Sigma,N}$. Similarly, the smooth component of the variance risk premium $\Pi_{t,t+1}^C$ is defined as

$$\Pi_{t,t+1}^C = F_{t,t+1} + d_{S_t}(\Sigma_{t-1,t} - b_{S_t}), \tag{14}$$

with

$$F_{t,t+1} = a_{S_t} + \phi^F(F_{t-1,t} - a_{S_{t-1}}) + \sigma_{F,S_t} v_t^F, \tag{15}$$

where $v_t^F \sim i.i.d. N(0, 1)$. Since the smooth component moves between regimes 0 and $N-1$, we impose the identifying restrictions $a_0 \le a_1 \le \dots \le a_{N-1} = a_N$ and $\sigma_{F,0} < \sigma_{F,1} < \dots < \sigma_{F,N-1} = \sigma_{F,N}$. Compatible with the Gaussianity assumption, volatilities of innovations, σ_{Σ,S_t} and σ_{F,S_t}, small enough relative to the levels, b_{S_t} and a_{S_t}, ensure positivity of $\Sigma_{t-1,t}$ and $F_{t,t+1}$. Note that the extension of (6)–(7) to a regime dependent model defined in (11)–(12) naturally assumes that physical and risk neutral expectations are driven by the same regime indicator S_t and that agents make expectations given the current regime. This allows direct computation of the physical expectation of $\Sigma_{t-1,t}$ in (12).

The extreme variance episodes in $RV_{t-1,t}$ and $SW_{t,t+1}$ are captured by the Nth regime ($S_t = N$). Under the physical dynamics, we assume a basic random additive shock

$$J_{t-1,t} = \begin{cases} 0 & S_t = 0, 1, \dots, N-1 \\ j + \sigma_J u_t & S_t = N, \end{cases} \tag{16}$$

with $u_t \sim i.i.d. N(0, 1)$. The advantage of (16) is that it can fit in a parsimonious manner rare and short lasting extreme variance episodes. Similarly, for the expectation under the risk neutral measure, we assume

$$E_t^Q[J_{t,t+1}] = \begin{cases} 0 & S_t = 0, 1, \dots, N-1 \\ j^Q + \sigma_{J,Q} u_t^Q & S_t = N, \end{cases} \tag{17}$$

with $u^Q \sim i.i.d. N(0, 1)$. Note that this setup assumes that discontinuities occur simultaneously in the physical and risk neutral variances. That is, if a shock occurs in the spot market at a point in time (of size $J_{t-1,t}$), this generates a reaction in the risk neutral market, and prices in the latter will adjust accordingly (by an amount $E_t^Q[J_{t,t+1}]$).

Finally, the regime dependent error terms ε_t^{RV} and ε_t^{SW} are expressed as

$$\varepsilon_t^{RV} = \sigma_{RV,S_t} e_t^{RV}, \text{ and} \tag{18}$$
$$\varepsilon_t^{SW} = \sigma_{SW,S_t} e_t^{SW}, \tag{19}$$

with $e_t^{RV} \sim i.i.d. N(0, 1)$ and $e_t^{SW} \sim i.i.d. N(0, 1)$. As discussed at the beginning of this section, in order to identify σ_J and $\sigma_{J,Q}$, we impose the restriction $\sigma_{i,0} < \sigma_{i,1} < \cdots < \sigma_{i,N-1} = \sigma_{i,N}$, $i = RV, SW$. The restriction implies that in periods of instability the accuracy of the estimated proxies, $RV_{t-1,t}$ and $SW_{t,t+1}$, for their latent counterparts deteriorates. In addition, the restriction imposes the same variance of the measurement error in the highest level variance of variance regime ($S_t = N - 1$) and the extreme variance regime ($S_t = N$), implying that the quality of the measurements remains invariant to the occurrence of extreme shocks.

In sum, by accounting explicitly for the occurrence of extreme variance events and filtering out measurement error noise for the variables of interest, our approach allows for a more precise estimation of the VRP. Also, being relatively simple and flexible, the model can easily account for features advocated in the literature and typically revealed from the data such as uncertainty deriving from heteroskedasticity, dynamic characteristics such as level shifts, persistence as well as time varying correlation structure and dependence between the VRP, the risk neutral and the physical variance.

The model above can be written in state space form and can be estimated using straightforward filtering techniques. To do this, we define for measurement and transition equations of the physical variance $\Sigma_{t-1,t}^* = \Sigma_{t-1,t} + j_{S_t}$ and $\mathcal{E}_t^{RV} = \varepsilon_t^{RV} + \sigma_J u_t \mathbf{1}_{S_t=N}$, with $\text{Var}(\mathcal{E}_t^{RV}) = \sigma_{RV,S_t}^2 + \sigma_J^2 \mathbf{1}_{S_t=N} = \varsigma_{RV,S_t}^2$. Then we can rewrite (11) as

$$\begin{aligned} \Sigma_{t-1,t}^* &= (b_{S_t} + j_{S_t}) + \phi^{\Sigma}(\Sigma_{t-2,t-1}^* - b_{S_{t-1}} - j_{S_{t-1}}) + \sigma_{\Sigma,S_t} v_t^{\Sigma} \\ &= k_{S_t}^{(1)} + \phi^{\Sigma}(\Sigma_{t-2,t-1}^* - k_{S_{t-1}}^{(1)}) + \sigma_{\Sigma,S_t} v_t^{\Sigma} \end{aligned} \tag{20}$$
$$RV_{t-1,t} = \Sigma_{t-1,t}^* + \mathcal{E}_t^{RV} \tag{21}$$

where $k_{S_t}^{(1)}$ is a regime switching drift parameter.

We also re-parameterize and reduce the measurement and transition equations of the risk neutral expectation of the variance to functions of the latent states F and Σ. We can first compute $E_t^P[\Sigma_{t,t+1}]$ as

$$\begin{aligned} E_t^P[\Sigma_{t,t+1}] &= b_{S_t} + \phi^{\Sigma}(\Sigma_{t-1,t} - b_{S_t}) \\ &= b_{S_t} + \phi^{\Sigma}(\Sigma_{t-1,t}^* - b_{S_t} - j_{S_t}) \\ &= (b_{S_t}(1 - \phi^{\Sigma}) - \phi^{\Sigma} j_{S_t}) + \phi^{\Sigma} \Sigma_{t-1,t}^*. \end{aligned} \tag{22}$$

Second, the term in (14) which links the variance to the VRP can be written as

$$\begin{aligned} d_{S_t}(\Sigma_{t-1,t} - b_{S_t}) &= d_{S_t}(\Sigma_{t-1,t}^* - b_{S_t} - j_{S_t}) \\ &= -d_{S_t}(b_{S_t} + j_{S_t}) + d_{S_t} \Sigma_{t-1,t}^*. \end{aligned} \tag{23}$$

By defining $F_{t,t+1}^* = F_{t,t+1} + j_{S_t}^Q + b_{S_t}(1 - \phi^{\Sigma} - d_{S_t}) - j_{S_t}(\phi^{\Sigma} + d_{S_t}) = F_{t,t+1} + k_{S_t}^{(2)} - a_{S_t}$, we can write the transition equation for the latent state $F_{t,t+1}^*$ as

$$\begin{aligned} F_{t,t+1}^* &= k_{S_t}^{(2)} + \phi^F(F_{t-1,t} - a_{S_{t-1}}) + \sigma_{F,S_t} v_t^F \\ &= k_{S_t}^{(2)} + \phi^F(F_{t-1,t}^* - k_{S_{t-1}}^{(2)}) + \sigma_{F,S_t} v_t^F, \end{aligned} \tag{24}$$

where $k_{S_t}^{(2)}$ is a Markov-state dependent parameter defining the drift.

Following the same rationale for the physical variance, we define for the risk neutral variance $\mathcal{E}_t^{SW} = \varepsilon_t^{SW} + \sigma_{J,Q} u_t^Q \mathbf{1}_{S_t=N}$, with $\text{Var}(\mathcal{E}_t^{SW}) = \sigma_{SW,S_t}^2 + \sigma_{J,Q}^2 \mathbf{1}_{S_t=N} = \varsigma_{SW,S_t}^2$. The reduced form for the measurement equation of $SW_{t,t+1}$ in (12) is given by

$$\begin{aligned} SW_{t,t+1} &= F_{t,t+1}^* + (d_{S_t} + \phi^{\Sigma})\Sigma_{t-1,t}^* + \mathcal{E}_t^{SW} \\ &= F_{t,t+1}^* + h_{S_t}^{(12)} \Sigma_{t-1,t}^* + \mathcal{E}_t^{SW}. \end{aligned} \tag{25}$$

Putting the pieces together, the reduced form model in its state space form is thus

$$\begin{bmatrix} RV_{t-1,t} \\ SW_{t,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ h_{S_t}^{(12)} & 1 \end{bmatrix} \begin{bmatrix} \Sigma_{t-1,t}^* \\ F_{t,t+1}^* \end{bmatrix} + \begin{bmatrix} \mathcal{E}_t^{RV} \\ \mathcal{E}_t^{SW} \end{bmatrix}, \tag{26}$$

with

$$\begin{bmatrix} \Sigma_{t-1,t}^* \\ F_{t,t+1}^* \end{bmatrix} = \begin{bmatrix} k_{S_t}^{(1)} \\ k_{S_t}^{(2)} \end{bmatrix} + \begin{bmatrix} \phi^{\Sigma} & 0 \\ 0 & \phi^F \end{bmatrix} \left(\begin{bmatrix} \Sigma_{t-2,t-1}^* \\ F_{t-1,t}^* \end{bmatrix} - \begin{bmatrix} k_{S_t}^{(1)} \\ k_{S_t}^{(2)} \end{bmatrix} \right) + \begin{bmatrix} \sigma_{\Sigma,S_t} & 0 \\ 0 & \sigma_{F,S_t} \end{bmatrix} \begin{bmatrix} v_t^{\Sigma} \\ v_t^F \end{bmatrix}. \tag{27}$$

The model in its state-space form is estimated using Kim (1994) filter, which consists of a combination of the Kalman and Hamilton filters, in a quasi maximum likelihood framework, see Kim and Nelson (1999) for details. Starting values

Table 1
Sample statistics.

	$VIX^2_{t,t+1}$	$RV_{t-1,t}$	Payoff	$\Pi^{RW}_{t,t+1}$	$\Pi^{AR(1)}_{t,t+1}$	$\Pi^{AR(2)}_{t,t+1}$	$\Pi^{ARX1}_{t,t+1}$	$\Pi^{ARX2}_{t,t+1}$
Mean	38.796	20.179	17.584	17.584	17.602	17.568	17.568	17.602
Stddev	34.331	35.755	29.918	19.502	19.231	19.651	19.574	13.642
Skewn.	3.456	8.126	−5.873	−2.394	2.345	2.341	2.308	2.441
Kurt.	20.43	94.362	77.303	39.276	10.466	10.480	10.252	10.854
Min	9.048	1.865	−350.28	−180.68	−18.830	−25.428	−22.755	1.652
Max	298.9	479.58	124.45	116.52	125.910	125.07	125.820	91.840
$\rho(1)$	0.804	0.648	0.274	0.263	0.567	0.554	0.576	0.639

Notes: This table reports sample statistics for variables used to estimated the VRP. $\Pi^i_{t,t+1}$, $i = RW, AR(1), AR(2), ARX1$ and $ARX2$, is computed using (3) where the expectation under the physical measure is obtained using model i and the risk neutral expectation is approximated by the squared VIX. The first order sample autocorrelation coefficient is denoted by $\rho(1)$. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

are set using a combination of rank statistics and sample moments. A similar procedure has been used by Egloff et al. (2010) who estimate a VRP term structure model.

3. Model based VRPs

3.1. Nonparametric variance estimators

The quadratic variation between t and $t+\tau$ defined in (2) is estimated by the realized variance, see Andersen et al. (2001a) and Andersen et al. (2001b), as

$$RV_{t,t+\tau} = \sum_{i=1}^{I} \left(\ln(S_{t+i}/S_{t+i-1})\right)^2, \tag{28}$$

with I the number of observations between t and $t+\tau$.

The risk neutral expectation of the quadratic variation in (3) can be written as a portfolio of European call and put options with weights inversely proportional to the square of the options' strike prices, see Bakshi and Madan (2000), Britten-Jones and Neuberger (2000) and Jiang and Tian (2005), as

$$E_t^Q[QV_{t,t+\tau}] = \frac{2}{B_{t,\tau}} \left(\int_0^{\mathcal{F}_t} \frac{P_{t,\tau}(K)}{K^2} dK + \int_{\mathcal{F}_t}^{\infty} \frac{C_{t,\tau}(K)}{K^2} dK \right) + \mathcal{O}\left(\left(\frac{d\mathcal{F}_t}{\mathcal{F}_{t-}} \right)^3 \right), \tag{29}$$

where $B_{t,\tau}$ is the price of a time t zero-coupon bond maturing at time $t+\tau$, $\mathcal{F}_t$ is the forward price, and $P_{t,\tau}(K)$ and $C_{t,\tau}(K)$ are, respectively, the prices of put and call options with strike price K. The first term in (29) represents the risk neutral variance expectations in absence of price discontinuities. The second term is the compensator of the discontinuous component, see Carr and Wu (2009) for details. Eq. (29) is estimated as

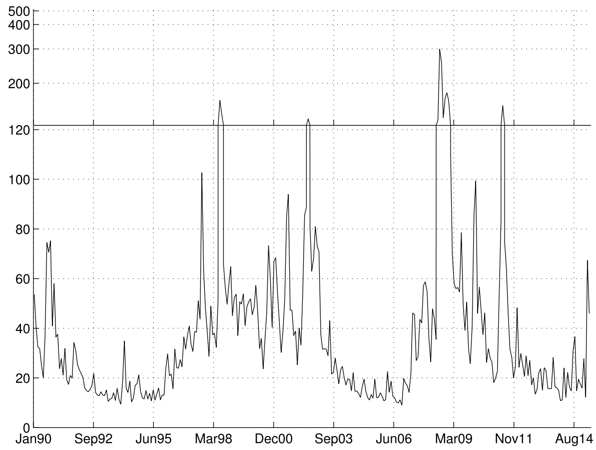
$$SW_{t,t+1} = 2e^{r_\tau} \left[\sum_{K_i \le K_0} \frac{\Delta K_i}{K_i^2} P_{t,\tau}(K_i) + \sum_{K_i \ge K_0} \frac{\Delta K_i}{K_i^2} C_{t,\tau}(K_i) \right] - \left[\frac{\mathcal{F}_t}{K_0} - 1 \right]^2, \tag{30}$$

where r_τ is the risk free rate and K_0 is the at the money strike price.

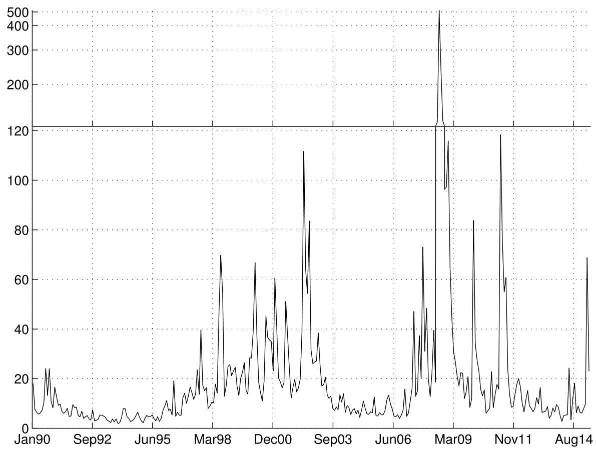
3.2. Data

S&P 500 data is sampled on non overlapping one month intervals with a monthly horizon for the computation of the physical and risk neutral variance expectations (i.e. $\tau = 1$). The data runs from January, 1990, to September, 2015. This setup is similar to Drechsler and Yaron (2011) and Bollerslev et al. (2014), among others. The variable $SW_{t,t+1}$ is the CBOE volatility index (VIX), see CBOE (2015). We denote as $VIX^2_{t,t+1}$ the last day of the month squared value of VIX, obtained from Datastream, divided by 12. The variable $RV_{t,t+\tau}$ is computed using 5 min returns along with the squared close-to-open overnight return and is obtained from the Oxford-Man Institute.

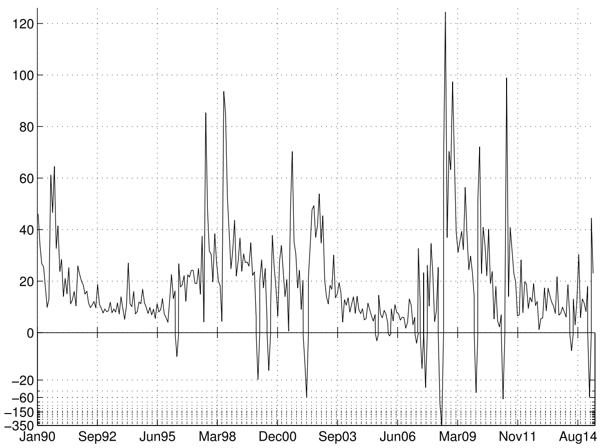
Fig. 1 plots $VIX^2_{t,t+1}$ and $RV_{t,t+\tau}$ in panel (a) and (b), respectively. The squared VIX generally shows smoother and more persistent dynamics than the realized variance, see also Table 1 which contains descriptive statistics. This is because, over the one-month observation period, shocks are incorporated into the medium-term expectations gradually. The realized variance shows the usual characteristic discontinuous behaviour emphasized by extreme and short lived variance events in periods of market instability, e.g. the dot-com bubble burst and the global financial crisis, due to the aggregation of shocks. In normal to volatile market conditions, the squared VIX lies above the realized variance. We see clear evidence of temporal causality between the occurrence of extreme shocks in the spot market and their inclusion in next period



(a) Monthly squared VIX



(b) Monthly realized variance



(c) Monthly variance swap payoff

Fig. 1. Risk neutral variance expectation, realized variance and variance swap payoff.

Table 2
Parameter estimates for our model.

State Var.	Parameter	Estimate	Standard error	t-statistic
$\Sigma_{t-1,t}$	b_0	8.184	1.165	7.023
	$b_1,\ b_2$	15.575	1.778	8.758
	ϕ^{Σ}	0.875	0.027	31.814
	$\sigma^2_{\Sigma,0}$	3.034	0.861	3.521
	$\sigma^2_{\Sigma,1},\ \sigma^2_{\Sigma,2}$	14.029	4.850	2.892
$\Pi^C_{t,t+1}$	a_0	12.164	1.785	6.813
	$a_1,\ a_2$	15.426	3.084	5.001
	ϕ^C	0.945	0.026	35.358
	d_0	0.785	0.190	4.133
	d_1	1.255	0.279	4.488
	d_2	1.131	0.283	3.990
	$\sigma^2_{C,0}$	0.770	0.719	1.071
	$\sigma^2_{C,1},\ \sigma^2_{C,2}$	4.135	3.632	1.138
$J_{t-1,1}$	j	43.392	5.126	8.464
	σ^2_J	502.954	163.625	3.073
$E^Q_t[J_{t,t+1}]$	j^Q	43.541	6.417	6.784
	$\sigma^2_{J,Q}$	750.955	233.115	3.221
ε^{RV}_t	$\sigma^2_{RV,0}$	1.901	0.627	3.027
	$\sigma^2_{RV,1},\ \sigma^2_{RV,2}$	60.448	13.244	4.564
ε^{SW}_t	$\sigma^2_{SW,0}$	10.968	1.983	5.529
	$\sigma^2_{SW,1},\ \sigma^2_{SW,2}$	48.197	16.859	2.858
P	p_{00}	0.862	0.031	26.999
	p_{11}	0.590	0.076	7.761
	p_{12}	0.128	0.039	3.234
	p_{22}	0.584	0.101	5.744
Log-likelihood				1506.321

Notes: This table provides maximum likelihood estimates for our model. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

option implied expectations. Fig. 1(c), displays the ex-post payoff from a synthetic variance swap contract defined as $VIX^2_{t,t+1} - RV_{t,t+1}$, see Carr and Wu (2009). Although the variance swap payoff does not measure the ex ante expectation and thus the variance premium, it is generally accepted as a valid approximation. The figure shows clustering of calm and volatile periods. Although the payoff is positive on average, on several occasions it is negative with extreme values around the beginning of the global financial crisis (up to −350.28), see Table 1.

3.3. Estimation results

In this section, we report results for a three regime specification, with regimes denoted respectively: $S_t = 0$, low variance level, i.e. calm markets; $S_t = 1$, high variance level, i.e. volatile markets; $S_t = 2$, extreme variance events. Given the monthly data frequency, we find that the three regime specification strikes the best compromise in terms of model fit, while remaining computationally feasible.

Table 2 reports parameter estimates for our model. We note that the change in the level and variance of the latent factors between the base regime and the high level/high variance of the variance regime is substantial. For example, the variance of the signals increase by a factor of about six for $F_{t,t+1}$ and five for $\Sigma_{t-1,t}$, respectively, and this accounts for the uncertainty deriving from heteroskedasticity. Moreover, the large marginal increase in the measurement error variance between the second and third regime, σ_J and σ_J^Q respectively, shows that the filter identifies the extreme variance events. Finally, the transition probability matrix reveals a highly persistent base regime, while the other two are relatively short lived, with expected durations equal to 7.24, 2.44 and 2.40 months, respectively. Note that the results reported in Table 2 refer to the model with a restricted transition probability matrix, i.e. $p_{02} = p_{20} = 0$, since the unrestricted model is rejected using a standard likelihood ratio test. Departures from regime 2 to regime 3 and back imply that the extreme variance regime constitutes systematically a departure from regime 2. Departures from regime 1 to regime 2 and back

Table 3
Extreme variance clusters.

Event	Date	$VIX^2_{t,t+1}$	$RV_{t-1,t}$	Length	avg $\Pi^C_{t,t+1}$	Sentiment
Asian Crisis	1997–10	102.61	39.58	1	24.33	Fear
Russian Crisis	1998–08	163.39	46.87	3	30.59	Fear
09/11	2001–09	84.96	51.16	2	22.05	Fear
Dot-com bubble burst	2002–07	85.49	111.67	4	24.98	Optim.
Global financial crisis	2008–09	129.30	128.26	7	52.05	Optim.
Flash crash	2010–05	85.71	83.79	2	21.43	Optim.
US debt downgrade	2011–08	83.32	118.36	4	25.40	Optim.

Notes: This table reports extreme variance clusters over the period January, 1990, to September, 2015. Length is the estimated length of the cluster measured in months. avg $\Pi^C_{t,t+1}$ means the average of $\Pi^C_{t,t+1}$ over the cluster.

Table 4
Times series regressions for $RV_{t,t+1}$.

Model name	Regressors		β_0	β_1	β_2	Adj.R^2
AR(1)	$RV_{t-1,t}$	–	7.122 (4.272)	0.647 (11.52)		41.9
AR(2)	$RV_{t-1,t}$	$RV_{t-2,t-1}$	7.402 (3.877)	0.670 (7.440)	−0.035 (−0.505)	42.0
ARX1	$RV_{t-1,t}$	$VIX^2_{t-1,t}$	7.684 (2.756)	0.660 (7.114)	−0.020 (−0.284)	42.0
ARX2	$RV_{t-1,t}$	$VIX^2_{t,t+1}$	0.111 (0.092)	0.381 (4.953)	0.328 (4.873)	45.5

Notes: This table report parameter estimates from time series models for $RV_{t,t+1}$. Newey–West t-statistics are in brackets and adjusted R^2s are expressed in percentages. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

accommodate the heteroskedastic behaviour of the variance. The steady state probabilities for the regimes are equal to 60.8%, 29.8% and 9.2%, respectively, showing that extreme variance episodes, while rare, occur around 10% of the time.

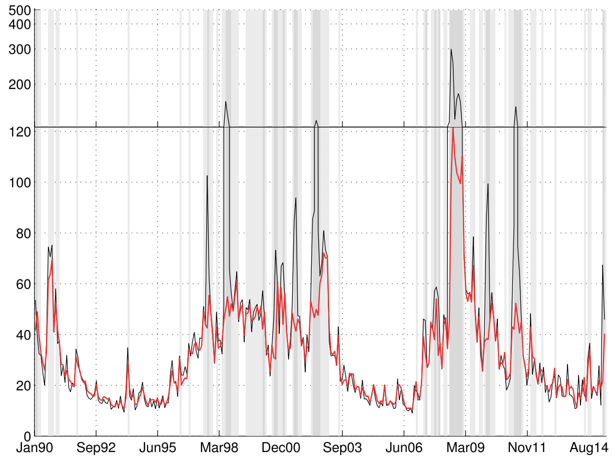
Table 2 also shows that with respect to the dynamics of the latent states, the stationary idiosyncratic component of $\Pi^C_{t,t+1}$, i.e., $F_{t,t+1}$, exhibits a strong persistence, with an autoregressive coefficient of 0.945, as well as the common factor, $\Sigma_{t-1,t}$, with an autoregressive coefficient of 0.875. For the latter, the effect of filtering out the extreme observations, which allows extracting the smooth component of the process and correctly estimating the mean reversion of the physical variance, is particularly striking as the equivalent coefficient of an AR(1) estimated directly on $RV_{t-1,t}$ equals 0.653, see Table 4. The reason for this is that in our model the observed extreme realizations get substantially down-weighted in the updating step of the latent states $\Pi^C_{t,t+1}$ and $\Sigma_{t-1,t}$. The usual negative bias on the autoregressive coefficient typically due to the presence of such extreme realizations appears clearly, and as discussed below, has serious implications on the estimation of the VRP. Also, the link between $\Sigma_{t-1,t}$ and $\Pi^C_{t,t+1}$ shows evidence of time variation especially when shifting between regime 1 and 2, as indicated by the difference $d_1 - d_0 = 0.470$.

Fig. 2 plots the filtered latent states, $\Sigma_{t-1,t}$ and $\Pi^C_{t,t+1}$, from our model as well as their combination into the latent risk neutral variance expectation. The shaded areas identify high vol-of-vol (light), i.e regime 2, and extreme variance event (dark), i.e. regime 3. For comparison, we also consider two alternative models for producing the variance expectations under the physical measure, the random walk and the AR(1). Our model fits well the data and correctly identifies the occurrence of extreme variance episodes and isolate the smooth component. The random walk model (black line), though providing reasonable predictions in periods of calm markets (indicated by the dark-grey shaded areas in the figure), suffers the problem of entirely projecting extreme realizations into the prediction, see Fig. 2(b). The AR(1) model (blue line) overestimates the variance in calm markets due to the well known bias in the autoregressive parameter in presence of extremes.

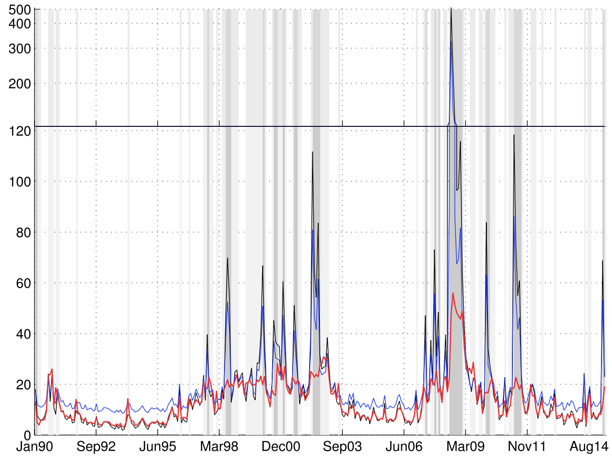
Finally, Fig. 2(c) (red line) shows that refining the estimation of the variance expectations under the physical dynamics and that of its risk neutral expectations, by explicitly accounting for heterogeneity and discontinuities, the smooth component of the variance risk premium is always positive ranging from 5.43 to 65.68 and with a persistence of 0.95. Furthermore, the variance risk premium increases sensibly, systematically and promptly in periods of market instability, for instance July and September 2002 (Stock market downturn of 2002), September 2008 (subprime crisis peak), May 2010 (Flash crash) or August 2011 (European sovereign debt crisis peak and US downgrading) among other episodes, see Table 3. The RW model implies a VRP which is sharply reduced and may even become negative in periods of market turmoil when the premium for bearing risk should be the highest and the AR(1) model implies a VRP which shrinks towards zero, eventually crossing the zero lower bound, in periods of calm markets.

The positivity of the VRP implied by our model, provides a solution to the puzzling feature pointed out by Bekaert and Hoerova (2014), characterizing the VRP series as estimated in Bollerslev et al. (2009), that when realized variances show extreme peaks, the VRP series often becomes very negative. Bekaert and Hoerova (2014) find this counterfactual as it is unlikely that during periods of intense market stress, there is a sudden increase in risk appetite.

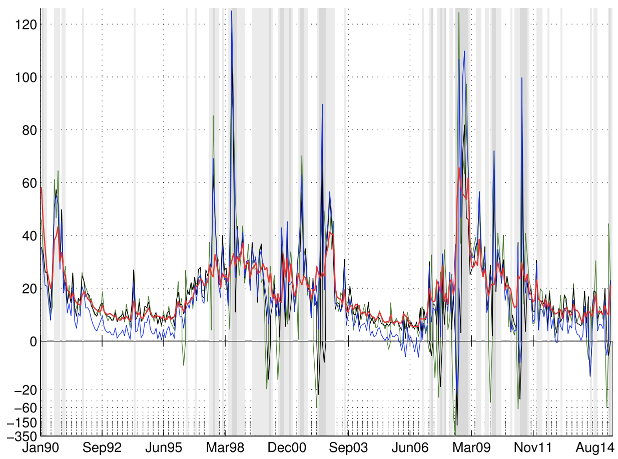
Table 3 provides information on extreme variance clusters identified by our model and triggered by several historical episodes. These clusters last from one to seven months with an average length of 3.3 months. We also identify 9 other isolated extreme variance events which, despite that they cannot be immediately linked to specific historical events, concentrate on the months around the dot-com bubble burst and the period immediately prior to the global financial



(a) Squared VIX (black) vs. filtered risk neutral expectations (red)



(b) RV predictions (red) vs. AR(1) (blue) and RW (black)



(c) Smooth component of the VRP, $\Pi^{C}_{t,t+1}$, (red) vs. $\Pi^{AR(1)}_{t,t+1}$ (blue), $\Pi^{RW}_{t,t+1}$ (black) and variance swap payoff (green)

Fig. 2. Risk-neutral variance expectations, realized variances and variance risk premia. The shaded areas identify the high vol-of-vol (light) and extreme variance event (dark) regimes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Return predictability with existing VRPs.

$\Pi^{i}_{t,t+1}$		Horizon (h)							
		1	2	3	4	5	6	9	12
$\Pi^{RW}_{t,t+1}$	$a_1(h)$	0.050	0.042	0.043	0.041	0.034	0.027	0.015	0.012
		(4.073)	(4.538)	(7.901)	(7.844)	(6.589)	(4.645)	(2.611)	(2.161)
	Adj.R^2	5.08	6.91	10.7	12.7	10.3	7.46	3.37	2.48
$\Pi^{AR(1)}_{t,t+1}$	$a_1(h)$	0.033	0.032	0.031	0.031	0.028	0.025	0.016	0.013
		(1.941)	(2.843)	(3.282)	(4.001)	(4.008)	(3.633)	(2.378)	(1.975)
	Adj.R^2	1.89	3.65	5.27	6.78	6.83	6.30	3.76	2.83
$\Pi^{ARX2}_{t,t+1}$	$a_1(h)$	0.033	0.031	0.031	0.031	0.028	0.025	0.016	0.013
		(1.984)	(2.834)	(3.272)	(4.037)	(4.048)	(3.638)	(2.370)	(1.980)
	Adj.R^2	1.97	3.68	5.34	6.92	6.96	6.37	3.78	2.87

Notes: This table reports parameter estimates for model (31). The constant is not reported to save space. Newey–West t-statistics in brackets. Adjusted R^2s in percentages. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

crisis. Section 4.1 further investigates these events, in particular the agents' reaction, and infer their sentiment towards future market states.

3.4. Benchmark models

A widely used approach to estimate the VRP assumes random walk (RW) dynamics for $QV_{t,t+1}$ such that $E^P_t[QV_{t,t+1}] = QV_{t-1,t}$. When $QV_{t-1,t}$ is estimated by the realized variance, $RV_{t-1,t}$, this results in $\Pi^{RW}_{t,t+1} = VIX^2_{t,t+1} - RV_{t-1,t}$. This way of computing the VRP, suggested by Bollerslev et al. (2009) and Bollerslev et al. (2014), only uses observable quantities at time t and is therefore directly useable in financial applications.

Drechsler and Yaron (2011) and Bekaert and Hoerova (2014) compute $E^P_t[QV_{t,t+1}]$ by projecting realized variance measures on a set of predictor variables including autoregressive dynamics and implied volatility. Mueller et al. (2015) follow a similar approach when studying bond VRP's. For sake of comparison, we consider the VRP implied from AR(p), $p = 1, 2$ models ($\Pi^{AR(1)}_{t,t+1}$ and $\Pi^{AR(2)}_{t,t+1}$) and two AR(1) specifications that include the ex-ante option implied variance expectation of the previous ($VIX^2_{t-1,t}$) or current ($VIX^2_{t,t+1}$) period as an additional regressor ($\Pi^{ARX1}_{t,t+1}$ and $\Pi^{ARX2}_{t,t+1}$).

Table 4 reports parameter estimates for the models fitted to realized variance. Because the additional parameters in AR(2) and ARX1 are insignificant, we consider as benchmarks $\Pi^{AR(1)}_{t,t+1}$ and $\Pi^{ARX2}_{t,t+1}$ together with $\Pi^{RW}_{t,t+1}$.

4. What is generating return predictability?

The VRP has been shown to be an important predictor of future aggregate stock market returns. Bollerslev et al. (2009) and Drechsler and Yaron (2011) demonstrate that, in addition to consumption risk, volatility risk plays an important role in generating returns, see also (Bonomo et al., 2015) for a framework starting from a daily frequency.

Table 5 provides slope coefficient estimates and (adjusted) R^2, the measure of return predictability, from the in-sample regressions

$$\frac{1}{h}\sum_{j=1}^{h} r_{t+j} = a_0(h) + a_1(h)\Pi^{i}_{t,t+1} + u_{t+h,t}, \tag{31}$$

where $\Pi^{i}_{t,t+1}$ denotes the VRP for $i = RW, AR(1)$ and $ARX2$, respectively, h denotes the forecast horizon and r_t denotes the monthly excess return for month t obtained from Datastream. We consider eight forecast horizons from one month up to one year. We find an inverse U-shaped pattern in the R^2 peaking at 4 months horizon with an R^2 of 12.7% when $\Pi^{RW}_{t,t+1}$ is used, and at the 5 month horizon with R^2 of 6.83% and 6.96% when $\Pi^{AR(1)}_{t,t+1}$ and $\Pi^{ARX2}_{t,t+1}$ are used respectively. These results suggest an interesting puzzle because more robust and realistic dynamics for the variance process fall short in terms of predictability compared to the simpler yet unrealistic random walk specification for the variance, see Bekaert and Hoerova (2014) for an extensive model comparison. In the next section, we define risk factors which, together with the continuous part of the VRP, explain where those differences stem from.

4.1. Return predictability and investors' sentiment

While our model yields the path of the smooth component of the variance risk premium $\Pi^{C}_{t,t+1}$, it also allows extracting the extreme variance episode indicator $I_{S_t=2}$. Using the notation in Kim and Nelson (1999), the indicator function equals

Table 6
Sentiment indicators associated with extreme market events.

Indicator	SI_j	Market performance	Response of Expect. to Jump	Size of the reaction	Construction of SI_i
Fear	ΔJ_t^+		$\Delta J_t > 0$	$\lvert\Delta J_t\rvert$	$\lvert\Delta J_t\rvert I_{\Delta J_t>0}$
Optimism	ΔJ_t^-		$\Delta J_t < 0$		$\lvert\Delta J_t\rvert I_{\Delta J_t<0}$
Fear&Bear	$\Delta J_{t,r_t^-}^+$	$r_t < 0$	$\Delta J_t > 0$	$\lvert\Delta J_t\rvert$	$\lvert\Delta J_t\rvert I_{\Delta J_t>0} I_{r_t<0}$
Optim.&Bear	$\Delta J_{t,r_t^-}^-$		$\Delta J_t < 0$		$\lvert\Delta J_t\rvert I_{\Delta J_t<0} I_{r_t<0}$
Fear&Bull	$\Delta J_{t,r_t^+}^+$	$r_t > 0$	$\Delta J_t > 0$	$\lvert\Delta J_t\rvert$	$\lvert\Delta J_t\rvert I_{\Delta J_t>0} I_{r_t>0}$
Optim.&Bull	$\Delta J_{t,r_t^+}^-$		$\Delta J_t < 0$		$\lvert\Delta J_t\rvert I_{\Delta J_t<0} I_{r_t>0}$

Notes: This table summarizes the sentiment indicators and their interaction with the state of the market.

Table 7
Descriptive statistics for sentiment indicators and their interaction with the market performance.

SI_j	Indicator	Mean	Std. deviation	Minimum	Maximum	N. obs.
ΔJ_t^+	Fear	36.025	21.947	2.071	86.785	12
ΔJ_t^-	Optimism	34.747	48.452	2.791	229.810	20
$\Delta J_{t,r_t^-}^+$	Fear&Bear	39.643	22.429	11.849	86.785	9
$\Delta J_{t,r_t^-}^-$	Optim.&Bear	41.128	56.292	8.526	229.810	14
$\Delta J_{t,r_t^+}^+$	Fear&Bull	25.172	20.036	2.071	37.809	3
$\Delta J_{t,r_t^+}^-$	Optim.&Bull	19.860	17.378	2.791	50.709	6

Notes: This table reports descriptive statistics for the attitude and sentiment indicators. The statistics are computed considering only the non-zero values for a given indicator. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

one if $\max\Big(P(S_t = j|\psi_t)\Big) = P(S_t = 2|\psi_t)$, where $P(S_t = j|\psi_t)$ is the posterior probability of state $j = 0, 1, 2$ and ψ_t the information set up to and including t. Using the filtered $\Sigma_{t-1,t}$ and $\Pi_{t,t+1}$ together with the indicator $I_{S_t=2}$, we can estimate, up to some measurement error, the extent of the extreme variance event $J_{t-1,t}$ and its risk neutral expectation $E_t^Q[J_{t,t+1}]$. To measure the type and size of agents' reactions to the occurrence of extreme variance events, we define the quantity

$$\Delta J_t = E_t^Q[J_{t,t+1}] - J_{t-1,t}. \tag{32}$$

The interpretation of ΔJ_t relates to the jump risk premium, i.e. $E_t^Q[J_{t,t+1}] - E_t^P[J_{t,t+1}]$ as discussed in (4). In fact, ΔJ_t can be written as the sum of the expected jump differential, i.e. $E_t^P[J_{t,t+1} - J_{t-1,t}]$, and the jump risk premium. Since the latter represents the price that agents attach to extreme events, ΔJ_t contains direct information about the response of agents to the current market environment.

The sign of ΔJ_t indicates if a shock in the market is incorporated into the risk neutral expectations more or less than proportionally. We use this quantity to assess agents' different attitude towards risk and thus their sentiment towards future market states. Agents which react more than proportionally to an extreme shock on the market ($\Delta J_t > 0$), i.e. agents which become more conservative when forming variance expectations, signal an increase in risk aversion as the observed shock is expected to produce a long lasting aftermath. We define this as fear and denote it ΔJ_t^+. Conversely, a reaction that is less than proportional to the observed shock ($\Delta J_t < 0$), signals a reduction in risk aversion. More liberal variance expectations, after an abnormally large variance shock occurs, indicate that the impact of the extreme event is expected to die out fast. We define this as optimism and denote it ΔJ_t^-. The construction of these risk factors finds justification in Bollerslev and Todorov (2011) which show that compensation for continuous price moves and for disaster risk explain to a different extent aggregate return variation.

We argue that not only the direction and the size of the agents' reaction to extreme shocks has a relevant effect on future market performances but also that the effect may be asymmetric, i.e. systematically related to the current market performance at the moment the shock occurs. This hypothesis relates to the concept of good vs. bad volatility developed in Patton and Sheppard (2015). A fear episode in conjunction with a bad market performance, i.e. a negative excess market return, is denoted $\Delta J_{t,r_t^-}^+$, while the mirroring case of fear associated to bullish markets is denoted $\Delta J_{t,r_t^+}^+$. Similarly, optimism in a state of bearish markets is denoted $\Delta J_{t,r_t^-}^-$, while optimism in conjunction with rallying markets is denoted $\Delta J_{t,r_t^+}^-$.

Table 6 summarizes the sentiment indicators and their interaction with the state of the market and Table 7 reports descriptive statistics. Out of 32 extreme events, we observe 12 fear reactions against 20 optimism episodes, with average

Table 8
Adjusted R^2s from predictive return regressions.

Regressors		Horizon							
Π^i	$SI_{j,t}$	1	2	3	4	5	6	9	12
$\Pi^C_{t,t+1}$		−0.30	−0.17	0.16	0.84	1.32	1.28	0.65	1.01
$\Pi^C_{t,t+1}$	ΔJ_t	6.23	6.25	8.31	9.98	8.16	5.43	2.39	1.91
$\Pi^C_{t,t+1}$	ΔJ^+_t ΔJ^-_t	7.08	7.68	8.97	10.33	8.87	6.49	3.57	2.44
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r^-_t}$ $\Delta J^-_{t,r^-_t}$ $\Delta J^+_{t,r^+_t}$ $\Delta J^-_{t,r^+_t}$	7.37	8.93	9.27	10.38	8.89	6.54	4.23	3.26
$\Pi^C_{t,t+1}$	ΔJ^+_t	4.41	5.53	5.19	5.47	5.90	5.19	3.32	2.29
$\Pi^C_{t,t+1}$	ΔJ^-_t	2.98	2.60	4.70	6.52	4.95	2.99	1.07	1.25
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r^-_t}$	3.01	4.07	3.92	4.50	5.22	4.84	3.56	2.75
$\Pi^C_{t,t+1}$	$\Delta J^-_{t,r^-_t}$	2.68	1.62	3.95	6.10	4.76	2.93	1.23	1.37
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r^+_t}$	1.27	1.35	1.48	1.79	1.92	1.60	0.66	1.09
$\Pi^C_{t,t+1}$	$\Delta J^-_{t,r^+_t}$	−0.11	1.80	0.89	1.08	1.42	1.31	0.85	1.17
$\Pi^{RW}_{t,t+1}$		5.08	6.91	10.7	12.7	10.3	7.46	3.37	2.48
$\Pi^{AR(1)}_{t,t+1}$		1.89	3.65	5.27	6.78	6.83	6.30	3.76	2.83
$\Pi^{ARX2}_{t,t+1}$		1.17	2.68	3.78	5.02	5.43	5.36	3.41	2.56

Notes: This table reports the adjusted R^2s in percentage from the predictive return regressions in (33) with VRP, sentiment indicators and their interaction with the state of the market as regressors. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

sizes of 36.025 and 34.747, respectively. When we associate the size of the reaction to the sign of the market performance, we observe that when returns are negative the average reaction is about twice as large and more than double in occurrence when compared to positive returns. Note that the number of occurrences of the fear indicator when associated to bear markets counts for less than 3% of the total sample size, while the episodes classified as Optimism&Bear represent about 5%. As expected, fear and optimism reactions in conjunction with positive shocks in the market, turn out to be much more rare (3 and 6 occurrences, respectively, out of 309 observations).

4.2. Return predictability from model based VRPs

We assess the predictive ability of the sentiment indicators for the equity premium. We argue that the predictability previously attributed to $\Pi^i_{t,t+1}$ stems to a large extent from the effect of rare and extreme variance events as well as the reaction of agents to those events. To test this hypothesis, we augment the predictive regression model as

$$\frac{1}{h}\sum_{j=1}^{h} r_{t+j} = a_0(h) + a_1(h)\Pi^C_{t,t+1} + \sum_{j=1}^{m} b_j(h) SI_{j,t} + u_{t+h,t} \tag{33}$$

where SI_j is a set of sentiment indicators. In order to isolate the contribution of the variance premium and the investors' sentiment indexes, we do not consider other explanatory variables here. The inclusion of other economic predictors is discussed in Section 4.3.

Tables 8 and 9 report respectively the adjusted R^2s and the estimates for the predictive return regression in (33). When we regress the future aggregate market excess return on $\Pi^C_{t,t+1}$, we find no substantial return predictability at any horizon. In fact, the predictability attains a maximum value equal to 1.3% at the five-month horizon. Although a persistent $\Pi^C_{t,t+1}$ variable, we do not find increasing predictability, as documented e.g. for the dividend yield by Boudoukh et al. (2008), over the horizons we consider. The size of R^2 is in line with Huang et al. (2015), when using the investor sentiment index results of Baker and Wurgler (2006), who size the predictability of the monthly aggregate market excess return in the order of 1.5%. Thus, the results of Table 8 suggest that $\Pi^C_{t,t+1}$ indeed measures the economic uncertainty under normal market conditions.

This finding suggests that much of the return predictability ascribed to the variance risk premium computed as $\Pi^{RW}_{t,t+1}$ is effectively coming from the part of the premium related to how agents gauge extreme variance events, their prediction and compensation, see also Bollerslev et al. (2015). In our framework, the latter effects are proxied by the variable ΔJ_t and when including the size of the reaction of agents to unusual variance events, the R^2 of the regression sharply increases at every horizon, reaching 10% at the 4 month horizon.

Separating ΔJ_t in its two components ΔJ^+_t and ΔJ^-_t, Table 8 shows that the extra degree of freedom allowing for an asymmetric effect on future returns of conservative (fear) versus liberal (optimism) expectations improves the

Table 9
Parameter estimates from return predictive regressions.

	One-month horizon				Three-month horizon			
$\Pi^C_{t,t+1}$	0.006	0.024	0.011	0.011	0.017	0.029	0.022	0.024
	(0.187)	(0.854)	(0.038)	(0.402)	(0.829)	(1.719)	(1.215)	(1.276)
ΔJ_t		0.064				0.043		
		(4.384)				(7.248)		
ΔJ^+_t			0.107				0.065	
			(4.854)				(5.707)	
ΔJ^-_t			−0.049				−0.035	
			(−4.312)				(−5.245)	
$\Delta J^+_{t,r^-_t}$				0.097				0.060
				(3.738)				(5.526)
$\Delta J^-_{t,r^-_t}$				−0.049				−0.034
				(−4.006)				(−5.845)
$\Delta J^+_{t,r^+_t}$				0.177				0.095
				(11.224)				(2.822)
$\Delta J^-_{t,r^+_t}$				−0.053				−0.063
				(−0.503)				(−1.610)
Adj.R^2	−0.30	6.23	7.09	7.37	0.16	8.31	8.97	9.27
	Six-month horizon				Twelve-month horizon			
$\Pi^C_{t,t+1}$	0.023	0.029	0.023	0.022	0.015	0.017	0.014	0.013
	(1.432)	(1.709)	(1.335)	(1.282)	(1.061)	(1.128)	(0.950)	(0.884)
ΔJ_t		0.023				0.008		
		(5.717)				(2.474)		
ΔJ^+_t			0.044				0.019	
			(4.665)				(2.721)	
ΔJ^-_t			−0.015				−0.004	
			(−3.745)				(−1.413)	
$\Delta J^+_{t,r^-_t}$				0.045				0.024
				(4.489)				(3.868)
$\Delta J^-_{t,r^-_t}$				−0.016				−0.005
				(−4.155)				(−1.719)
$\Delta J^+_{t,r^+_t}$				0.036				−0.012
				(1.258)				(−0.437)
$\Delta J^-_{t,r^+_t}$				−0.007				0.017
				(−0.370)				(1.272)
Adj.R^2	1.28	5.43	6.49	6.53	1.01	1.91	2.45	3.26

Notes: This table reports parameter estimates from the predictive return regressions in (33) with VRP, sentiment indicators and their interaction with the state of the market as regressors. Newey–West t-statistics are in brackets and adjusted R^2s are expressed in percentages. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

predictability at the short and long horizons. Moreover, Table 9 shows that the marginal impact of ΔJ^+_t is systematically at least double that of ΔJ^-_t and the two consistently carry opposite signs. The null hypothesis of no asymmetry is always rejected at standard significance levels. Therefore, agents effectively price the fear for market instability only when shocks are perceived as long lasting.

Interacting the variables ΔJ^+_t and ΔJ^-_t with the market performance indicator further increases the R^2s and this especially so at horizons at which return predictability is known to be difficult to capture, i.e. one to three month horizon. This shows that the asymmetry, with respect to the state of the market, in the way each sentiment indicator explains future aggregate market returns cannot be rejected. The inverse U-shaped relationship between horizon and R^2 is preserved but the curve that we obtain is somewhat flatter, i.e. starting and ending with higher R^2s. For example, the one month horizon R^2 is 7.37 against 6.23 of the regression including only $\Pi^C_{t,t+1}$ and ΔJ_t. Similarly, the one year horizon R^2s are respectively 3.26 and 1.91.

With respect to the state of the market, as expected, we find that fear in conjunction with negative returns impacts the future market performance over longer horizons. In fact, the coefficients associated with $\Delta J^+_{t,r^-_t}$ (Fear&Bear) are significant at all horizons exhibiting the slowly decaying behaviour discussed above. Contrary, the effect of fear associated to bullish markets dies out much faster as shown by the insignificant coefficients starting from the six month horizon. For the case of optimistic variance expectations, $\Delta J^{-}_{t,r_t^-}$ (Optimism&Bear) contributes to explain market returns at all horizons, while optimism in bullish markets contributes only marginally. This result supports the hypothesis of an asymmetric response of the excess aggregate market returns to the sign of the shock in the aggregate price.

The natural question to ask is what justifies the sharp difference in return predictability of the existing models in the literature typically used to estimate the VRP. Using our sentiment indicators, we can pin down which component in the VRP based on the random walk and on autoregressive type models effectively explains future return variation. To be precise, since the random walk implied VRP is completely exposed to extreme variance realizations, both in the realized variance and its risk neutral expectations, our results suggest that the degree of predictability of $\Pi^{RW}_{t,t+1}$ has to be dominated by that of ΔJ^{+}_{t} and ΔJ^{-}_{t}. To test this hypothesis, we regress the future excess returns on the difference between $\Pi^{RW}_{t,t+1}$ and the sum of the fear and optimism indicators. We find that the resulting R^2s drop at all horizons, e.g. from 12.7 to 3.2 at the four-month horizon. Similarly, we argue that return predictability associated with $\Pi^{AR(1)}_{t,t+1}$, which is instead based on smoother variance expectations, is mostly exposed to extreme realizations of the risk neutral expectations, thus being dominated by ΔJ^{+}_{t}. Subtracting the fear indicator (ΔJ^{+}_{t}) from AR(1) implied $\Pi^{AR(1)}_{t,t+1}$, the return predictability in fact aligns to the same levels observed for the smooth component $\Pi^{C}_{t,t+1}$.

4.3. Predictability with other economic predictors

We compare the predictive power of our sentiment indicators with several well known macro-finance predictors of market returns. This allows to examine whether our sentiment indicators' predictive power is driven by omitted economic variables related to business cycle fundamentals or changes in investor risk aversion. To test this, we augment the predictive regression to take the following form

$$\frac{1}{h}\sum_{j=1}^{h} r_{t+j} = a_0(h) + \sum_{j=1}^{m} a_j(h) SI_{j,t} + \sum_{i=1}^{k} b_i(h) X_{i,t} + u_{t+h,t}. \tag{34}$$

The set of regressors, $X_{i,t}$, we consider are:

- *Basic financial variables*: Dividend–payout ratio (DPO): difference between the log of dividends and log of earnings on the S&P 500 index; Price–earnings ratio (PE): difference between the log of earnings and log of prices on the S&P 500 index, where earnings are measured using a one-year moving sum; Dividend yield (DY): difference between the log of dividends and log of lagged prices on the S&P 500 index. All variables are downloaded from Robert Shiller's website.
- *Yield related variables*: Term spread (TMS): difference between the long-term yield and Treasury bill rate (10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity); Default yield spread (DFY): difference between Moody's Seasoned BAA- and AAA-rated corporate bond yields; Default return spread (DFR): difference between long-term corporate bond and long-term government bond returns (Moody's Seasoned AAA Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity). All variables are downloaded from the St. Louis Federal Reserve website.
- *Traditional risk factors*: Small minus big (SMB) and high minus low (HML) are constructed using value-weighted portfolios of NYSE, AMEX and NASDAQ stocks formed on size and book-to-market. SMB is the average return on the small portfolios minus the average return on the big portfolios, HML is the average return on the value portfolios minus the average return on the growth portfolios. MOM is the Carhart (1997) momentum factor. All variables are downloaded from Kenneth R. French's website.
- *Macro-economic variables*: consumption–wealth ratio (CAY): residual of regressing consumption on asset wealth and labour income from Lettau and Ludvigson (2001) and downloaded from Lettau's website. Re-spanned on monthly frequency as in Bollerslev et al. (2009); Inflation (CPI): calculated from the consumer price index for all urban consumers as used in Giglio et al. (2016); Output gap (OG): deviation of the logarithm of total industrial production from a trend that includes both a linear component and a quadratic component, see Cooper and Priestley (2009). We also include a cubic trend to capture better the recent economic downturn. The variables CPI and total industrial production are downloaded from the St. Louis Federal Reserve website.
- *Implied volatility*: The squared CBOE implied volatility index (VIX) downloaded from Datastream.

The left panels of Tables 10 and 11 consider the predictive regressions with each single economic variable, for respectively the one/three month and six/twelve month horizons. We report the slope coefficients of the regressions along with the adjusted R^2s. Out of the 14 economic predictors, only DY exhibit significant predictive power for the market aggregate future return at the 10% or better significance levels for all the considered horizons. The R^2s increase from 0.49% to 12.26% when the horizon increases from one to twelve months. The variable OG has significant predictive power from the three month horizon onwards with R^2s from 3.06% to 17.17% and finally MOM is significant from the sixth month horizon with largest R^2 of 1.33% at the twelve month horizon.

Table 10
Predictive return regressions with economic predictors.

One-month horizon								
	$\hat{r}_t^{(h)} = \hat{a}_0 + \hat{b}_1 X_{1,t}$		$\hat{r}_t^{(h)} = \hat{a}_0 + \sum_{j=1}^{4} \hat{a}_j SI_{j,t} + \hat{b}_1 X_{1,t}$					
$X_{1,t}$	$\hat{b}_1$	Adj.R^2	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{b}_1$	Adj.R^2
Π^C	6e−3	−0.30	0.10*	−0.05*	0.18*	−0.05	0.01	7.37
DPO	−3e−3	−0.33	0.10*	−0.05*	0.18*	−0.05	0.06	7.30
PE	−0.73	0.09	0.10*	−0.05*	0.19*	−0.03	−0.93	7.96
DY	1.31*	0.49	0.10*	−0.05*	0.19*	−0.05	1.69*	8.64
TMS	−0.08	−0.27	0.10*	−0.05*	0.18*	−0.05	0.02	7.30
DFR	−0.42	−0.10	0.10*	−0.04*	0.19*	−0.03	−0.44	7.52
DFY	0.69	0.12	0.10*	−0.04*	0.18*	−0.03	0.31	7.37
SMB	0.06	−0.12	0.10*	−0.05*	0.18*	−0.05	0.05	7.48
HML	−0.09	0.15	0.10*	−0.05*	0.17*	−0.05	−0.09	7.76
MOM	−0.05	−0.05	0.10*	−0.05*	0.18*	−0.05	−0.04	7.48
LTR	−0.10	0.01	0.10*	−0.05*	0.18*	−0.05	−0.07	7.49
CAY	−10.65	−0.12	0.10*	−0.05*	0.18*	−0.04	−12.95	7.60
CPI	33.90	−0.25	0.10*	−0.05*	0.18*	−0.05	−8.49	7.30
OG	−8.99	0.55	0.10*	−0.05*	0.18*	−0.05	−8.77*	8.13
VIX	0.00	−0.32	0.11*	−0.04*	0.19*	−0.04	−0.01	7.37
Three-month horizon								
	$\hat{r}_t^{(h)} = \hat{a}_0 + \hat{b}_1 X_{1,t}$		$\hat{r}_t^{(h)} = \hat{a}_0 + \sum_{j=1}^{4} \hat{a}_j SI_{j,t} + \hat{b}_1 X_{1,t}$					
$X_{1,t}$	$\hat{b}_1$	Adj.R^2	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{b}_1$	Adj.R^2
Π^C	0.02	0.16	0.06*	−0.03*	0.10*	−0.06	0.02	9.27
DPO	0.23	−0.18	0.07*	−0.03*	0.10*	−0.05	0.30	8.67
PE	−0.47	0.16	0.07*	−0.03*	0.11*	−0.04	−0.57	9.12
DY	1.39*	2.26	0.07*	−0.03*	0.11*	−0.05	1.64*	12.00
TMS	−0.04	−0.29	0.07*	−0.03*	0.10*	−0.05	0.02	8.43
DFR	−0.35	0.11	0.07*	−0.03*	0.11*	−0.04	−0.35	8.81
DFY	0.40	0.10	0.07*	−0.03*	0.10*	−0.04	0.11	8.44
SMB	−0.06	0.24	0.06*	−0.03*	0.11*	−0.05	−0.06	9.03
HML	−0.06	0.29	0.07*	−0.03*	0.09*	−0.05*	−0.07	9.10
MOM	−0.03*	0.96	0.07*	−0.03*	0.09*	−0.06*	−0.06	9.58
LTR	−0.15*	1.90	0.06*	−0.03*	0.10*	−0.05	−0.14*	10.28
CAY	−0.39	−0.22	0.07*	−0.03*	0.10*	−0.05	−1.46	8.43
CPI	45.80	0.05	0.07*	−0.03*	0.10*	−0.04	17.62	8.47
OG	−11.37*	3.06	0.06*	−0.03*	0.10*	−0.05	−11.32*	12.29
VIX	0.00	−0.24	0.06*	−0.03*	0.09*	−0.05	0.00	8.50

Notes: This table reports parameter estimates from the predictive return regressions in (34) with economic predictors and sentiment indicators as regressors. Newey–West t-statistics are in brackets and adjusted R^2 are expressed in percentages. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

The right panels of Tables 10 and 11 report multivariate regressions to test for incremental predictive power of the sentiment indicators and their interaction with the state of the market, $SI_{j,t}$. The results show that the estimates of the slope coefficients associated with the sentiment indicators are indeed in line, in terms of sign and size, with those reported in Table 9. More precisely, the coefficient a_1 associated with Fear&Bear remains statistically significant at all horizons when the regression is augmented by the economic predictors. The coefficients of the indicators Optimism&Bear and Fear&Bull, i.e. a_2 and a_3, are significant at the one and three month horizons but less so at the six and twelve month horizons. The coefficient a_4 associated with Optimism&Bull is significant several times at the twelve month horizon.

With respect to the magnitude of the adjusted R^2s of these augmented regressions, they are substantially larger than those based on the economic predictors alone. For instance, at the one month horizon, the R^2s increase from about zero percent on average to at least 7%. Regressions combining the sentiment indicators with economic predictors found individually significant seems to increase predictability in an additive way. For example at the six month horizon, DY alone has an R^2 of 4.83% compared to an R^2 of 11.26% when the sentiment indicators are included in the model. This confirms that the latter indicators bring in genuinely different explanatory power. In sum, these results demonstrate that the sentiment indicators contain considerable complementary forecasting information beyond what is contained in the usual economic predictors.

5. International evidence

To study the properties of the VRP and its return predictability beyond the S&P 500 index, we estimate our model on seven other market indices. Apart from the Dow Jones industrial average (DJIA), we consider the following foreign market indices: the STOXX Europe 50 index covering 18 European countries, the CAC 40 from France, the AEX index from the Netherlands, the FTSE 100 from the United Kingdom, the SMI from Switzerland and the NIKKEI 225 from Japan.

Table 11
Predictive return regressions with economic predictors.

Six-month horizon								
	$\hat{r}_t^{(h)} = \hat{a}_0 + \hat{b}_1 X_{1,t}$		$\hat{r}_t^{(h)} = \hat{a}_0 + \sum_{j=1}^4 \hat{a}_j SI_{j,t} + \hat{b}_1 X_{1,t}$					
$X_{1,t}$	$\hat{b}_1$	Adj.R^2	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{b}_1$	Adj.R^2
Π^C	0.02	1.28	0.05*	−0.02*	0.04	−0.01	0.02	6.53
DPO	0.37	0.38	0.05*	−0.01*	0.04	−0.00	0.35	5.80
PE	−0.36	0.19	0.05*	−0.01*	0.05*	0.02	−0.49	6.11
DY	1.46*	4.83	0.05*	−0.01*	0.05*	−0.00	1.60*	11.26
TMS	0.03	−0.30	0.05*	−0.01*	0.04	0.01	0.07	5.35
DFR	−0.22	−0.03	0.05*	−0.01*	0.05*	0.02	−0.33	5.79
DFY	−0.02	−0.33	0.05*	−0.01	0.04	0.01	−0.03	5.19
SMB	−0.05	0.61	0.05*	−0.01*	0.05	0.01	−0.05	5.97
HML	−0.08*	1.27	0.05*	−0.01*	0.03	−0.00	−0.08*	6.79
MOM	−0.04*	0.92	0.05*	−0.01*	0.04	−0.00	−0.04*	6.32
LTR	−0.09	1.02	0.05*	−0.01*	0.04	0.01	−0.08	6.23
CAY	7.18	1.15	0.05*	−0.01*	0.04	0.01	5.81	5.47
CPI	−45.38	0.34	0.05*	−0.01*	0.04	−0.00	−56.58	6.12
OG	−12.96*	8.80	0.05*	−0.01*	0.04	−0.00	−12.70*	13.93
VIX	0.00	1.06	0.04*	−0.02*	0.03	−0.01	0.01	5.93
Twelve-month horizon								
	$\hat{r}_t^{(h)} = \hat{a}_0 + \hat{b}_1 X_{1,t}$		$\hat{r}_t^{(h)} = \hat{a}_0 + \sum_{j=1}^4 \hat{a}_j SI_{j,t} + \hat{b}_1 X_{1,t}$					
$X_{1,t}$	$\hat{b}_1$	Adj.R^2	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{b}_1$	Adj.R^2
Π^C	0.01	1.01	0.02*	−0.01*	0.01	−0.02	0.01	3.26
DPO	0.35	0.91	0.03*	−0.00	−0.01	0.02	0.32	3.38
PE	−0.51	1.58	0.03*	−0.00	0.00	0.03	−0.60	5.00
DY	1.66*	12.26	0.03*	−0.00	0.01	0.02	1.70*	15.47
TMS	0.17	1.75	0.03*	−0.00	−0.01	0.03*	0.18	4.82
DFR	−0.21	0.19	0.03*	−0.00	0.00	0.03*	−0.31	3.46
DFY	−0.24	0.17	0.03*	−0.00	−0.01	0.02	−0.17	2.63
SMB	−0.04	0.36	0.03*	−0.00	−0.01	0.03*	−0.03	2.90
HML	−0.05	0.71	0.03*	−0.00	−0.02	0.02*	−0.05	3.52
MOM	−0.04*	1.33	0.03*	−0.00	−0.01	0.02*	−0.04*	4.05
LTR	−0.05	0.38	0.03*	−0.00	−0.01	0.02*	−0.04	2.98
CAY	16.23	3.44	0.02*	−0.00	−0.01	0.02*	15.60	5.86
CPI	−61.15	1.82	0.03*	−0.01	−0.01	0.01	−64.68	4.57
OG	−13.19*	17.17	0.02*	−0.00	−0.01	0.02*	−12.97*	19.28
VIX	0.00	0.82	0.02*	−0.01	−0.02	0.02	0.00	2.89

Notes: This table reports parameter estimates from the predictive return regressions in (34) with economic predictors and sentiment indicators as regressors. Newey–West t-statistics are in brackets and adjusted R^2 are expressed in percentages. The sample used is monthly data from January, 1990, to September, 2015, totalling 309 observations.

5.1. Data and parameter estimates

The monthly data for the panel of international indices spans the period from January, 2000, to September, 2015, totalling 189 observations. Implied variances are estimated by the CBOE volatility index applied to each index options dataset. Interest rates (to compute excess returns) are extracted from Datastream. Realized variances, available from Oxford-Man Institute, are based on 5 min returns along with the squared close-to-open overnight return.

Table 12 presents parameter estimates of our model for the seven indices. Generally speaking, we find similar results to the S&P 500 index case. Regarding the dynamics of $\Sigma_{t-1,t}$, we find values ranging from 0.754 to 0.914, with SMI and AEX being the least and the most persistent, respectively. The idiosyncratic component of $\Pi^C_{t,t+1}$ is generally very persistent with mean reversion coefficients between 0.830 and 0.962 for SMI and DJIA respectively. In terms of regime duration, the first regime lasts on average 6 months, the second regime on average 1.5 months, and the third regime about 3 months with small variability across indices. The model isolates the smooth component well as can be seen from the distribution of $J^P_{t-1,1}$ and $E^Q_t[J^Q_{t,t+1}]$ exhibiting a large location and variance. Finally, the link between $\Sigma_{t-1,t}$ and $\Pi^C_{t,t+1}$, represented by the parameter d_{S_t}, shows evidence of time variation, though it disappears for CAC in periods of high to extreme volatility.

Fig. 3 displays the smooth component of the VRP from our model, $\Pi^C_{t,t+1}$, along with the VRP from a model that assumes random walk dynamics for the realized variance, $\Pi^{RW}_{t,t+1}$. The shaded areas identify high vol-of-vol (light), i.e regime 2, and extreme variance event (dark), i.e. regime 3. We see that the positivity of $\Pi^C_{t,t+1}$ over the sample period is generally satisfied with only a few exceptions which are numerically small. These occurrences may appear because the short sample (189 observations) does not allow to precisely separate the signal from the noise. Regarding the regime capturing the extreme shocks on the market, i.e. the dark-grey areas in the figure, we see similarities across markets. It identifies the well known market turmoil events since 2000. For all markets, we observe an increase in $\Pi^C_{t,t+1}$ coinciding

Table 12
Parameter estimates for our model — international evidence.

State Var.	Parameter	Index						
		DJIA	STOXX	CAC	AEX	FTSE	SMI	NIKKEI
$\Sigma_{t-1,t}$	b_0	10.456	12.133	17.672	8.765	10.620	9.906	22.391
	$b_1,\ b_2$	22.738	23.188	37.024	22.570	22.827	22.356	30.500
	ϕ^{Σ}	0.806	0.845	0.837	0.914	0.828	0.754	0.789
	$\sigma^2_{\Sigma,0}$	6.656	5.461	12.307	12.265	3.301	3.505	15.183
	$\sigma^2_{\Sigma,1},\ \sigma^2_{\Sigma,2}$	24.304	54.175	64.770	35.233	36.920	96.508	25.401
$\Pi^C_{t,t+1}$	a_0	4.859	8.678	[2.916]	[2.843]	7.564	[1.501]	12.517
	$a_1,\ a_2$	7.048	11.258	9.232	8.659	9.813	6.429	20.696
	ϕ^C	0.962	0.862	0.838	0.904	0.950	0.830	0.909
	d_0	0.122	0.652	0.400	0.534	0.930	0.474	0.480
	d_1	0.917	0.829	[0.189]	0.783	[0.538]	0.479	0.783
	d_2	0.382	0.996	[0.510]	1.064	1.684	0.722	0.966
	$\sigma^2_{C,0}$	[0.569]	0.684	[0.008]	0.786	[0.019]	[1.048]	[0.236]
	$\sigma^2_{C,1},\ \sigma^2_{C,2}$	3.919	13.235	22.690	11.109	2.589	7.655	9.078
$J^P_{t-1,1}$	j^P	49.398	56.825	59.015	67.353	48.092	60.479	31.837
	$\sigma^2_{J^P}$	728.49	1765.5	2031.8	2256.9	1131.1	1461.5	580.84
$E^Q_t[J^Q_{t,t+1}]$	j^Q	44.971	56.982	47.217	59.403	32.617	37.340	42.055
	$\sigma^2_{J^Q}$	306.57	513.57	623.61	1600.3	387.49	233.09	544.81
ε^{RV}_t	$\sigma^2_{RV,0}$	1.951	7.496	13.900	[4.972]	8.149	4.692	13.780
	$\sigma^2_{RV,1},\ \sigma^2_{RV,2}$	142.49	87.216	114.24	144.08	92.130	59.119	55.206
ε^{SW}_t	$\sigma^2_{SW,0}$	11.314	5.449	14.218	7.320	12.163	4.842	[0.904]
	$\sigma^2_{SW,1},\ \sigma^2_{SW,2}$	11.791	45.938	125.19	58.811	41.953	33.536	29.662
P	p_{00}	0.847	0.637	0.801	0.742	0.837	0.827	0.777
	p_{11}	0.628	0.341	0.247	0.205	0.407	0.439	0.696
	p_{12}	0.061	0.122	0.173	0.035	0.029	0.117	0.085
	p_{22}	0.751	0.667	0.674	0.598	0.681	0.679	0.649

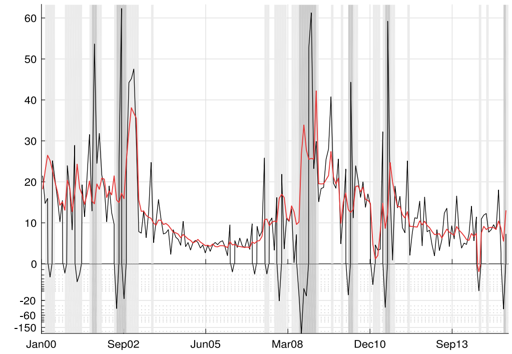
Notes: This table provides maximum likelihood estimates for our model. Parameters that are insignificant at a ten percent level are in [ ]. The sample used is monthly data from January, 2000, to September, 2015, totalling 189 observations.

with the peak of the global financial crisis (September, 2008), which instead triggers an unrealistic sharp fall in the VRP computed under the random walk dynamics.

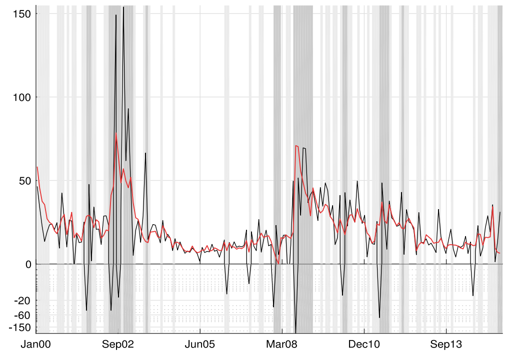
5.2. Predictive return regressions

Table 13 displays the adjusted R^2s of predictive return regressions with the VRP, sentiment indicators and their interaction with the state of the market as regressors. With respect to the smooth part of the VRP, $\Pi^C_{t,t+1}$ generally has no predictive power on future aggregate market excess returns for any horizons. An exception is SMI, showing sizeable R^2s up to the five month horizon with a maximum of 4.45 at the three month horizon. When we add to the predictive regression our sentiment indicators and their interaction with the state of the market, the R^2s globally increase in the same fashion as observed in the case of S&P 500 index and particularly so at the short and long horizons. This result clearly indicates that the information stemming from extreme variance episodes bears most of, if not all, the predictive power on future aggregate market returns. The most striking example is DJIA, where the interaction of the sentiment indicators with the sign of the market performance explain up to 13.8% of the future return variation at the four month horizon. The STOXX and FTSE indices show the same sharp increase in return predictability though to a lower extent. For instance, in the case of FTSE, the four indicators contribute to an increase in return predictability from 0.33 to 6.81 at the four month horizon. The exception is SMI where adding the sentiment indicators to the predictive regression only marginally increases the predictive power at all horizons and often significantly so.

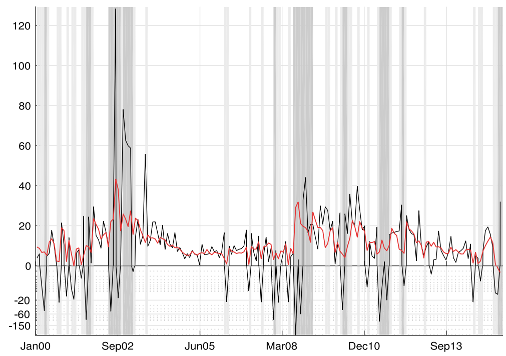
Compared to the performance of the VRPs based on random walk expectations, the R^2s of the predictive return regressions including the sentiment indicators and their interaction with the state of the market as explanatory variables largely dominates over most if not all horizons for all indices. The results reported for DJIA, which shows stronger gains at the short and long horizons, displays very similar patterns and levels as observed for the S&P 500 index. For FTSE, the extra degree of flexibility provided by the sentiment indicators, produces gains in predictability from 3.36 to 6.71 and 2.38 to 5.44 at the three and six month horizon, respectively. Another example is SMI with gains from −0.29 to 5.79 and from 0.52 to 3.87 at the three and six month horizon, respectively. Thus, the international evidence shows for several



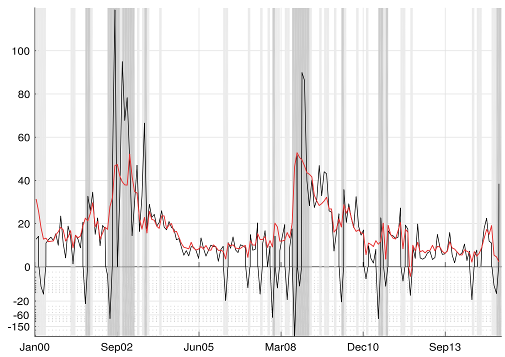
(a) DJIA



(b) STOXX

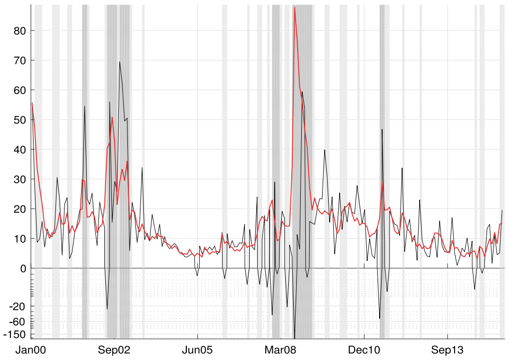


(c) CAC

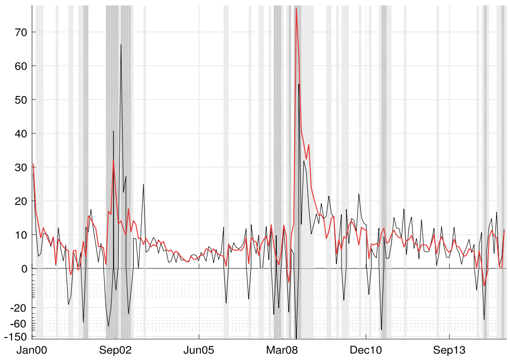


(d) AEX

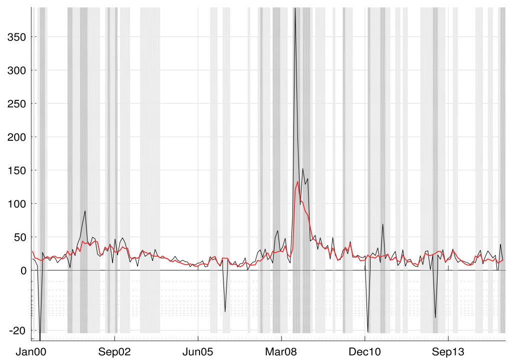
Fig. 3. $\Pi^{C}_{t,t+1}$ (red) vs $\Pi^{RW}_{t,t+1}$ (black). The shaded areas identify the high vol-of-vol (light) and extreme variance event (dark) regimes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(e) FTSE



(f) SMI



(g) NIKKEI

Fig. 3. (*continued*).

markets that the sentiment indicators have significantly more predictive power than what is obtained with the random walk VRPs.

6. Conclusion

This paper proposes a flexible approach for retrieving the variance risk premia which is essential for risk management, asset pricing and portfolio management. The method delivers realistic estimates of the market price of risk. Moreover, our model allows identifying not only the smooth component associated with the price of risk in periods of normal market activity but also to infer the occurrence and size of extreme variance events and the risk neutral expectation of these. Based on this decomposition we construct sentiment indicators based on agents' expectations response to extreme variance shocks.

Our empirical application to the S&P 500 index documents the importance of allowing for interactions, discontinuities and occurrence of extreme variance events. The filtered variance predictions and the risk neutral variance expectations match the level and the dynamics of their observable counterparts along the entire sample. Furthermore, the resulting

Table 13
Adjusted R^2 from predictive return regressions.

Regressors		Horizon							
Π^i	$SI_{j,t}$	1	2	3	4	5	6	9	12
DJIA									
$\Pi^C_{t,t+1}$		−0.56	−0.56	−0.49	−0.22	0.26	0.17	0.13	0.62
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r_t^-}$ $\Delta J^-_{t,r_t^-}$ $\Delta J^+_{t,r_t^+}$ $\Delta J^-_{t,r_t^+}$	7.20	8.18	10.44	13.80	9.87	7.68	7.03	7.18
$\Pi^{RW}_{t,t+1}$		3.30	5.24	10.63	14.06	10.31	6.84	1.41	0.53
STOXX									
$\Pi^C_{t,t+1}$		−0.56	−0.51	−0.35	−0.39	−0.51	−0.57	−0.22	−0.02
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r_t^-}$ $\Delta J^-_{t,r_t^-}$ $\Delta J^+_{t,r_t^+}$ $\Delta J^-_{t,r_t^+}$	1.53	1.90	2.09	3.10	2.71	0.29	0.24	1.28
$\Pi^{RW}_{t,t+1}$		−0.15	0.68	0.78	1.68	1.31	−0.08	−0.22	0.23
CAC									
$\Pi^C_{t,t+1}$		−0.51	−0.14	0.16	0.06	−0.42	−0.57	−0.44	−0.13
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r_t^-}$ $\Delta J^-_{t,r_t^-}$ $\Delta J^+_{t,r_t^+}$ $\Delta J^-_{t,r_t^+}$	2.44	4.53	5.72	6.21	4.39	1.60	1.74	2.34
$\Pi^{RW}_{t,t+1}$		1.31	3.59	4.60	4.33	4.13	1.92	2.03	2.83
AEX									
$\Pi^C_{t,t+1}$		0.14	0.16	0.72	0.33	−0.24	−0.53	−0.26	−0.45
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r_t^-}$ $\Delta J^-_{t,r_t^-}$ $\Delta J^+_{t,r_t^+}$ $\Delta J^-_{t,r_t^+}$	2.97	2.99	3.60	3.18	3.00	1.22	1.39	0.81
$\Pi^{RW}_{t,t+1}$		0.63	1.14	1.58	1.70	0.94	0.52	0.26	0.10
FTSE									
$\Pi^C_{t,t+1}$		0.14	0.16	0.72	0.33	−0.24	−0.53	−0.26	−0.45
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r_t^-}$ $\Delta J^-_{t,r_t^-}$ $\Delta J^+_{t,r_t^+}$ $\Delta J^-_{t,r_t^+}$	3.61	6.00	6.71	6.81	6.46	5.44	2.57	2.12
$\Pi^{RW}_{t,t+1}$		−0.37	0.71	3.36	3.93	2.74	2.38	0.38	0.45
SMI									
$\Pi^C_{t,t+1}$		2.29	2.75	4.45	4.21	2.50	0.56	−0.58	−0.58
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r_t^-}$ $\Delta J^-_{t,r_t^-}$ $\Delta J^+_{t,r_t^+}$ $\Delta J^-_{t,r_t^+}$	7.91	5.52	5.79	5.23	4.83	3.87	0.63	−0.17
$\Pi^{RW}_{t,t+1}$		0.65	−0.48	−0.29	0.73	0.83	0.52	−0.24	−0.36
NIKKEI									
$\Pi^C_{t,t+1}$		−0.53	−0.33	0.31	0.35	−0.28	−0.46	0.26	1.10
$\Pi^C_{t,t+1}$	$\Delta J^+_{t,r_t^-}$ $\Delta J^-_{t,r_t^-}$ $\Delta J^+_{t,r_t^+}$ $\Delta J^-_{t,r_t^+}$	2.12	3.06	1.43	0.74	0.35	1.65	5.42	3.05
$\Pi^{RW}_{t,t+1}$		−0.31	−0.29	−0.35	−0.05	−0.49	−0.56	−0.56	−0.36

Notes: This table reports the adjusted R^2s in percentage from the predictive return regressions in (33) with VRP, sentiment indicators and their interaction with the state of the market as regressors. The sample used is monthly data from January, 2000, to March, 2015, totalling 189 observations.

filtered variance risk premium satisfies the positivity constraint, is persistent and it appropriately reacts to changes in level and variability.

We address a puzzle in the existing literature related to return predictability by disentangling the predictability stemming from the part of the variance risk premium associated with normal sized price fluctuations from that associated with extreme tail events, i.e. tail risk. In particular, using our sentiment indicators we show that predictability comes almost entirely from the way agents adapt their expectations when exposed to extreme variance realizations and our results essentially show that future market performance is driven by fear and optimism with respect to the current state of the market. The documented predictive power is distinct from that of alternative economic predictors.

We extend our empirical results to seven other indices and find that the filtered VRP exhibits similar characteristics across the different markets. Return predictability due to the smooth component emerges in some cases and in combination with the sentiment indicators often dominates that of benchmark VRP measures for all markets.

References

Ait-Sahalia, Y., Karaman, M., Mancini, L., 2018. The term structure of variance swaps, risk premia and the expectation hypothesis. Working paper.

Andersen, T.G., Bollerslev, T., Diebold, F., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. Rev. Econ. Stat. 89, 701–720.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001a. The distribution of realized stock return volatility. J. Financ. Econ. 61 (1), 43–76.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001b. The distribution of realized exchange rate volatility. J. Amer. Statist. Assoc. 96 (453), 42–55.

Andersen, T.G., Bondarenko, O., Gonzalez-Perez, M.T., 2015a. Exploring return dynamics via corridor implied volatility. Rev. Financ. Stud. 28 (10), 2902–2945.

Andersen, T.G., Fusari, N., Todorov, V., 2015b. Parametric inference and dynamic state recovery from option panels. Econometrica 83 (3), 1081–1145.

Andersen, T.G., Fusari, N., Todorov, V., 2015c. The risk premia embedded in index options. J. Financ. Econ. 117 (3), 558–584.

Andersen, T.G., Fusari, N., Todorov, V., 2016. The pricing of tail risk and the equity premium: Evidence from international option markets. Working Paper.

Augustyniak, M., Bauwens, L., Dufays, A., 2019. The factorial hidden Markov volatility model. J. Bus. Econom. Statist. 37 (2), 696–709.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. J. Finance 61, 1645–1680.

Bakshi, G., Madan, D., 2000. Spanning and derivative-security valuation. J. Financ. Econ. 55 (2), 205–238.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. J. Finance 59 (4), 1481–1509.

Bardgett, C., Gourier, E., Leippold, M., 2019. Inferring volatility dynamics and risk-premia from the S&P500 and VIX markets. J. Financ. Econ. 131 (3), 593–618.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc. Ser. B Stat. Methodol. 64 (2), 253–280.

Bauwens, L., Dufays, A., Rombouts, J., 2013. Marginal likelihood for Markov switching and change-point GARCH models. J. Econometrics 178 (3), 508–522.

Bauwens, L., Hafner, C., Laurent, S., 2012. Handbook of Volatility Models and Their Applications. Wiley.

Bauwens, L., Laurent, S., Rombouts, J., 2006. Multivariate GARCH models: A survey. J. Appl. Econometrics 21, 79–109.

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

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

Bollerslev, T., Marrone, J., Xu, L., Zhou, H., 2014. Stock return predictability and variance risk premia: Statistical inference and international evidence. J. Financ. Quant. Anal. 49, 633–661.

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Rev. Financ. Stud. 22 (11), 4463–4492.

Bollerslev, T., Todorov, V., 2011. Tails, fears, and risk premia. J. Finance 66 (6), 2165–2211.

Bollerslev, T., Todorov, V., Xu, L., 2015. Tail risk premia and return predictability. J. Financ. Econ. 118 (1), 113–134.

Bonomo, M., Garcia, R., Meddahi, N., Tédongap, R., 2015. The long and the short of the risk-return trade-off. J. Econometrics 187 (2), 580–592, Econometric Analysis of Financial Derivatives.

Boudoukh, J., Richardson, M., Whitelaw, R.F., 2008. The myth of long-horizon predictability. Rev. Financ. Stud. 21 (4), 1577–1605.

Britten-Jones, M., Neuberger, A., 2000. Option prices, implied price processes, and stochastic volatility. J. Finance 55 (2), 839–866.

Carhart, M., 1997. On persistence in mutual fund performance. J. Finance 52, 57–82.

Carr, P., Wu, L., 2009. Variance risk premiums. Rev. Financ. Stud. 22 (3), 1311–1341.

CBOE, 2015. The CBOE volatility index–VIX. White Paper. Available at https://www.cboe.com/micro/vix/vixwhite.pdf.

Cooper, I., Priestley, R., 2009. Time-varying risk premiums and the output gap. Rev. Financ. Stud. 22 (7), 2801–2833.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. J. Financ. Econ. 7 (2), 174–196.

Dew-Becker, I., Giglio, S., Le, A., Rodriguez, M., 2017. The price of variance risk. J. Financ. Econ. 123 (2), 225–250.

Drechsler, I., Yaron, A., 2011. What's vol got to do with it. Rev. Financ. Stud. 24 (1), 1–45.

Egloff, D., Leippold, M., Wu, L., 2010. The term structure of variance swap rates and optimal variance swap investments. J. Financ. Quant. Anal. 45 (5), 1279–1310.

Engle, R.F., Gallo, G.M., 2006. A multiple indicators model for volatility using intra-daily data. J. Econometrics 131 (1), 3–27.

Francq, C., Zakoian, J., 2010. GARCH Models: Structure, Statistical Inference and Financial Applications. Wiley.

Giglio, S., Kelly, B., Pruitt, S., 2016. Systemic risk and the macroeconomy: An empirical evaluation. J. Financ. Econ. 119 (3), 457–471.

Gruber, P.H., Tebaldi, C., Trojani, F., 2015. The price of the smile and variance risk premia. Swiss Finance Institute Research Paper No. 15–36.

Huang, D., Jiang, F., Tu, J., Zhou, G., 2015. Investor sentiment aligned: A powerful predictor of stock returns. Rev. Financ. Stud. 28 (3), 791–837.

Jiang, G.J., Tian, Y.S., 2005. The model-free implied volatility and its information content. Rev. Financ. Stud. 18 (4), 1305–1342.

Kim, C.-J., 1994. Dynamic linear models with Markov-switching. J. Econometrics 60 (1–2), 1–22.

Kim, C.-J., Nelson, C.R., 1999. State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications. MIT Press Books, The MIT Press.

Kozhan, R., Neuberger, A., Schneider, P., 2013. The skew risk premium in the equity index market. Rev. Financ. Stud. 26 (9), 2174–2203.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. J. Finance 56 (3), 815–849.

Mueller, P., Vedolin, A., Yen, Y.-M., 2015. Bond variance risk premia, LSE Working paper.

Patton, A., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. Rev. Econ. Stat. 97(3), 683–697.

Wang, C.S.-H., Bauwens, L., Hsiao, C., 2013. Forecasting a long memory process subject to structural breaks. J. Econometrics 177 (2), 171–184, Dynamic Econometric Modeling and Forecasting.

Wu, L., 2011. Variance dynamics: Joint evidence from options and high-frequency returns. J. Econometrics 160 (1), 280–287.