The Price of the Smile and Variance Risk Premia

PETER H. GRUBER, CLAUDIO TEBALDI, and FABIO TROJANI*

First version: February 2014. This version: November 11, 2015

ABSTRACT

In a tractable stochastic volatility model, we identify the price of the smile as the price of the unspanned risks traded in SPX option markets. The price of the smile reflects two persistent volatility and skewness risks, which imply a downward sloping term structure of low-frequency variance risk premia in normal times. In periods of distress, the term structure is upward sloping and dominated by a high-frequency premium for jump variance. This dichotomy is consistent with the puzzling skew sensitivities of option markets with credit-constrained intermediaries and it builds a challenge for many reduced-form and structural models of stochastic volatility.

JEL classification: G10, G12, G13
Keywords: Price of the Smile, Price of Volatility, Option Pricing, Stochastic Volatility, Unspanned Skewness, Financial Constrains, Financial Intermediation, Financial Crisis, Factor Models, Matrix Jump Diffusions, Variance Swaps, Skew Swaps.

* Peter H. Gruber is from the Università della Svizzera Italiana, Lugano; Claudio Tebaldi is from IGIER, Bocconi University, and a CAREFIN fellow; Fabio Trojani is from the University of Geneva and a Senior Chair of the Swiss Finance Institute. All authors gratefully acknowledge the financial support of the CAREFIN foundation. Peter Gruber and Fabio Trojani gratefully acknowledge the financial support from the Swiss National Science Foundation (Project 150198, " Higher order robust resampling and multiple testing methods ") and the Swiss Finance Institute (Project "Term structures and cross-sections of asset risk premia"). We thank participants to the 2008 FINRISK Research Day, the X. Workshop on Quantitative Finance, the 6th World Congress of the Bachelier Society, the European Summer Symposia in Financial Markets, the 2010 and 2015 EFA meetings, the 2014 SoFIE Conference on Skewness, Heavy Tails, Market Crashes, and Dynamics, and the finance seminars at the Vienna University of Technology, Bocconi University, the University of Geneva, the EPFL Lausanne, the University of Bonn, the University of Manchester, the Luxembourg School of Finance and Banque de France, for valuable comments. The usual disclaimer applies.

Understanding the properties of the market price of volatility risk is a key issue in financial economics, as the recent macroeconomic literature has shown that time-varying uncertainty is a source of risk with real economic effects (e.g., Bloom (2009), Gourio (2012) and Gourio (2014)). The recent financial literature has reached a consensus on the fact that aggregate uncertainty shocks are priced in modern financial markets, by estimating a typically negative average excess return for long variance portfolios over different investment horizons. However, far less is known about (i) which characteristics of the volatility generate the volatility risk premium, (ii) the time-variation of the premium and its dependence on the investment horizon, (iii) the relation between the volatility risk premium and the premia of the unspanned risks traded in option markets, and (iv) the link between volatility risk premia and market equity premia.

While all these questions are key for understanding the properties of the price of volatility risk, addressing them in a single coherent arbitrage-free model is a challenge. First, evidence in Bates (2000), Gruber, Tebaldi and Trojani (2010), Calvet, Fearnley, Fisher and Leippold (2013) and Andersen, Fusari and Todorov (2015), among others, shows that volatility risk originates from multiple risks with distinct persistence properties. These risks comove dynamically and can contribute in distinct ways to the price of volatility. Second, volatility risk is a particular form of unspanned risk, which is tradeable in liquid option markets together with others risks, such as skewness risk. Therefore, its price needs to be studied consistently with the price of all tradeable unspanned risks, i.e., consistently with the price of the smile; see Buraschi and Jackwerth (2001), among others.[1] Third, the specification of the volatility has to be consistent with the cross-sectional and time series properties of the option-implied volatility smile, because shocks to option-implied volatilities reveal useful information for identifying the unspanned risks driving option prices. Importantly, such an identification is impaired by most existing arbitrage-free models, such as Bates (2000), because of the counterfactually tight relation these models induce between volatility and skewness, as noted in, e.g., Gruber et al. (2010), Andersen et al. (2015) and Constantinides and Lian (2015).

In this paper, we study the relation between the price of the smile and variance risk premia, using a single coherent arbitrage-free framework that systematically addresses questions (i)-(iv). We jointly identify the unspanned risks traded in option markets and the price

[1] This consistency precludes unplausible arbitrage opportunities between option portfolios that trade distinct characteristics of unspanned risks. Unspanned risks can be traded in option markets using either (i) dynamically hedged option portfolios, which typically require model assumptions about different option price sensitivities with respect, e.g., to time-varying volatility and skewness, or (ii) model-free static option portfolios that are dynamically delta hedged in the forward market; see Bergomi (2004), Bergomi (2005), Bergomi (2008), Bergomi (2009), Kozhan, Neuberger and Schneider (2010), Schneider and Trojani (2014a) and Schneider and Trojani (2014b), among others.

of the smile, by estimating a parsimonious three-factor stochastic volatility model in the class of matrix affine jump diffusion (AJD) proposed by Leippold and Trojani (2008). Our model features two important properties that are useful to better identify and interpret the unspanned risks traded in option markets and the price of the smile. First, a specification based on three interdependent risks, which are mutually-exciting and have distinct degrees of persistence. Second, a stochastic skewness and a price of the smile that are not linearly spanned by the level of the spot volatility. By weakening the tight link between volatility and skewness induced by most arbitrage-free models, our model produces a more accurate description of the cross-section and the time series of option-implied volatilities. In a similar way, the weaker link between the price of the smile and volatility in our model allows us to better identify risk premium components that are genuinely driven by a time-varying skewness.[2]

We estimate the model using a simple two-step procedure, based on the joint information of a panel of S&P500 index (SPX) option prices and a panel of excess returns of option volatility portfolios, in the sample period from January 1996 to January 2013. In the first step, we exploit the information from the panel of option prices to estimate the physical and risk neutral dynamics of the unspanned risks driving the SPX implied volatility surface, together with the risk-neutral parameters of the jump component in the underlying's returns. In this way, we identify the time series of unspanned risks traded in the SPX option market and the price of the smile. In the second step, we estimate the parameters of the jump variance risk premium, from a simple arbitrage-free regression of the payoffs of synthetic variance swaps on the unspanned risks identified in the first step. With this identification strategy, we estimate the time series of unspanned risks in option markets, the price of the smile and the term structure of variance risk premia, without relying on direct information about S&P 500 index returns or a complete specification of the index equity premium.

The main findings resulting from our model estimation are the following. First, our model produces an excellent fit of the cross-sectional and the time series properties of SPX option-implied volatility smiles, relative to benchmark arbitrage-free models in the literature, such as a two-factor Bates (2000) model. Importantly, the improved fit of the smile induces a good identification of unspanned risks, in terms of three interdependent state variables X_{22}, X_{12} and X_{11} driving the volatility dynamics in our model. These state variables capture risk and risk premium dynamics at three distinct frequencies, inducing different contributions to the price of the smile, over time and in dependence of the investment horizon.

[2]Our model specification nests a number of important affine stochastic volatility models in the literature, such as Bates (2000) two-factor jump diffusion or Heston (1993)-type two-factor volatility models.

Second, the unspanned risks in our model have a natural interpretation as directly observable characteristics of the smile. The first state variable (X_{22}) has a half life of about five weeks and models the high frequency component of the spot volatility. It closely follows the 30-days at-the-money implied variance, with a weekly correlation of 91%. The second state variable (X_{12}) has a half life of about one quarter and captures the component of the stochastic skewness that is not generated by shocks in the spot volatility. It closely targets the 30-days implied skew, with a weekly correlation of −89%. The third state variable (X_{11}) has a half life of more than one year and models the low frequency component of the spot volatility. It closely traces the 12 months at-the-money implied variance, with a weekly correlation of 91%. According to this evidence, the unspanned risks traded in option markets are naturally summarized by three dynamically correlated state variables, which closely reflect transient implied volatility shocks, moderately persistent implied skewness shocks and highly persistent implied volatility shocks.

Third, we find that the price of the smile is negative, highly time-varying and fully explained by the level of the two more persistent risks X_{11} and X_{12}. Therefore, these state variables can be interpreted as risk premium factors driving the price of unspanned risks in option markets. The structure of the price of the smile is striking. We document that the price of the most persistent volatility risk X_{11} is proportional to the level of X_{11} itself. In this sense, a shock in long-term implied volatilities is also a shock in the risk premium of long-term implied volatility risk. In contrast, the price of the moderately persistent skewness risk X_{12} is explained by the levels of both X_{11} and X_{12}. Therefore, shocks in short-term implied skewness only partly explain the risk premium for short-term implied skewness risk. Finally, we find that the price of the highly transient volatility risk X_{22} is proportional to the level of X_{12}. In this sense, shocks in short-term implied skewness are directly linked to shocks in the risk premium for high-frequency volatility risk. However, since risks X_{22} and X_{12} are only weakly correlated, a shock in high-frequency volatility risk can materialize even in absence of a variation in its price. These findings have natural implications for the term structures of the price of the smile, defined by the risk premia of swap cantracts with floating leg $\frac{1}{\tau}\int_0^\tau X_{ijt}dt$, where $\tau > 0$ is the swap time to maturity. As this term structure depends on linear combinations of risks X_{11} and X_{12} alone, it is naturally downward sloping. It is typically steeper for the persistent risks X_{11} and X_{22}, while it flattens out early, at horizons of about one quarter, for the more transient risk X_{22}.

Fourth, we address the relation between the unspanned risks traded in options markets, the price of the smile and the excess returns of popular variance swaps. The time-variation of all unspanned risks implies highly time-varying and unambiguously negative variance risk premia, ranging between zero and −16 percent squared (−11 percent squared), on an

annualized basis, for a monthly (an annual) investment horizon. At short horizons of one month, variance risk premia are almost completely explained by a time-varying premium for pure jump variance risk, while at horizons of twelve months the premium for diffusive variance explains about one quarter of the total variance risk premium. This structure of the variance risk premium is fully consistent with the properties of the unspanned risks and the price of the smile estimated with our model. Indeed, at short horizons the dynamics of the pure jump variance risk premium is completely explained by the probability of a jump in returns, which is a linear function of all unspanned risks X_{11}, X_{12} and X_{22}. In contrast, at longer horizons variance risk premia are more systematically explained by the term structure of the price of the smile, which is a linear function of the two more persistent unspanned risks X_{11} and X_{12} alone.

Fifth, the unspanned risks and the price of the smile in our model generate sharp implications for the term structure of variance risk premia. While this term structure is usually downward sloping, reflecting the higher price of option insurance at longer horizons, it can become strongly upward sloping for short periods of time. This feature is explained by the interplay of the term structure of the price of the smile with the term structure of the expected number of future large negative returns. The first of these term structures implies a downward sloping term structure of the risk premia for exposure to future diffusive variance or jump intensity realizations. When jump intensities are sufficiently large, the second term structure is typically downward sloping, due to the mean reversion of unspanned risks. This yields an upward sloping term structure of pure jump variance risk premia, because the variance risk premium of a sure jump of unknown size is negative. Thus, the term structure of variance risk premia is upward sloping, whenever the term structure of pure jump variance risk premia dominates the term structure of the smile at short horizons. We document that an upward sloping term structure typically emerges in conditions of market distress. These market conditions naturally reflect transient states, in which short-term market downside risk can become extraordinarily high and insurance against a sudden large market downturn is particularly expensive. Therefore, it is plausible that precisely in such states the Value At Risk constraints of credit-constrained suppliers of option insurance are particularly binding; see again Constantinides and Lian (2015).

Finally, we test our model specification and findings along several additional dimensions, by means of different robustness checks, documented in detail in Section III. of the Online Appendix. We verify in a more nonparametric predictive regression context the role of unspanned risks traded in option markets as risk premium factors for equity and variance risk premia. Predictive regressions results suggest that unspanned risks have an economically relevant predictive power, both for S&P 500 index and variance swap excess returns, with a

dominating contribution to the predictive power deriving from the two persistent unspanned risks. The affine specification of variance risk premia in our model is preferred by the out-of-sample predictability results, providing additional support to our specification of the price of variance risk. In contrast, affine predictive relations for S&P 500 index excess returns are dominated by a nonlinear threshold specification, in which unspanned risks generate a significant degree of predictability only in states of sufficiently persistent volatility. This finding further motivates an identification of variance risk premia that does not directly depend on an affine specification of index equity premia.

Review of the Literature.

Our work borrows from an enormous literature that has studied the economic sources of volatility variations, the dynamics of the option-implied volatility smile, the origins of a negative variance premium and the relation with market equity premia. We contribute to this literature along several dimensions.

First, we make use of a novel specification of stochastic volatility, which parsimoniously identifies the multi-frequency unspanned risks traded in option markets, together with the price of the smile and the term structure of variance risk premia. Following Heston (1993)'s seminal model, it has been early recognized that volatility is a multi-frequency object dependent on risks with distinct persistence and variability properties. Bates (2000) was the first to estimate with a panel of option prices a tractable two-factor model for index returns. Subsequent papers, such as Huang and Wu (2004) and Christoffersen, Heston and Jacobs (2009), have quantified the improvements in the fit of the option-implied volatility smile using two-factor models with independent volatility components. The more recent literature has explored reacher three-factor specifications. Carr and Wu (2009a), Gruber et al. (2010) and Andersen et al. (2015), among others, clearly improve on the fit of the smile provided by benchmark two-factor models. Our specification of stochastic volatility is different and complements these approaches, based on three mutually-exciting risks that follow an affine jump diffusion on the state space of symmetric positive definite matrices, under the physical and the pricing measures.[3] This approach yields a direct identification and interpretation of correlated unspanned risks. By embedding a dynamic skewness component disconnected from the spot volatility, it also avoids the puzzling skew sensitivities of benchmark arbitrage-free models noted in Constantinides and Lian (2015). Our goal is also different from Carr and Wu (2009a) and Andersen et al. (2015). Carr and Wu (2009a) identify in a semi-structural

[3] See, among others, Gourieroux (2006), da Fonseca, Grasselli and Tebaldi (2008) and Buraschi, Porchia and Trojani (2010) for examples and applications of affine matrix-valued diffusions, as well as Leippold and Trojani (2008) for a broad class of affine matrix jump diffusion processes.

model the economic channels generating the equity volatility, by mapping them on a leverage, a volatility feedback and a self-exciting component. Andersen et al. (2015) identify with a penalized nonlinear least-squares approach the unspanned risks revealed by a panel of SPX options. They remain agnostic about the specification of the price of the smile and study the role of unspanned risks as risk premium factors for equity and volatility risk. We rely on a different approach to study in a coherent no-arbitrage setting the unspanned risks traded in option markets, the price of the smile and variance risk premia. Given the need for a parsimonious specification of physical and risk neutral state dynamics, we adopt a matrix jump diffusion with jumps in index returns. While it would be possible to introduce jumps in the volatility process, a parsimonious model would require not innocuous additional assumptions, in order to identify the jump components of the volatility and jump volatility risk premia.[4]

Second, our paper borrows from a large literature that has studied the trading of unspanned risks in option markets, the market price of volatility and the term structure of variance risk premia. In a first strand of this literature, Dupire (1993) and Neuberger (1994) where among the first to propose synthetic option portfolio strategies for trading proxies of realized variance, followed by Carr and Madan (1998), Demeterfi, Derman, Kamal and Zou (1999) and Britten-Jones and Neuberger (2000), among others. From the price of such portfolios, the price of variance can be measured in a model-free way, giving rise to a variety of synthetic variance swap contracts. Recent papers have focused on the properties of variance swaps in presence of jumps and on swap contracts for trading higher-order risks, such as, e.g., skewness and kurtosis.[5] A key insight of this literature, which motivates our work, is the tradeability of variance, skewness and higher-order unspanned risks by means of appropriate option portfolios. Given the no-arbitrage constraints prevailing in liquid option markets, the prices of these risks may be naturally interconnected and difficult to study in isolation. Therefore, a coherent treatment of the price of variance risk naturally calls for a joint arbitrage-free specification of the unspanned risks traded in option markets and the price of the smile.

A second strand of this literature has established the existence of a negative risk premium

[4]For instance, Carr and Wu (2009a) assume that the probability of a co-jump in returns and volatility follows a pure-jump single-factor dynamics. Such an assumption restricts the jump variance risk premia to be perfectly correlated across horizons, which we feel excessively constrains the term structure of variance risk premia for our analysis.

[5]Martin (2012), Neuberger (2012) and Bondarenko (2014) introduce definitions of variance swap payoffs robust to jumps. Kozhan et al. (2010) propose a synthetic skew swap to study skewness vs. variance risk premia, while Schneider and Trojani (2014b) trade and price fear using Hellinger skew swaps. More broadly, Schneider and Trojani (2014a) introduce power divergence swaps for trading general nonlinear risks and characterize in a model-free way the premia for divergence risks of different orders.

for market volatility and has studied its properties. Buraschi and Jackwerth (2001) test the spanning properties of option markets and conclude in favour of models with priced unspanned risks, such as stochastic volatility or jump risk. Bakshi and Kapadia (2003) provide first direct evidence on a negative variance risk premium using delta-hedge call option positions. Similar evidence is obtained by Wu (2011), using the payoffs of (over-the-counter) variance swaps, and by Carr and Wu (2009b), using synthetic variance swaps on several underylings. Todorov (2010) and Bollerslev and Todorov (2011) conclude that variance risk premia are dominated by a premium for jump variance risk, which tends to increase after a negative jump has occurred. Our findings are consistent with the evidence in this strand of the literature. Importantly, we obtain a new decomposition of variance risk premia, into the contribution of three economically interpretable unspanned risks. These risks feature distinct persistence features and are priced very differently. With this decomposition, we document that in periods of distress the term structure of variance risk premia is partly disconnected from the price of the smile at the short end, when volatility and jump risk are large and particularly transient. In contrast, the long end of the term structure of variance risk premia is more systematically spanned by the price of the smile and the risk premia for persistent risks.

A third strand of this literature has studied the properties of the term structure of variance risk premia. Ait-Sahalia, Karaman and Mancini (2012) and Filipovic, Gourier and Mancini (2015) estimate an affine and a quadratic two-factor volatility model, based on (over-the-counter) variance swap rates of maturities between two months and two years, focusing on the implications for the term structure of equity vs. variance risk premia and on the structure of optimal portfolios with variance swaps, respectively. The first paper estimates an unambiguously negative and downward sloping term structure of variance risk premia. The second paper documents that the optimal portfolio contains an important long-short position in long versus short term variance swaps, which earns the premium implied by the decreasing term structure of variance risk premia and simultaneously limits portfolio losses when volatility increases. Dew-Becker, Giglio, Le and Rodriguez (2014) estimate a discrete-time version of Ait-Sahalia et al. (2012)'s model, using different sets of variance swaps with maturities from 1 month to 14 years. They document that a steep term structure of variance risk premia at the short end is a strong puzzle for recent parameterizations of structural long-run risk models, such as Drechsler and Yaron (2011) and Wachter (2013), but less so for models with a time-varying stock market exposure to rare disasters, such as Gabaix (2012). Our paper studies the term structure of variance risk premia with a different approach, by recognizing the joint tradeability of variance, skewness and higher-order unspanned risks by means of appropriate option portfolios. In this way, we jointly identify in a parsimonious three-factor model the unspanned risks traded in option markets, the price of the smile and

the term structure of variance risk premia. In contrast to the literature, we find that while the term structure of variance risk premia is downward sloping in normal times, it can be strongly inverted in periods of financial distress when jump variance risk premia are unusually large.[6] Importantly, both low- and high-frequency unspanned risks are key to understand these dynamics. The long end of the term structure is strongly connected to the price of the smile, which is jointly explained by a moderately persistent skewness risk and a highly persistent volatility risk. In contrast, the short end of the term structure depends also on a high-frequency volatility risk, which can be very substantial in phases of market distress. As this risk is not a risk premium factor, it does not directly contribute to variations in the price of the smile. However, since it is positively related to the probability of a negative jump in returns, it can generate sizable jump variance risk premia in states of large and transient volatility, together with a steep inverted term structure of variance risk premia.

The insights provided by our findings can help to understand different structural mechanisms for explaining variance risk premia. Our evidence indicates that the equilibrium mechanisms implied by long-run risk models might be useful to understand the persistent dynamics of the long end of the term structure of variance risk premia. In contrast, the term structure dynamics in periods of distress may be better explained by high-frequency volatility shocks, which materialize without affecting much the price of the smile. Such a mechanism may be rationalized by structural rare disaster models with a time-varying stock market exposure that is only weakly correlated with aggregate consumption shocks, as in Gabaix (2012). More generally, the multi-frequency structure of variance risk premia in our model is consistent with a price of volatility risk that can depend on high frequency shocks generated by situations of financial distress. Adrian and Rosenberg (2008) decompose market volatility into two weakly persistent components with a half-life of less than a quarter, which are priced in the cross-section of stock returns. They explicitly interpret their highest frequency volatility component as a proxy of skewness risk reflecting the tightness of financial constraints. Adrian and Shin (2010) show that expansions and contractions of repo and commercial paper funding predict variations in option-implied volatility, while Adrian, Moench and Shin (2013) document empirically the link between financial intermediaries balance sheets and asset prices. Muir (2013) emphasizes the high-frequency character of financial crises and explains in a theoretical model with financial intermediation why the term structure of the price of volatility and the term structure of variance risk premia can be inverted in phases of financial turmoil. Barras and Malkhozov (May 2014) document that the variance risk premia inferred from option markets contain a component related

[6]These features have likely implications for optimal portfolios including variance swaps, as shorting long against short term variance swaps when the term structure is inverted can be potentially very costly.

to measures of the financial standing of intermediaries, which explains the difference with the variance risk premium estimated from equity markets. The dynamics of variance risk premia estimated by our model, in particular the high-frequency character of the inverted term structure of variance risk premia in periods of financial distress, is consistent with the economic intuition motivating this literature.

Finally, our paper is related to the literature studying the risk premium factors for market returns and their relation to variance risk premia. Bollerslev, Tauchen and Zhou (2009) where the first to document the predictive power of variance risk premia, proxied by the difference of implied and realized volatilities, for future S&P500 index returns. More recently, Andersen et al. (2015) identify from a panel of SPX options three unspanned risks, in order to address their role as risk premium factors for index and variance risk premia. They find that a single risk capturing option-implied skewness unspanned by volatility has a large predictive power for variance and S&P 500 index returns. Our predictive regression results are consistent with these findings and generate a number of additional insights. Coherently with Andersen et al. (2015), we document that two persistent unspanned risks disconnected from the spot volatility jointly exhaust the large (in-sample) predictive power for S&P 500 index returns; see Section III.C. of the Online Appendix. Interestingly, we find that these risk premium factors naturally span the risk premium factor for unspanned option-implied skewness in Andersen et al. (2015).[7] Importantly, we provide a clear interpretation of these risk premium factors, in terms of observable option-implied volatility and skew components, and we identify their relation with the price of the smile and variance risk premia in a coherent no-arbitrage model. In this way, we obtain a more structural understanding of the multi-frequency dynamics of the term structures of the price of the smile and variance risk premia. Finally, we document that the relation between unspanned risks and S&P 500 index equity premia is possibly not affine and dependent on the frequency-composition of the volatility. This evidence further supports the robustness of our two-step identification of the price of the smile and variance risk premia.

The rest of the paper is organized as follows. Section **I.** introduces our three-factor stochastic volatility model, together with the closed-form expressions for the term structure of variance risk premia. Section **II.** presents our empirical findings, while Section **III.** concludes and highlights avenues for future developments.

[7] To illustrate, full sample regressions of the predictive factor for unspanned skewness in Andersen et al. (2015) on our two persistent unspanned risks produces significant results and R^2s of about 95%. We are grateful to Nicola Fusari for having provided us with the time series of the unspanned risks identified in Andersen et al. (2015).

I. Model

Our model is characterized by three mutually exciting unspanned risks, a time-varying skewness partly disconnected from the spot volatility and a compensation for variance risk that can vary separately from the spot volatility. To motivate our modelling approach with respect to well-known benchmarks in the literature, we embed Bates (2000) two-factor jump diffusion in a more general state dynamics, within the class of matrix affine jump diffusions (AJD) proposed in Leippold and Trojani (2008).

A. Bates (2000) Benchmark Volatility Model

In Bates (2000) model, returns are driven by two independent unspanned risks and follow a Poisson-Normal jump process. We denote by S_t the value of an equity index at time t, by r and q the (constant) interest rate and dividend yield, and by v_{1t}, v_{2t} the two volatility components. Under the risk-neutral probability measure $\mathbb{Q}$, the return dynamics is:

$$\frac{dS_t}{S_{t-}} = (r - q - \lambda_t \bar{k})dt + \sqrt{v_{1t}}dz_{1t} + \sqrt{v_{2t}}dz_{2t} + kdN_t \ , \tag{1}$$

where z_1, z_2 are independent standard Brownian motions and the volatility components have the dynamics:

$$dv_{it} = (\alpha_i - \beta_i v_{it})\, dt + \sigma_i \sqrt{v_{it}} dw_{it} \quad ; \quad i = 1, 2 \quad , \tag{2}$$

where w_1 and w_2 are independent standard Brownian motions, having correlation ρ_1 and ρ_2 with z_1 and z_2, respectively. Return jumps kdN_t feature an affine jump intensity

$$\lambda_t := P_t(dN_t = 1)/dt = \lambda_0 + \lambda_1 v_{1t} + \lambda_2 v_{2t} \ , \tag{3}$$

and a jump size k with expected value $\bar{k} = E^{\mathbb{Q}}(k)$.[8] The well-known volatility feedback effect is captured by the (stochastic) covariance between returns and diffusive variance $v_{1t} + v_{2t}$:

$$Cov_t\left(dS_t/S_{t-}, d(v_{1t} + v_{2t})\right) = \rho_1 v_{1t} + \rho_2 v_{2t} \ . \tag{4}$$

In addition to the volatility feedback effect, the time varying jump intensity (3) generates a jump-driven channel for stochastic return skewness.

Two features of Bates (2000) model are interesting for motivating our modelling ap-

[8]Different assumptions can be made on the distribution of log return jumps $\ln(1+k)$. Bates (2000), e.g., assumes $\ln(1+k) \sim N\left(\ln\left(1+\bar{k}\right) - \frac{\delta^2}{2}\, , \delta^2\right)$. Alternative specifications include a double exponential or similar distributions.

proach. First, v_{1t} and v_{2t} are mutually independent. Therefore, they may be difficult to interpret in terms of observable, potentially correlated, unspanned risks traded in option markets. Second, jump intensity (3) and volatility-feedback effect (4) are functions of v_{1t} and v_{2t} alone, meaning that shocks to risk neutral skewness always correlate with a shock to the spot volatility. This feature produces a tight link between risk-neutral volatility and risk-neutral skewness, which is hardly consistent with the data and can impair the identification of unspanned skewness risks in the smile dynamics.[9] Figure 1 documents more systematically this important aspect.

[Insert Figure 1 about here.]

In Figure 1, we scatter plot two measures of the short-term risk-neutral skewness and the slope of the term structure of at-the-money implied volatilities. In panel A, we compute such measures both model-free (grey data points), based on the panel of SPX options in the time span from January 1996 to January 2013, and using the fitted parameters and states of Bates (2000) model (black data points). We isolate in two ways the effect of the level of the volatility on the risk-neutral skewness and the slope of the implied volatility term structure. First, we scale both proxies by the 30-days at-the-money implied volatility. Second, we stratify the sample in four subsamples associated with different at-the-money implied volatilities. In this way, we document two important stylized facts. First, in each scatter plot the model-free proxies have a very large degree of variability, reflecting a dynamics of the SPX volatility smile that is partially disconnected from the level of the implied volatility. Second, the two proxies of the smile in Bates (2000) model are linked by a virtually deterministic relation, showing that the two-factor specification counterfactually constrains the dynamics of the smile. A different evidence emerges from the scatter plots in panel B of Figure 1. These plots are produced by our three-factor model in Section II.*B.* and are clearly better consistent with the loose link between risk-neutral skewness and steepness of the implied volatility term structure in the data.[10]

[9] In models with perfectly correlated spot volatility and jump intensity, Constantinides and Lian (2015) highlight a counterfactual negative relation between (i) the risk-neutral skewness and (ii) the risk-neutral variance or the disaster index. As demonstrated in Figure 1 of the Online Appendix, these puzzling skew sensitivities do not arise in an unconstrained two-factor Bates (2000) model with not perfectly correlated volatility and jump intensity.

[10] Gruber et al. (2010) discuss the role of dynamically interacting components of short-run volatility, long-run volatility and unspanned skewness for parsimoniously capturing the dynamics of the option-implied volatility surface.

B. The Three-Factor Matrix AJD Volatility Model

Motivated by the previous evidence, we nest Bates (2000) model in a broader three-factor dynamics, driven by a matrix-valued diffusion X of symmetric and positive definite matrices

$$X_t := \begin{pmatrix} X_{11t} & X_{12t} \\ X_{12t} & X_{22t} \end{pmatrix} . \tag{5}$$

Positive definiteness of matrix X_t allows us to coherently specify diagonal elements X_{11} and X_{22} as (diffusive) variance risks, while out-of-diagonal element X_{12} can be used to model shocks in jump intensities and volatility feedbacks that are unspanned by X_{11} and X_{22}.

B.1. State Dynamics

We obtain a tractable affine dynamics for state X_t, using the Wishart diffusion of Bru (1991).

Assumption 1 *Symmetric positive semi-definite process X_t follows the affine dynamics*[11]

$$dX_t = [\beta Q'Q + MX_t + X_t M']dt + \sqrt{X_t}dB_tQ + Q'dB_t'\sqrt{X_t} \ , \tag{6}$$

where $\beta > 1$, M, Q are 2×2 parameter matrices and B is a 2×2 standard Brownian motion under risk-neutral martingale measure $\mathbb{Q}$. $\sqrt{X_t}$ denotes the symmetric square root of X_t.

Note that when matrices M or Q are not diagonal, all states X_{11t}, X_{22t} and X_{12t} are dynamically interconnected, because their drifts and volatilities depend on all state variables in equation (6). When M and Q are diagonal, vector (X_{11t}, X_{22t}) is an autonomous Markov process with components distributed as independent Heston (1993) volatility processes. For this case, the state dynamics of Bates (2000) model is nested in equation (6).

B.2. Risk-Neutral Return Dynamics and Nested Models

We specify the risk-neutral return dynamics by the following affine jump-diffusion.

Assumption 2 *Under risk neutral probability measure $\mathbb{Q}$, the dynamics of S_t is given by:*

$$\frac{dS_t}{S_{t-}} = (r - q - \lambda_t \overline{k})dt + tr(\sqrt{X_t}dZ_t) + kdN_t \ , \tag{7}$$

[11] Positive semi-definiteness (positive definiteness) of X_t follows if $\beta > 1$ ($\beta > 3$), ensuring that the volatility components cannot cross (reach) the zero boundary.

where X_t follows the dynamics (6),

$$Z_t = B_t R + W_t \sqrt{I_2 - RR'} \ , \tag{8}$$

with $tr(\cdot)$ denoting the trace operator, W another 2×2 standard Brownian motion, independent of B, and R a 2×2 matrix such that $I_2 - RR'$ is positive semi-definite. Return jumps follow a compound Poisson process kdN_t with jump intensity $\lambda_t = \lambda_0 + tr(\Lambda X_t)$, for $\lambda_0 \geq 0$, a 2×2 matrix Λ and an iid jump size k such that $\bar{k} = E^{\mathbb{Q}}[k]$. The distribution of log return jumps $J := ln(1+k)$ is a double exponential with parameter $\lambda^+, \lambda^- > 0$ and density:[12]

$$f(J) = \frac{\lambda^+ \lambda^-}{\lambda^+ + \lambda^-} \left[e^{-\lambda^- J^- - \lambda^+ J^+} \right] \ , \tag{9}$$

with $J^+ := \max(J, 0)$ ($J^- := \max(-J, 0)$) the positive (negative) part of log return jumps.

The expression for the return variance in model (7) is given by:

$$Var_t \left(dS_t / S_{t-} \right) = tr(X_t) + \lambda_t E(k^2) = X_{11t} + X_{22t} + \lambda_t E(k^2) \ . \tag{10}$$

This shows that state variable X_{12} is instantaneously unrelated to the (diffusive) spot variance $Var_t \left(dS_t^c / S_t^c \right) = tr(X_t)$. However, X_{12} does in general affect the jump intensity $\lambda_t = \lambda_0 + tr(\Lambda X_t)$ and the volatility feed-back effect, because:

$$Cov_t \left(\frac{dS_t}{S_{t-}}, d(X_{11t} + X_{22t}) \right) = 2tr(R'QX_t) \ . \tag{11}$$

In summary, whenever RQ' and Λ are not diagonal matrices, state variable X_{12} impacts the jump-driven volatility, the jump-driven skewness and the diffusive skewness. At the same time, it is absent from the diffusive variance.

Remark 3 *A diagonal matrix $R'Q$ (Λ) gives rise to Bates (2000) specification of volatility feedbacks (stochastic intensities). Thus, a diagonal model in Assumption 2 (with diagonal matrices M, Q, R and Λ) is simply Bates (2000) model.*[13] *Note that all diagonal models in our setting have independent volatility risks, as well as jump intensities and volatility feedbacks that are linear functions of the (diffusive) variance risks. Table 1 of the Online Appendix gives a summary of benchmark models related to Assumption 2. We denote by SV_{rq} diffusion*

[12] We adopt a double exponential distribution for risk-neutral log return jumps because of its parsimony and flexibility. We have estimated our model also using a normal distribution, as in Bates (2000), and we have obtained similar results to those reported in the paper.

[13] In order to fully nest Bates (2000)- and Heston (1993)-models in our setting, we specify β as a diagonal matrix B when both Q and M are diagonal.

and by SVJ_{rq} jump diffusion models, according to the numbers r and q of state variables and skewness components disconnected from volatility, respectively. For comparison, we also report the total number of parameters necessary for a complete specification of the risk-neutral and the physical dynamics in our two-step estimation approach.

It is useful to note that the three-factor return specification in Assumption 2 is only slightly less parsimonious than Bates (2000) model, with three additional parameters. In constrast, a three-factor Bates (2000)-type model implies seven additional parameters. Parsimony of our three-factor specification helps the good identification of parameters and unspanned risks in our two-step estimation procedure.

B.3. Option Valuation

Assumption 2 and Assumption 7 yield closed-form risk-neutral transforms in our matrix AJD setting, which are useful to compute the prices of plain vanilla options; see Carr and Madan (1999) and Duffie, Pan and Singleton (2000), among others. Following Leippold and Trojani (2008), the exponentially affine conditional Laplace transform for $Y_T := \log(S_T)$ is given by:

$$\Psi(\tau;\gamma) := E_t\left[\exp\left(\gamma Y_T\right)\right] = \exp\left(\gamma Y_t + tr\left[A(\tau)X_t\right] + B(\tau)\right) , \tag{12}$$

where $\tau = T - t$ and $A(\tau) := C_{22}(\tau)^{-1}C_{21}(\tau)$, with 2×2 matrices $C_{ij}(\tau)$ and scalar $B(\tau)$ given in closed form in Section **I.***A.* of the Online Appendix.

Remark 4 *In contrast to Bates (2000)-model, the computation of the risk neutral transform cannot be reduced to calculations that involve only scalar exponential and logarithmic functions, because $C_{ij}(\tau)$ and $B(\tau)$ depend on a matrix exponential and a matrix logarithm. This makes the computation of Laplace transform (12) computationally more intensive by at least one order of magnitude. We obtain an efficient computation using a version of the Cosine-FastFourierTransform (COS) method proposed by Fang and Oosterlee (2008).*

B.4. Affine Market Price of Risk and Physical Dynamics

Consistently with Leippold and Trojani (2008), we can price the matrix B of Brownian shocks in Assumption 2, driving the unspanned risks in our model, with a stochastic discount factor that preserves an affine dynamics under the physical probability measure $\mathbb{P}$. The resulting market price of risk is detailed in the next assumption.

Assumption 5 *The change of measure from the physical probability $\mathbb{P}$ to the risk neutral probability $\mathbb{Q}$ is such that:*

$$dB^* = dB - \left(\sqrt{X_t}\Gamma + \frac{1}{2\sqrt{X_t}}(\beta^* - \beta)Q' \right) dt \ , \tag{13}$$

where either $\beta^* > 3$ or $\beta^* = \beta$, Γ is a 2×2 parameter matrix and B^* is 2×2 standard Browian motion under the physical probability measure.[14]

Remark 6 *Given Assumption 5, the $X-$dynamics under the physical probability $\mathbb{P}$ is:*

$$dX_t = [\beta^* Q'Q + M^* X_t + X_t(M^*)']dt + \sqrt{X_t}dB_t^* Q + Q'dB_t^{*\prime}\sqrt{X_t} \ , \tag{14}$$

where

$$M^* = M + \Gamma Q \ . \tag{15}$$

When $\beta \neq \beta^*$, the condition $\beta^* > 3$ implies that process X is positive definite under probability $\mathbb{P}$. In all other cases, we require $\beta^* = \beta$, i.e., a completely affine market price of risk.[15] *When $\beta \neq \beta^*$, we obtain an extended affine market price of risk that allows the price of the volatility factors X_{11t} and X_{22_t} to be inversely related to the (diffusive) volatility. Cheridito, Filipovic and Kimmel (2007) propose a class of yield curve models with an extended affine market price of risk in the context of affine models with standard state spaces.*

A convenient feature of Assumption 5 is that the price of a shock in any of the unspanned risks can depend on all other risks X_{11}, X_{22} and X_{12}. This feature implies a price of the smile having rich interconnections, as can be seen from the difference of the $\mathbb{P}$ and $\mathbb{Q}$ expectation of a shock in state X, which summarizes the instantaneous risk premia of unspanned risks:

$$\frac{1}{dt}(E^{\mathbb{P}} - E^{\mathbb{Q}})[dX_t] = (\beta^* - \beta)Q'Q + \Gamma Q X_t + X_t Q'\Gamma' \ . \tag{16}$$

This matrix of risk premia is state dependent when $M^* - M = \Gamma Q \neq 0$. Moreover, when ΓQ is a diagonal matrix with diagonal components D_1, D_2, the risk premia of unspanned risks are disconnected:

$$\frac{1}{dt}(E^{\mathbb{P}} - E^{\mathbb{Q}}) \left[d \begin{pmatrix} X_{11t} & X_{12t} \\ X_{12t} & X_{22t} \end{pmatrix} \right] = (\beta^* - \beta)Q'Q + \begin{pmatrix} 2D_1 X_{11t} & (D_1 + D_2)X_{12t} \\ (D_1 + D_2)X_{12t} & 2D_2 X_{22t} \end{pmatrix} ,$$

[14] $1/\sqrt{X_t}$ denotes the unique inverse square root of symmetric positive definite matrix X_t.

[15] Empirically, our model estimations suggest that $\beta^* < 3$, which effectively reduces the number of parameters that need to be estimated.

because the risk premium of each unspanned risk is proportional to the level of the risk itself. This is the situation emerging, e.g., in Bates (2000)-type models. In cases where matrix ΓQ is not diagonal, the risk premium of the diagonal risk X_{11t} (X_{22t}) is an affine function of both X_{11t} (X_{22t}) and X_{12t}, while the risk premium of the out-of-diagonal risk X_{12t} is an affine function of all unspanned risks. In this case, the compensation for diffusive variance risk in the model can vary in a way partly disconnected from the diffusive variance itself.

B.5. Stochastic Discount Factor

In our model, three types of shocks can be priced: (i) a diffusive shock in index returns, (ii) a diffusive shocks in the $X-$dynamics for the unspanned risks and (iii) a jump-type shock in index returns. According to Assumption 2, these shocks correspond to the scalar Brownian shock dW_t, the matrix-valued Brownian shock dB_t and the compound poisson shock $(e^J - 1)dN_t$, respectively. Given the incompleteness of our model setting, a multiplicity of stochastic discount factors for pricing these shocks exists. Existence of a stochastic discount factor is ensured by a corresponding density for an equivalent change of measure from the physical to the risk neutral probability. Given suitable matrix processes $\{\Gamma_{1t}\}$, $\{\Gamma_{2t}\}$ and our double-exponential specification for log return jumps, such a density can take the form:

$$\begin{aligned}\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} &= \exp\left\{tr\left(-\int_0^T \Gamma_{1t}dW_t^* + \frac{1}{2}\int_0^T \Gamma_{1t}'\Gamma_{1t}dt - \int_0^T \Gamma_{2t}dB_t^* + \frac{1}{2}\int_0^T \Gamma_{2t}'\Gamma_{2t}dt\right)\right\} dt \\ &\quad \times \prod_{i=1}^{N_T^*} \exp\left\{-(\lambda^- - \lambda^{*-})J_i^{*-} - (\lambda^+ - \lambda^{*+})J_i^{*+} + \ln\left(\frac{1/\lambda^{*-} + 1/\lambda^{*+}}{1/\lambda^- + 1/\lambda^+}\right)\right\},\end{aligned}$$

where the second line defines the change of measure for return jumps. This choice implies a double exponential distribution with density (9) for return jumps, having parameters λ^{*+}, λ^{*-} and λ^+, λ^- under the physical and the risk neutral distribution, respectively. Our implicit specification of Γ_{2t} in Assumption 5 is:

$$\Gamma_{2t} = \sqrt{X_t}\Gamma + \frac{1}{2\sqrt{X_t}}(\beta^* - \beta)Q' \ . \tag{17}$$

Together with Assumption 1, this implies a well-defined change of probability measure for pricing the B^*-shocks driving the unspanned risks in our model. It is straightforward to introduce a well-defined concrete specification of the market price of risk Γ_{1t} for $W-$shocks. For instance, under Assumptions 1 and 5, the specification

$$\Gamma_{1t} = \sqrt{X_t}\Delta + \frac{\mu_0 - (r - q)}{\sqrt{X_t}} \ , \tag{18}$$

where Δ is a 2×2 parameter matrix and $\mu_0 > r - q$ a scalar parameter, implies a change of measure with affine dynamics for index returns under the physical and the risk-neutral probability measures.[16] However, it is important to realize that this last assumption is not necessary for the validity of our identification of unspanned risks and the price of the smile, which is robust with respect to the form of the market price of risk Γ_{1t} defining the underlying density process.

B.6. Term Structure of Variance Risk Premia

We characterize the risk premia of variance swap contracts that can be synthetized by a dynamically delta hedged static option portfolio, consistently with the definition of the CBOE (2009) VIX index. As shown in Neuberger (1994), among others, the floating leg $RV_{t+\tau}(\tau)$ of these contracts is proportional to the delta-hedged payoff of a log contract:

$$
\begin{aligned}
RV_{t+\tau}(\tau) \ &:= \ \frac{2}{\tau}\left[-\ln(S_{t+\tau}/S_t) + \int_t^{t+\tau} dS_s/S_{s-}\right] && (19)\\
&= \ \frac{1}{\tau}\int_t^{t+\tau} \frac{1}{S_s^2} d[S,S]_s^c + \frac{2}{\tau}\sum_{t\leq s\leq t+\tau} \mathcal{E}(S_s/S_{s-}) \ , && (20)
\end{aligned}
$$

where $[S,S]_s^c$ is the continuous index quadratic variation and $\mathcal{E}(S_s/S_{s-}) := -\ln(S_s/S_{s-}) + S_s/S_{s-} - 1$ the Itakura-Saito divergence of a jump in index returns at time s.[17] Since the variance risk premium $VRP_t(\tau)$ is the difference of the $\mathbb{P}$ and $\mathbb{Q}$ expectations of $RV_{t+\tau}(\tau)$, Assumptions 1 and 5 give:

$$
VRP_t(\tau) = (E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{1}{\tau}\int_t^{t+\tau} tr(X_s)ds\right] + (E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{2}{\tau}\sum_{t\leq s\leq t+\tau} \mathcal{E}\left(\frac{S_s}{S_{s-}}\right)\right] \ .
$$

The first term on the right hand side is the diffusive variance risk premium, i.e., the premium contribution from continuous index shocks. The second term is the jump variance risk premium, i.e., the premium contribution from index return jumps. Under our previous assumptions, the first term is affine in state X and is by construction spanned by the term structure of the price of the smile. The second term is not fully specified, because we did not yet specify the structure of the jump intensity and the expected Itakura-Saito divergence under the physical measure. These specifications need to consider the challenges

[16] Section **I.** *C.* of the Online Appendix provides a formal proof of the fact that under Assumptions 1 and 5 the market price of risk specifications (17) and (18) imply a well defined density process $\{\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_T}\}_{T\geq 0}$ that is a martingale.

[17] See Schneider and Trojani (2014a) for details.

in identifying the term structure of variance risk premia from a time series of unspanned risks X_{11}, X_{12}, X_{22} and from the realized payoffs of variance swaps for different maturities τ. An affine specification of the physical intensity λ_t^* allows for a simple linear identification based on standard regression methods and requires the estimation of five additional parameters in the most general case, of which only four are identified.[18] In our empirical analysis, we investigated different affine specifications and found that a specification with proportional physical and risk-neutral intensities is preferred by the out-of-sample analysis of variance swap payoffs. Given the indistinguishability of λ_t^* and $E^{\mathbb{P}}[\mathcal{E}(S_t/S_{t-})]$ from the term structure of variance risk premia, we assume identical physical and risk-neutral intensities for identification. This identification assumption is consistent with the stochastic discount factor specification detailed in Section *B.5.*

Assumption 7 *Physical and risk-neutral intensities are identical:* $\lambda_t^* := \lambda_t = \lambda_0 + tr(\Lambda X_t)$.

For later reference, we denote by $\beta_\Lambda^* = \mathbb{E}^{\mathbb{P}}[\mathcal{E}(1+k)]/\mathbb{E}^{\mathbb{Q}}[\mathcal{E}(1+k)]$ the ratio of the physical and the risk neutral expected Itakura-Saito divergence of index return jumps. With this notation, we have under Assumption 7:

$$(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}}) \left[\frac{2}{\tau} \sum_{t \le s \le t+\tau} \mathcal{E}\left(\frac{S_s}{S_{s-}}\right) \right] = E^{\mathbb{Q}}[\mathcal{E}(1+k)](\beta_\Lambda^* E_t^{\mathbb{P}} - E_t^{\mathbb{Q}}) \left[\frac{2}{\tau} \int_t^{t+\tau} tr(\Lambda X_s) ds \right] .$$

Therefore, the jump variance risk premium is affine in the unspanned risks X_{11}, X_{12} and X_{22}.

Proposition 1 *Given Assumptions 1, 2, 5 and 7, the variance risk premium for time to maturity* $\tau > 0$ *is given by:*

$$VRP_t(\tau) = VRP_t^c(\tau) + VRP_t^d(\tau) \ , \tag{21}$$

where diffusive and jump variance risk premia $VRP_t^c(\tau)$ *and* $VRP_t^d(\tau)$ *read explicitly:*

$$\begin{aligned} VRP_t^c(\tau) &= tr[X_\infty^{\mathbb{P}} - X_\infty^{\mathbb{Q}} + A_\tau^{\mathbb{P}}(X_t - X_\infty^{\mathbb{P}}) - A_\tau^{\mathbb{Q}}(X_t - X_\infty^{\mathbb{Q}})], \\ VRP_t^d(\tau) &= 2E^{\mathbb{Q}}[\mathcal{E}(1+k)] tr[\Lambda(\beta_\Lambda^* X_\infty^{\mathbb{P}} - X_\infty^{\mathbb{Q}} + \beta_\Lambda^* A_\tau^{\mathbb{P}}(X_t - X_\infty^{\mathbb{P}}) - A_\tau^{\mathbb{Q}}(X_t - X_\infty^{\mathbb{Q}}))] \ , \end{aligned}$$

with 2×2 *matrices* $X_\infty^{\mathbb{Q}}$, $X_\infty^{\mathbb{P}}$ *such that:*

$$\beta Q'Q = X_\infty^{\mathbb{Q}} M + M' X_\infty^{\mathbb{Q}} \ ; \ \beta^* Q'Q = X_\infty^{\mathbb{P}} M^* + M^{*\prime} X_\infty^{\mathbb{P}} \ , \tag{22}$$

[18] An affine specification of the form $\lambda_t^* = \lambda_0^* + tr(\Lambda^* X_t)$, where Λ^* is a symmetric 2×2 matrix, implies four additional parameters. As expected jump realized variance under the physical measure depends on the product of $E_t^{\mathbb{P}}[\int_t^{t+\tau} \lambda_s^* ds]$ and the (constant) expected Itakura-Saito divergence $E^{\mathbb{P}}[\mathcal{E}(S_t/S_{t-})]$, only four parameters are identifiable.

*and linear matrix operators $A_\tau^{\mathbb{P}}(\cdot)$ and $A_\tau^{\mathbb{Q}}(\cdot)$ defined explicitly in Section **I.**B. of the Online Appendix.*

Remark 8 *(i) With the exception of parameter β_Λ^*, all states and parameters in Proposition 1 are identifiable from the information in the panel of index option prices. (ii) It is easy to see that when matrices Q, M and M^* are diagonal, $VRP_t^c(\tau)$ depends only on unspanned risks X_{11} and X_{22}. If additionally matrix Λ is diagonal, then $VRP_t^d(\tau)$ only depends on X_{11} and X_{22}, inducing a perfect correlation between shocks in variance risk premia and the spot (diffusive) variance. This situation emerges in Bates (2000) and similar models. In non-diagonal models, the term structure of variance risk premia depends also on shocks to risk X_{12}, which are partially separated from shocks to the spot volatility.*

It is important to realize that jump variance risk premia reflect two types of risk, future jump intensity risk and future pure jump variance risk. This naturally motivates the decomposition of $VRP_t^d(\tau)$ into the sum of a jump intensity and a pure jump variance risk premium $VRP_t^{dc}(\tau)$ and $VRP_t^{dj}(\tau)$, respectively:[19]

$$VRP_t^{dc}(\tau) = (E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\int_t^{t+\tau} tr(\Lambda X_s)ds\right] \cdot E^{\mathbb{Q}}[\mathcal{E}(1+k)] \ , \tag{23}$$

$$VRP_t^{dj}(\tau) = E_t^{\mathbb{P}}\left[\int_t^{t+\tau} tr(\Lambda X_s)ds\right] \cdot \left(E^{\mathbb{P}}[\mathcal{E}(1+k)] - E^{\mathbb{Q}}[\mathcal{E}(1+k)]\right) \ . \tag{24}$$

Under our assumptions, both components of the premium are affine in state X. However, while $VRP_t^{dc}(\tau)$ is by construction spanned by the term structure of the price of the smile, $VRP_t^{dj}(\tau)$ is spanned by the term structure of expected average future jump intensities. This distinction is key to understand the relation between the price of the smile and the term structure of variance risk premia.

B.7. Model-Free Variance Swap Payoffs and Variance Risk Premia

Denoting by F_t the S&P 500 index future for maturity $\tau \geq t$, the payoff of a variance swap with maturity τ is the following delta-hedged payoff of a static option portfolio:

$$\begin{aligned} RV_{t+\tau}^e(\tau) := RV_{t+\tau}(\tau) - E_t^{\mathbb{Q}}[RV_{t+\tau}(\tau)] &= \frac{2}{\tau}\left[\int_0^\infty \frac{O_{t+\tau}(K)}{K^2}dK + \int_t^{t+\tau}\left(\frac{1}{F_{s-}} - \frac{1}{F_t}\right)dF_s\right] \\ &\quad -\frac{2}{\tau}\int_0^\infty \frac{E_t^{\mathbb{Q}}[O_{t+\tau}(K)]}{K^2}dK \ , \end{aligned} \tag{25}$$

[19]The second component of the premium is zero if and only if jump variance risk is not priced by the stochastic discount factor of Section II.*B.5.*

where for $K < F_t$ ($K \geq F_t$) payoff $O_{t+\tau}(K) := (K - S_{t+\tau})^+$ ($O_{t+\tau}(K) := (S_{t+\tau} - K)^+$) is the payoff of an out-of-the-money European put (call) option on index futures, with maturity τ and strike price K. We compute $RV^e_{t+\tau}(\tau)$ in a model-free way, using the panel of SPX options and the time-series of high-frequency S&P 500 index futures prices. This motivates the second step of our identification procedure for the term structure of variance risk premia, in which we estimate parameter β^*_Λ from an arbitrage-free linear regression of payoff $RV^e_{t+\tau}(\tau)$ on unspanned risks X_{11}, X_{12}, X_{22}. The explicit form of this regression is directly motivated by Proposition 1 and is detailed in the next result.

Proposition 2 *For any $\tau > 0$, define the following variables:*

$$
\begin{aligned}
Y_{t+\tau}(\tau) &:= RV^e_{t+\tau}(\tau) - VRP^c_t(\tau) - 2E^{\mathbb{Q}}[\mathcal{E}(1+k)]tr[\Lambda(X^{\mathbb{Q}}_\infty + A^{\mathbb{Q}}_\tau(X_t - X^{\mathbb{Q}}_\infty))] \ , \quad (26)\\
U_t(\tau) &:= 2E^{\mathbb{Q}}[\mathcal{E}(1+k)]tr[\Lambda(X^{\mathbb{P}}_\infty + A^{\mathbb{P}}_\tau(X_t - X^{\mathbb{P}}_\infty))] \ . \quad (27)
\end{aligned}
$$

Given Assumptions 2-7 and maturities $\tau_1 < \ldots < \tau_n$, the following is an arbitrage-free linear regression model,

$$
\begin{pmatrix} Y_{t+\tau_1}(\tau_1) \\ \vdots \\ Y_{t+\tau_n}(\tau_n) \end{pmatrix} = \beta^*_\Lambda \begin{pmatrix} U_t(\tau_1) \\ \vdots \\ U_t(\tau_n) \end{pmatrix} + \begin{pmatrix} \eta_{t+\tau_1}(\tau_1) \\ \vdots \\ \eta_{t+\tau_n}(\tau_n) \end{pmatrix} \ , \quad (28)
$$

where error term $\eta_{t+\tau}(\tau) := (\eta_{t+\tau_1}(\tau_1), \ldots, \eta_{t+\tau_n}(\tau_n))'$ is such that $E^{\mathbb{P}}_t[\eta_{t+\tau}(\tau)] = 0$.

Note that all quantities in the definition of $Y_{t+\tau}(\tau)$ and $U_t(\tau)$ above are computable from the synthetic variance swap payoffs $RV^e_{t+\tau}(\tau)$ and a first-step estimation that estimates parameters $\beta, \beta^*, M, M^*, Q, \lambda_0, \Lambda, \lambda^+, \lambda^-$ and filtered states $\{\hat{X}_t\}$ from the panel of SPX options. This insight allows us to separate the estimation of parameter β^*_Λ from the estimation of all other parameters in the model, using our two-step identification procedure for the price of the smile and the term structure of variance risk premia.

II. Empirical Analysis

A. Data and Estimation

We collect from OptionMetrics daily data of end-of-day prices of S&P 500 index options (SPX), traded at the Chicago Board Options Exchange, for the sample period from January 1996 to January 2013 and maturities up to one year.[20] The sample consists of 4298 trading

[20]We obtain end-of-day midquotes as simple averages of end-of-day bid and ask call prices and force the put-call parity to hold when calculating the implied dividend yields.

days, which we reduce to 883 weekly observations (each Wednesday). In order to allow for an out-of sample evaluation of our model, we further split these 883 observations into an in-sample period (from January 1996 to December 2002) with 359 observations and an out-of sample period (from January 2003 to January 2013) with 524 observations. We apply a number of standard filtering procedures outlined, e.g., in Bakshi, Cao and Chen (1997).

For our first-step estimation of model parameters and unspanned risks in option markets, we make use of all options with an absolute Black-Scholes delta between 0.1 and 0.9. On average, this gives about 138 option prices per trading day, having an average time to maturity of 130 days and an average moneyness $S/K = 0.99$. Table 2 of the Online Appendix presents a summary of the main characteristics of our option data set. For the calculation of the model-free variance payoffs in equation (25), we make use of all available options. The delta hedging component in the variance payoff is computed using tick-by-tick data for the S&P 500 index future traded at the CBOE, obtained from tickdata.com and sampled at 60 second intervals.

In the first step of the estimation, we use the panel of SPX in-sample observations to estimate the structural model parameters, together with the time series of unspanned risks X_{11t}, X_{22t} and X_{12t}. The time series of these states uncovers their distinct roles as drivers of tradeable unspanned risks in option markets. The parameter estimates shed light on the dynamic interactions between unspanned riks, as well as their relation to the price of the smile. We estimate the model parameters Q, M, M^*, R, λ_0, Λ, λ^+, λ^-, β, β^*, by maximizing the likelihood defined on the option-implied volatility forecasting errors in a Kalman filter. For identification, we require matrices M, M^*, Λ, R and Q to be triangular, giving a total of 20 parameters to estimate. We borrow from Bates (2000) and conveniently discretize the matrix transition dynamics for state process X, accounting for the variability of conditional first and second moments. For the observation equation, we assume Gaussian errors and account for a potential autocorrelation of option pricing errors. Details on the estimation procedure are provided in Section **II.** of the Online Appendix.

In the second step, we estimate the term structure of variance risk premia, by estimating parameter β^*_Λ in Assumption 7 in a set of simple arbitrage-free linear regressions of realized variance swap payoffs on the time series of model-implied variance risk premia. Precisely, we first compute synthetic variance swap payoffs for maturities $\tau_1, \tau_2, \ldots, \tau_n = 1, 2, 3, 4, 5, 6, 9, 12$ months and construct a time series of in-sample weekly observations for variables $Z_t := (Y_{t+\tau_i}(\tau_i), U_t(\tau_i))_{i=1,\ldots,n}$ in linear model (28), where $t = 1, \ldots, N$ and the in-sample sample size is $N = 359$. We then estimate the single unknown parameter β^*_Λ in equation (28) with a pooled linear regression.

B. Option Pricing Performance and Model Fit

We quantify the option pricing performance and the statistical fit of our model (SVJ_{31}), in relation to the benchmark models in Table 1 of the Online Appendix. Since these models are linked to different degrees of parametrization and to state spaces of different dimensions, we control for overfitting using our in-sample (from January 1996 to December 2002) and out-of-sample (from January 2003 to January 2013) periods. Beyond good in-sample pricing performance and fit, we require that higher dimensional models achieve a stable pricing performance and statistical fit out-of-sample. The out-of-sample period includes phases of very low volatility and benign markets, such as the conundrum, as well as periods of very high volatility and market turmoil, such as the recent financial crisis and the US downgrade. It therefore represents a reasonably challenging benchmark. We estimate all models using only in-sample weekly data and compute proxies of pricing accuracy, such as the weekly absolute average implied volatility error, by computing option implied volatility pricing errors for each week of our in- and out-of-sample periods, using the filtered states implied by the in-sample parameter estimates. Finally, we compare the statistical fit of different models using the in- and the out-of-sample value of the average likelihood function, evaluated at the in-sample parameter estimates. Table 3 of the Online Appendix summarizes the resulting pricing performance and statistical fit across models.

The results indicate that our model produces the best pricing performance and statistical fit, both in- and out-of-sample. For instance, the pricing error is substantially reduced relative to a Bates (2000)-type model (SJV_{20}), by to about 19.4% in-sample and 29.2% out-of-sample, using the $MAIVE$ (mean absolute implied volatility) metric. The improvement of the in-sample (out-of-sample) value of the likelihood function is 4.5% (9.7%) and is statistically significant at conventional significance levels.[21] Our model also implies the smallest deteriorations in out-of sample performance. The out-of-sample $MAIVE$ is only 6.1% higher than the in-sample $MAIVE$. In contrast, in the SVJ_{20} (SVJ_{30}) models, the out-of-sample $MAIVE$ are 20.8% (10.7%) higher, respectively. Similarly, while the average out-of-sample likelihood in model SVJ_{31} is only 6.1% lower than the in-sample likelihood, the out-of-sample likelihood of the SVJ_{20} (SVJ_{30}) model is 11.1% (6.7%) lower. In summary, the improvements of our model are not due to overfitting, as the model's performance is quite similar in- and out-of-sample.

[21] Our model (SVJ_{31}) also improves with respect to a three-factor Bates (2000)-type model (SVJ_{30}), despite having four parameters less in its specification of the physical and risk-neutral dynamics. The improvement in pricing performance is about 3% in-sample and 5% out-of-sample, with respect to the $MAIVE$ metric. Our more parsimonious model SVJ_{31} also attains a higher average likelihood in- and out-of-sample, with improvements of 2.1% in-sample and 1.6% out-of-sample.

Table 4 of the Online Appendix compares $MAIVE$ pricing errors of our model and the benchmark Bates (2000) model (SVJ_{20}), across different moneyness and maturity bins (in days). It shows that our model especially improves on the modeling of out-of-the money options of maturities of 30 days or higher. To illustrate, the in-sample (out-of-sample) $MAIVE$ of model SVJ_{31} for maturities $\tau < 30$, $30 \leq \tau < 75$, $75 \leq \tau < 180$ and $\tau \geq 180$ is 10.0% (17.0%), 25.2% (31.5%), 4.1% (18.1%), 22.9% (33.5%) and 19.1% (27.1%) lower, respectively, than for model SVJ_{20}. Similarly, the in-sample (out-of-sample) $MAIVE$ of model SVJ_{31} for option deltas $|\Delta| < 0.2$, $0.2 \leq |\Delta| < 0.4$, $0.4 \leq |\Delta| < 0.6$, $0.6 \leq |\Delta| < 0.8$ and $|\Delta| \geq 0.8$ is 13.3% (29.8%), 17.7% (24.4%), 9.8% (13.6%), 19.2% (25.2%) and 28.0% (33.7%) lower, respectively. In summary, this evidence shows that our model clearly improves on the specification of the option-implied volatility smile of benchmark models.

C. Mutually Exciting Unspanned Risks

The times series of unspanned risks X_{11}, X_{22} and X_{12} are presented in Figure 2 and imply half-lives (volatilities) of 1.275, 0.277 and 0.108 years (0.009, 0.010 and 0.026), respectively, see also Table 6 of the Online Appendix.

[Insert Figure 2 about here.]

The diffusive variance $tr(X_t)$ is thus decomposed into two positive non-Markovian components with significantly different persistences and volatilities of volatility, where X_{22} has on average a larger contribution to the diffusive variance than X_{11}, besides being more volatile and less persistent. Unspanned risk X_{12} is positive most of the time and takes as X_{22} the largest values in periods of significant turmoil or distress, as during the recent financial crisis and around the US downgrade. Interestingly, the persistent unspanned risk X_{11} also spikes substantially in periods of distress, but often with a lag or a lead with respect to risks X_{12} and X_{22}. Such large variations in the relative importance of unspanned risks X_{11}, X_{22} and X_{12} have important implications for the structure of both the jump and diffusive variances. Panel B of Figure 2 quantifies this aspect by scaling X_{11t} and X_{12t} with the diffusive variance $tr(X_t) = X_{11t} + X_{22t}$. It appears that the fraction of persistent diffusive variance generated by X_{11} can vary essentially from zero to one. The top plot of Panel B in Figure 2 shows that this fraction is basically zero during the whole conundrum period, it is about one shortly before the collapse of the NASDAQ bubble and it rapidly increases from about 0.2 to 0.9 shortly after the US downgrade. The middle plot of Panel B shows that the relative importance of X_{12} also varies a lot. X_{12t} can be as large as 50% of $tr(X_t)$ during phases of market turnoil, e.g, shortly after the devaluation of the Thai Bhat, the beginning of the Russian

crisis, the Lehman default and the US downgrade.[22] Note that while X_{12} is not present in the diffusive variance, it has a significant contribution to the jump intensity. Therefore, it is directly related to the fraction of pure jump variance in the total variance of returns. Given the predominant role of X_{12} in periods of market distress, this risk is a key driver of the jump variance dynamics in these periods.

The time series of unspanned risks in option markets uncovers the dynamic properties of these state variables. In contrast, the estimated model parameters directly capture the dynamic interactions between these risks and their relation with the price of the smile. Table 1 presents the parameter estimates for our model and different benchmark models.

[Insert Table 1 about here.]

All parameters are significant, with the exception of constant λ_0 in the intensity process. Since we cannot reject the null hypothesis $\beta^* = \beta$, the data support a completely affine specification of the market price of risk in our matrix AJD setting. In all parameter matrices M, M^*, Λ and R, the out-of-diagonal element is strongly significant, indicating that option prices are better described by a three-factor matrix AJD than by a two-factor diagonal model with independent components. The estimated jump parameters $\lambda^- \ll \lambda^+$ directly reflect the negative risk-neutral skewness of the distribution of log return jumps.

The large negative coefficient M^*_{22} indicates that risk X_{22} has the strongest autonomous mean reversion, which induces the lowest persistence across unspanned risks in our sample. Since $Q_{22} \gg Q_{11}$, risk X_{22} also has the largest local volatility. Due to the positivity of M^*_{12}, X_{22}'s mean reversion is dampened (reinforced) in states where X_{12} is positive (negative). Recalling that X_{12} is positive most of the time, with stronger excursions during phases of market distress, this feature induces a mutually-exciting behaviour of risks X_{22} and X_{12} in such phases. Note that besides driving the high-frequency component of the diffusive variance, X_{22} also creates high-frequency movements in the jump variance, because parameter Λ_{22} is positive and significant. Thus, X_{22} is a high-frequency component of the total variance, featuring mutually exciting dynamics with X_{12} in phases of distress.

The negative coefficient M^*_{11} indicates that risk X_{11} is also mean-reverting, but clearly more persistent and less volatile than X_{22}, as $Q_{22} \gg Q_{11}$. The mean-reversion of X_{11} is dampened in phases of distress, so that overall the total diffusive variance follows a mutually-exciting dynamics with X_{12} in such periods. Positivity of parameter Λ_{11} shows that X_{11} is also a low-frequency component of the jump variance. Thus, X_{11} is a low-frequency component of the total variance, featuring mutually-exciting behaviour with X_{12} in periods of distress.

[22] The ratio $|X_{12t}|/tr(X_t)$ is less than 0.5 by construction, because of the positive definiteness of X_t.

The negative coefficients M^*_{11} and M^*_{22} indicate that risk X_{12} has an autonomous mean-reversion between the one of the high- and low-frequency risks X_{11} and X_{22}. The local mean reversion of X_{12} depends on X_{11} and X_{22} and is asymmetric. It is increased (dampened) in states where X_{12} is negative (positive), making X_{12} more persistent and mutually-exciting in phases of distress. By construction, X_{12} loads on the jump variance, via the jump intensity, but is absent from the diffusive variance. The large loading Λ_{12} indicates that X_{12} is a key state variable for the jump variance in periods of distress. Therefore, it has the interpretation of a risk factor for jump variance risk, or jump skewness risk, featuring mutually-exciting dynamics with X_{11} and X_{22} in phases of turbulences in financial markets.

D. The Price of the Smile

According to our estimation results in Table 1, we cannot reject the null hypothesis $M^*_{22} - M_{22} = 0$, implying that high-frequency risk X_{22} does not directly explain variations of the price of the smile. Since $M^*_{11} - M_{11} < 0$ and $M^*_{12} - M_{12} < 0$, a direct implication is that the instantaneous risk premium for $X_{22}-$shocks is proportional to risk X_{12}, while the risk premium for $X_{11}-$shocks is negative and proportional to X_{11} itself. In contrast, the risk premium for $X_{12}-$shocks is a linear combination of X_{11} and X_{12}, which is unambiguously negative in our sample period and largest in absolute value when risks X_{11} and X_{12} are large. The estimated instantaneous risk premia for $X_{11}-$, $X_{22}-$ and $X_{12}-$shocks read explicitly:[23]

$$\frac{1}{dt}(E^{\mathbb{P}}_t - E^{\mathbb{Q}}_t)[dX_{11t}] = -1.0776 X_{11t} \ , \tag{29}$$

$$\frac{1}{dt}(E^{\mathbb{P}}_t - E^{\mathbb{Q}}_t)[dX_{12t}] = -0.6283 X_{11t} - 0.5388 X_{12t} \ , \tag{30}$$

$$\frac{1}{dt}(E^{\mathbb{P}}_t - E^{\mathbb{Q}}_t)[dX_{22t}] = -1.2566 X_{12t} \ . \tag{31}$$

Given these estimated coefficients, we conclude that the more persistent risks X_{11} and X_{12} are risk premium factors that completely explain the price of the smile. The most persistent risk premium is the one for $X_{11}-$shocks, as intuitively expected. The second most persistent risk premium is the one of $X_{12}-$shocks, while the most transient risk premium is the one for $X_{22}-$shocks. The persistence properties of these risk premia have direct implications for the term structure of the price the smile. This term structure is defined by the risk premia

[23] Recalling that we cannot reject the hypotheses $\beta = \beta^*$ and $M^*_{22} = M_{22}$, we obtain (see Table 1):

$$\frac{1}{dt}(E^{\mathbb{P}}_t - E^{\mathbb{Q}}_t)[dX_t] = \begin{pmatrix} -0.5388 & 0 \\ -0.6283 & 0 \end{pmatrix} X_t + X_t \begin{pmatrix} -0.5388 & -0.6283 \\ 0 & 0 \end{pmatrix} .$$

The risk premium for $X_{12}-$shocks is always negative when X_{12} is positive. Given that $M^*_{11} - M_{11} < 0$, the contribution of X_{12} to its premium can be positive when state X_{12t} is negative.

of swap contracts that have as floating leg the realized unspanned risk $\frac{1}{\tau}\int_t^{t+\tau} X_{ijs}ds$, where $\tau \geq 0$ is the swap time to maturity. Consistently with the variance risk premium literature, the risk premium of these swaps is given by:

$$(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}}) \left[\frac{1}{\tau} \int_t^{t+\tau} X_{ijs}ds \right] \quad .$$

Figure 3 illustrates the time series of the term structure of the price of the smile for two fixed maturities $\tau = 3, 12$ months.

[Insert Figure 3 about here.]

Unspanned risks X_{11} and X_{12} imply an unambiguously negative term structure of the price of the smile. The term structure of the price of the smile for risk X_{22} is also downward sloping most of the time, but it can turn marginally upward sloping occasionally. Consistent with the previous findings, all components of the term structure of the price of the smile are larger in absolute value during phases of distress and reflect the different persistence properties of unspanned risks X_{11}, X_{12} and X_{22}. The slope of these term structures is procyclical, in the sense that it becomes more negative in periods of financial distress, and it is typically increasing (in absolute value) in the persistence of the unspanned risk. This feature is particularly apparent in the term structure of the smile of high-frequency risk X_{22}, which is essentially flat for times to maturity from 3 months on.

Using the model-implied expressions, we can decompose more precisely the term structure of the price of the smile, in terms of the contributions of each unspanned risk; see Table 7 of the Online Appendix for details. As expected from our previous findings, the term structure of the price of the smile is a linear function only of risks X_{11} and X_{12}, which are risk premium factors fully explaining the dynamics of the term structure. The relative contribution of X_{11} and X_{12} to the different components of this term structure is different, however. The term structure of the price of the smile for realized risk $\frac{1}{\tau}\int_0^{\tau} X_{11t}dt$ is a function of the level of X_{11} alone. Therefore, it reflects a low-frequency risk premium that implies a usually downward sloping term structure also at long horizons. The term structure of the price of the smile for realized risk $\frac{1}{\tau}\int_0^{\tau} X_{12t}dt$ is a function of both X_{11} and X_{12}, but it is predominantly influenced at the long end by the low frequency dynamics of X_{11}. Therefore, this term structure is also typically downward sloping at long horizons, even though in a less pronounced way than for realized risk $\frac{1}{\tau}\int_0^{\tau} X_{11t}dt$. Finally, the term structure for realized risk $\frac{1}{\tau}\int_0^{\tau} X_{22t}dt$ is predominantly determined by the less persistent dynamics of X_{12}. This explains its flatness for horizons above three months.

E. *Interpretation of Unspanned Risks Traded in Option Markets*

To gain some economic interpretation for the unspanned risks X_{11}, X_{12}, X_{22}, it is useful to link them to directly observable characteristics of the option-implied volatility smile. Given that the state X captures both unspanned risks and the price of the smile, such an approach can help to link the price of particular option strategies to the different components of the price of the smile.

We find the high-frequency risk X_{22} to be closely related to the 30-day at-the-money implied variance, with a weekly correlation of 91% and virtually identical persistence properties. Given their different persistence properties, unspanned risks X_{12} and X_{11} are only weakly correlated with X_{22}, with a weekly correlation of only 59% and 26%, respectively. Therefore, a good fraction of their conditional variation is orthogonal to X_{22}. We find that risk X_{12} is closely related to the 30-days option-implied skew, with a weekly correlation of -89%. Similarly, X_{11} closely targets the long end of the implied volatility curve and has a weekly correlation of 91% with the 12 months at-the-money implied variance. Figure 4 summarizes the properties of the time series of model-implied unspanned risks and compares these to the time series of the one-month implied variance, the one-month implied skew and the 12 month implied variance.

[Insert Figure 4 about here.]

From this evidence, we conclude that the term structure of the price of the smile, which is a linear function of risks X_{11} and X_{12} alone, is directly related to observable proxies of short term option-implied skewness and long-term option-implied volatility, respectively. These proxies contain a substantial component that is unspanned by high-frequency volatility shocks and are naturally related to the price of particular option portfolios, such as risk reversals or calendar spreads, which are designed to jointly profit from changes in short-term option-implied skewness and long-term option-implied volatilities. According to our findings, we can interpret the prices of these portfolios as observable risk premium factors that span the term structure of the price of the smile and are naturally related to the term structure of variance risk premia. From a different perspective, this evidence also stresses the importance of a joint treatment of the price of the smile and the price of variance risk.

F. *Term Structure of Variance Risk Premia*

The point estimate $\beta_\Lambda^* = 0.323$ in Table 1 for the pooled arbitrage-free linear regression of Proposition 2 yields a pure jump variance risk premium in equation (24) that is unambiguously negative, as intuitively expected. Together with the term structure of the price of the smile, this identifies the term structure of variance risk premia.

F.1. Variance Risk Premia

The model-implied variance risk premia for horizons $\tau = 1, 12$ months are plotted in Figure 5, together with their difference, as a proxy for the slope of their term structure.

[Insert Figure 5 about here.]

Variance risk premia are unambiguously negative and highly time-varying. They range from -0.1% to -16% (-0.4% to -11%) squared for horizons of $\tau = 1$ month (12 months) and provide a plausible description for the first conditional moment of variance swap payoffs. Consistent with intuition, the variability of variance swap payoffs around the conditional first moment is state dependent and can be extremely high during periods of distress. In such periods, when the price of option-implied insurance is large, variance risk premia are largest in absolute value, as, e.g., during the Asian and Russian crises in the late nineties, shortly before the collapse of the internet bubble in 2000, shortly after the Lehmann bankruptcy in September 2008 and the US downgrade in August 2011. The slope of the term structure of variance risk premia is most of the time negative, reflecting a more expensive price of option insurance for longer horizons. However, it can also be strongly upward sloping for short periods of time, roughly for 12% of the observations in our sample. The most prominent cases in which we observe an inversion of the term structure arise immediately after both the Lehmann default in September 2008 and the US downgrade in August 2011, when the spread between annualized 12 month and 1 month variance risk premia has been as large as $+5.8\%$ squared and $+2\%$ squared, respectively.

Table 8 of the Online Appendix makes use of the closed-form model expressions to decompose the term structure linearly, into the contribution of unspanned risks X_{11}, X_{12} and X_{22}. It shows that the loading of risk X_{11} (risks X_{12} and X_{22}) on the term structure increases (decreases) monotonically with the horizon in absolute value. These features imply a clearly different role of low- and high-frequency risks for the slope of the term structure. Indeed, while a positive shock in risk X_{11} decreases the slope, a positive shock in either X_{22} or X_{12} increases the slope. For instance, while the loading of risk X_{11} on the slope of the term structure between 12 and 3 months is -0.574, the loadings of X_{22} and X_{12} are 0.121 and 0.325, respectively; see Table 8 of the Online Appendix for details. These findings document the conceptually different impact of a shock in a low-frequency or a high-frequency unspanned risk on the term structure of variance risk premia.

F.2. Premia for Diffusive and Jump Variance Risk

Low- and high-frequency unspanned risks have different implications for different constituents of the variance risk premium, such as the diffusive and jump variance risk premium ($VRP_t^c(\tau)$

and $VRP_t^d(\tau)$) in Proposition 1. This is highlighted by Figure 6, where we plot these components for horizons of 3 months and 12 months, together with their difference as a measure of the slope of their term structure.

[Insert Figure 6 about here.]

Variance risk premia at all horizons are dominated by their jump component, as $VRP_t^d(\tau)$ is always at least 85% (65%) of the total premium for horizon $\tau = 3$ ($\tau = 12$) months. More importantly, the dynamics of the term structures of diffusive and jump variance risk premia are substantially different. The term structure of diffusive variance risk premia is unambiguously downward sloping, a finding that directly follows from the downward sloping term structure of the price of the smile in Section II.*D*.[24] While the term structure of jump variance risk premia is slightly downward sloping most of the time, it is upward sloping for about 28% of the observations in our sample. The absolute differences between 12 months and 3 months jump variance risk premia are typically very small when the term structure is downward sloping, i.e., smaller than for diffusive variance risk premia, but they can be very substantial otherwise. As intuitively expected, the term structure of jump variance risk premia is upward sloping in periods of distress, when short term option insurance is very expensive, e.g., immediately before the Lehmann default, with a spread of about +7% squared between 12 months and 3 months risk premia, or around the US downgrade, with a spread of almost +3% squared between 12 months and 3 months risk premia. The closed-form decompositions in Table 8 of the Online Appendix stress a second characterizing feature of the term structure of jump variance risk premia. Indeed, while the term structure of diffusive variance risk premia is spanned by the term structure of the price of the smile, which is itself a linear function or risks X_{11} and X_{12} alone, the term structure of jump variance risk premia depends on all unspanned risks, including high-frequency risk X_{22}. The upward sloping term structure arises in the case of large positive shocks to risks X_{22} and X_{12} and is therefore a primarily high-frequency phenomenon.

F.3. Premia for Pure Jump Variance Risk

Low- and high-frequency unspanned risks have different implications also for the decomposition of jump variance risk premia into the jump intensity and pure jump variance risk premia in equations (23) and (24), respectively. Similar to the diffusive variance risk premium, $VRP_t^{dc}(\tau)$ is a premium for unexpected future variations in jump intensity and is spanned by

[24]As the diffusive realized variance is the sum of the realized risks associated with X_{11} and X_{22}, the properties of the term structure of diffusive variance risk premia are directly inherited from the properties of the term structure of the price of the smile.

the price of the smile. Instead, $VRP_t^{dj}(\tau)$ is the product of an instantaneous jump variance risk premium $2(E^{\mathbb{P}}[\mathcal{E}(1+k^*)]-E^{\mathbb{Q}}[\mathcal{E}(1+k)])$ and the expected average jump intensity under the physical probability. Average expected intensities depend on all unspanned risks traded in option markets, because the estimated jump intensity process depends on all these risks as well:

$$tr(\Lambda X_t) = 25.67 \cdot X_{11t} + 40.43 \cdot X_{12t} + 15.98 \cdot X_{22t} \ . \tag{32}$$

As a consequence, the pure jump variance risk premium depends in a much more substantial way than all other components of the variance risk premium on high-frequency unspanned risks. These features have natural consequences for the term structure of pure jump variance risk premia. Figure 7 shows that this term structure is upward sloping for about 72% of the observations in our sample, while in all other cases it is virtually flat.

[Insert Figure 7 about here.]

This is intuitive, as in all cases where both X_{12} and X_{22} are small, expected jump intensities at all horizons are small and the term structure is flat. In contrast, when either risk X_{12} or risk X_{22} is unusually large, expected intensities are large at short horizons and much lower at longer horizons, because of the fast mean reversion of these risks. This feature induces a strongly upward sloping term structure of pure jump variance risk premia in phases of distress, because the risk premium for instantaneous jump variance risk is negative: $E^{\mathbb{P}}[\mathcal{E}(1+k)] - E^{\mathbb{Q}}[\mathcal{E}(1+k)] < 0$. The substantially larger loading of the intensity process on risks X_{11} and X_{12} also implies that the upward sloping term structure of pure jump variance risk premia is a typically high-frequency phenomenon.

III. Conclusions and Outlook

Motivated by the joint tradeability of variance and skewness risk, we study the price of variance consistently with the price of the unspanned risks traded in liquid option markets, i.e., the price of the smile. In this way, we preclude arbitrage opportunities between option portfolios that trade distinct characteristics of unspanned risks.

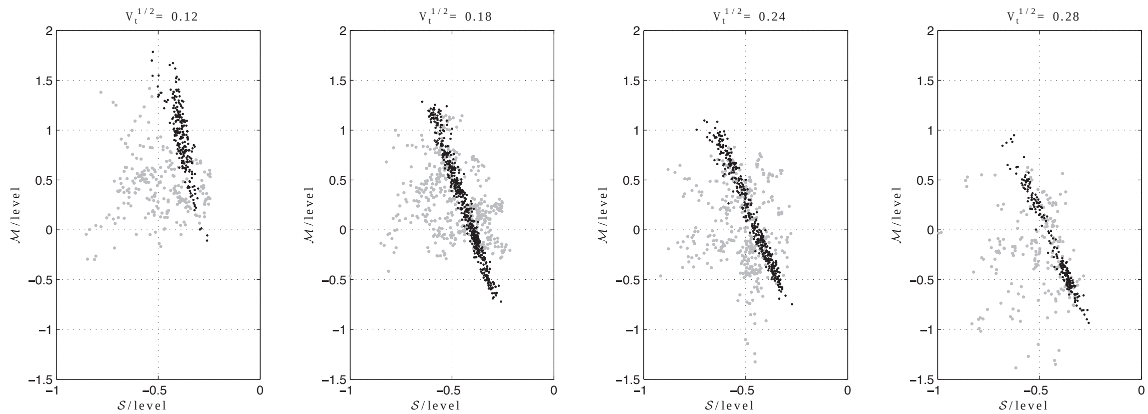
Coherently with the time series and cross-sectional properties of the SPX volatility smile, we identify the price of the smile and the term structure of variance risk premia with a parsimonious three-factor specification of stochastic volatility. In constrast to most existing arbitrage-free models, our specification incorporates three dynamically correlated and differently persistent state variables, which (i) are linked to a stochastic skewness and a price of the smile not spanned by the spot volatility and (ii) produce a sharp identification of the

uspanned risks traded in option markets.

We show that the unspanned risks identified with our approach have a clear interpretation in terms of observable properties of the smile, as they closely follow the one-month implied variance, the one-month implied skew and the one-year implied variance, respectively. Moreover, they induce distinct contributions to the price of the smile and the term structure of variance risk premia. Shocks in the transient component of the volatility generate high frequency variations of jump variance risk and the term structure of variance risk premia. In parallel, shocks in the two more persistent components of the volatility generate lower-frequency varations in the downward sloping term structure of the price of the smile. These features imply a term structure of variance risk premia that is downward sloping most of the time, but which can be strongly upward sloping in periods of market distress, when high-frequency risk escalates.

In a more general perspective, our findings concretely emphasize the importance of a joint specification of the price of second- and higher-order unspanned risks, as we find that the prices of high-frequency volatility risk and jump variance risk are both strongly related to the level of one month option-implied skewness. Such a joint specification is a strong challenge for a large class of reduced-form and structural models, which induce a counterfactually tight link between option-implied volatility and skewness. Our also results highlight the importance of structural equilibrium mechanisms that can explain a dichotomy between high-frequency unspanned risks and their time-varying price.

Panel A: Data vs. Bates (2000) model (SVJ_{20})



Panel B: Data vs. our model (SVJ_{31})

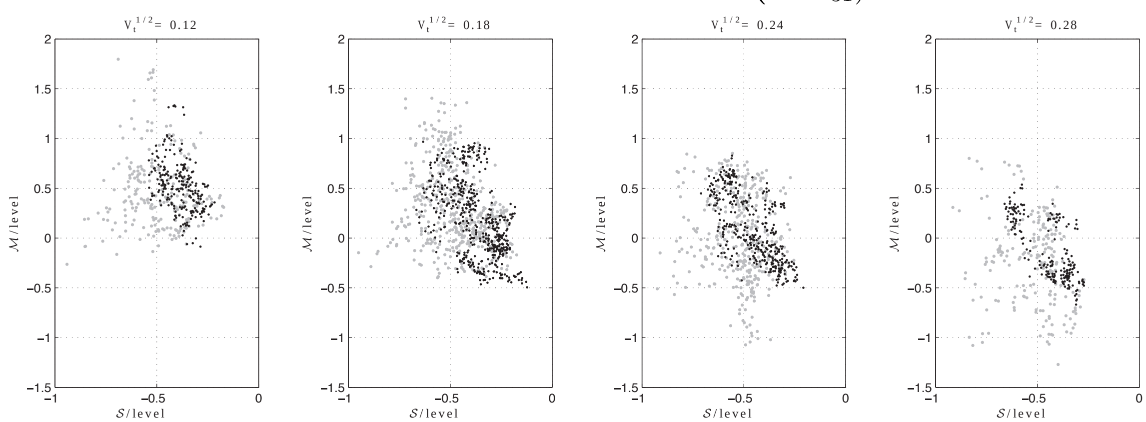
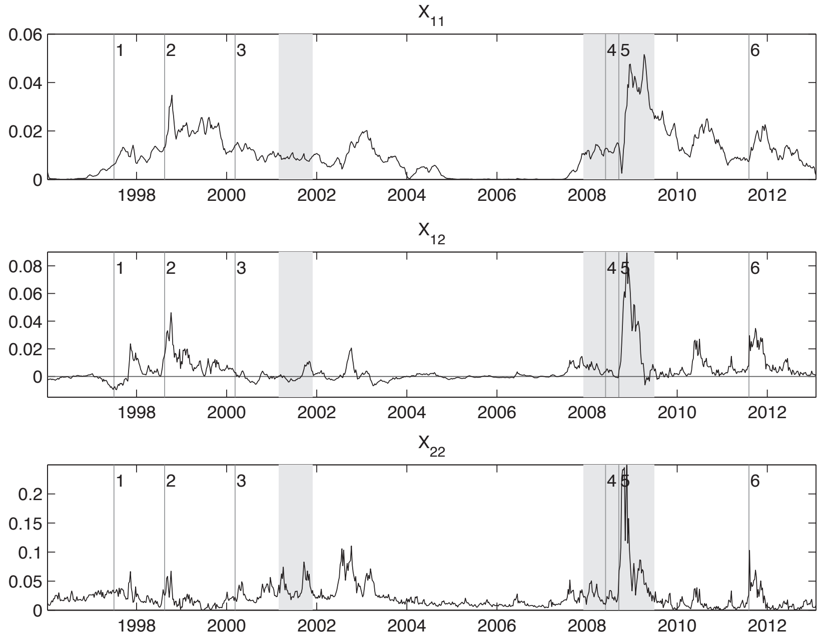


Figure 1: Option-implied skew $\mathcal{S}$ at 1-month horizon versus term structure $\mathcal{M}$, both normalized by the 30-day at the money implied volatility. Every dot corresponds to the volatility surface of one trading day. Grey dots: Data. Black dots – panel A: Fitted values of a two factor Bates model (SVJ_{20}). Black dots – panel B: Fitted values of our model (SVJ_{31}). We stratify by the 30-day at the money implied volatility at $\pm 5\%$ around the selected level, i.e., 17.1%-18.9% for the second panel. The exact calculation method for the option-implied skewness $\mathcal{S}$ and term structure $\mathcal{M}$ is explained in Section II.D. of the Online Appendix.

Panel A: Unspanned risks X_{11}, X_{12} and X_{22}



Panel B: Scaled unspanned risks and diffusive variance $tr(X_t)$

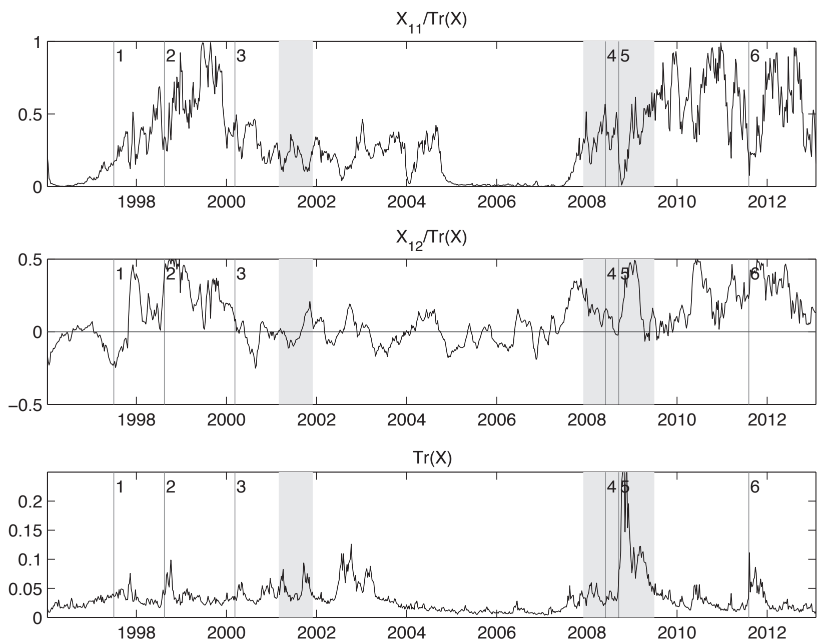


Figure 2: Panel A – Filtered unspanned risks $X_{11t}, X_{12t}, X_{22t}$. Panel B – Scaled unspanned risks $X_{11t}/tr(X_t), X_{12t}/tr(X_t)$ and diffusive variance $tr(X_t) := X_{11t} + X_{12t}$. Grey areas highlight NBER recessions; vertical lines indicate the following crisis events:

(1)	1997-07-02	Start of Asian Crisis	(4)	2008-05-30	Bear Sterns bailout
(2)	1998-08-17	Start of Russian Crisis	(5)	2008-09-15	Lehman bankruptcy
(3)	2000-03-10	NASDAQ maximum	(6)	2011-08-05	US downgrade

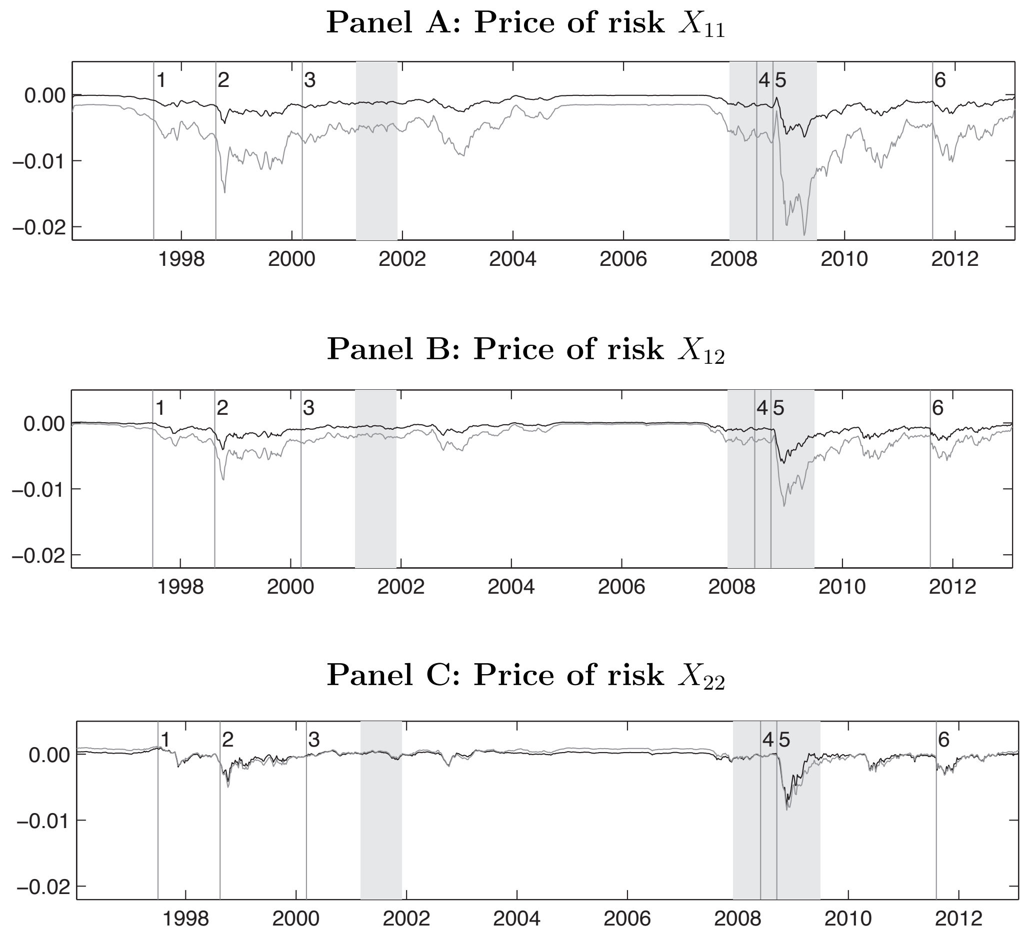
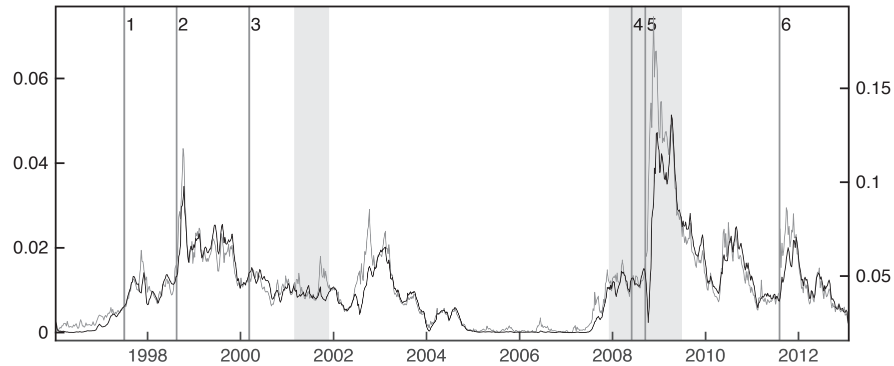
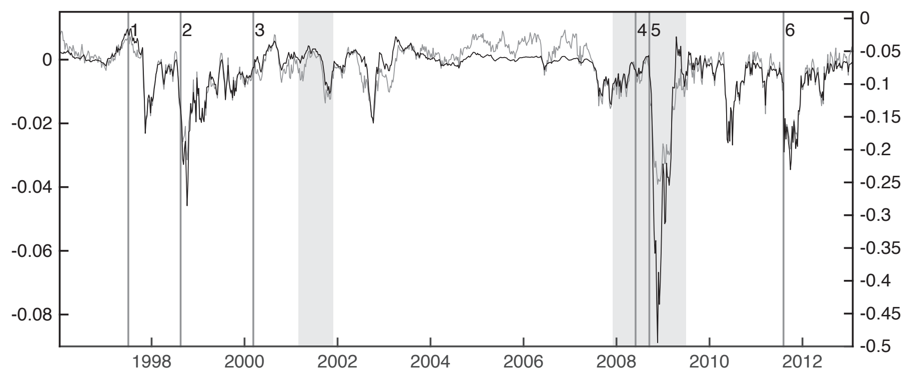


Figure 3: Term structure of the price of the smile. We plot for each unspanned risk X_{11}, X_{12} and X_{22} the corresponding component of the annualized difference $(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right]$ between physical and risk-neutral expectations of the integrated state, for horizons of 3 months (black lines) and 12 months (grey lines). Grey areas highlight NBER recessions; vertical lines indicate important crisis events as listed in the caption of Figure 2.

Panel A: Unspanned risk X_{11} as 12-month at-the-money implied variance



Panel B: Negative value of unspanned risk X_{12} as option-implied skew



Panel C: Unspanned risk X_{22} as 1-month at-the-money implied variance

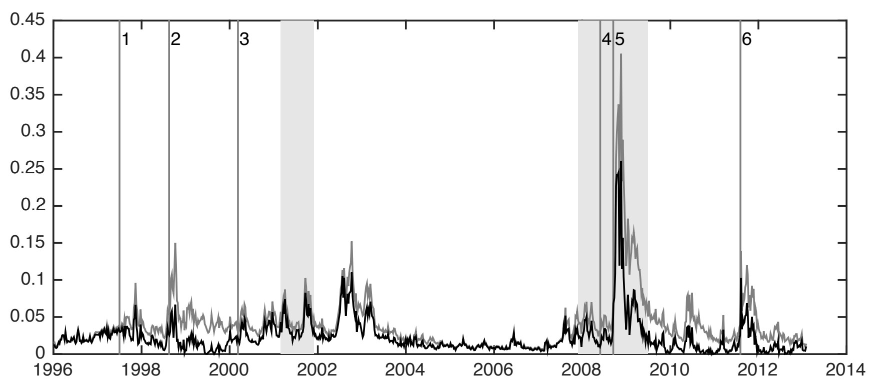
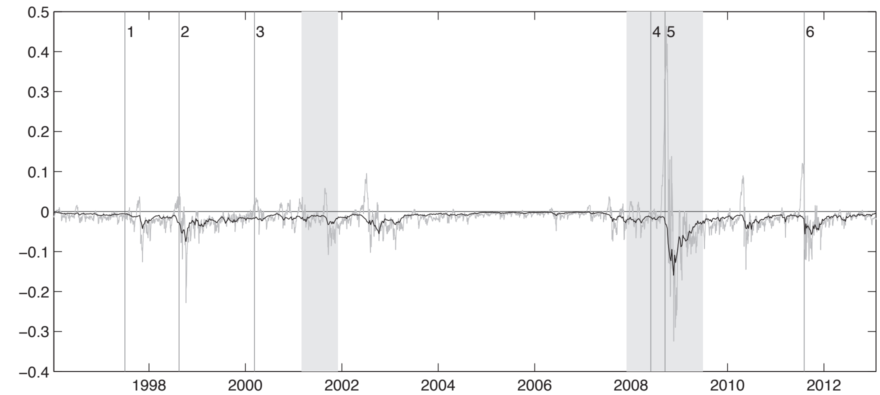
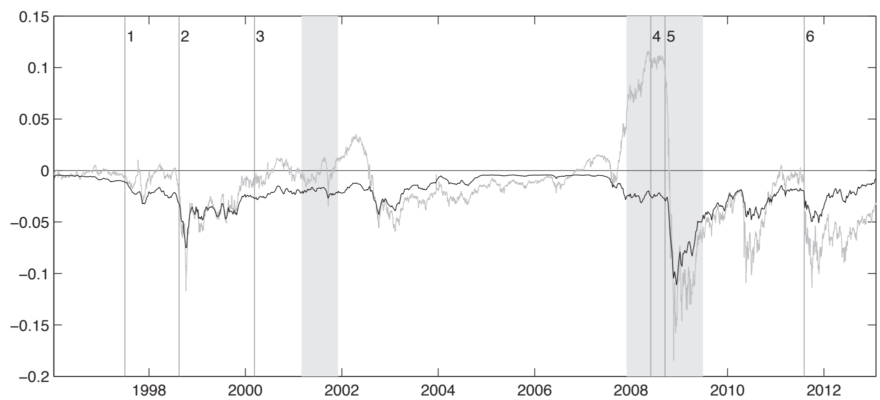


Figure 4: Unspanned risks as observable components of the volatility surface. Panel A: X_{11} (black line, left scale) and the 12-month at-the-money implied variance (grey line, right scale). Panel B: Negative value of X_{12} (black line, left scale) and option-implied skew $\mathcal{S}$ at 1-month horizon (grey line, right scale). Panel C: X_{22} (black line) and one-month at-the-money implied variance (grey line). See Section II.*D.* of the Online Appendix for the calculation method of skew $\mathcal{S}$. Grey areas highlight NBER recessions; vertical lines indicate important crisis events as listed in the caption of Figure 2.

Panel A: 1 month variance risk premium



Panel B: 12 months variance risk premium



Panel C: Term structure of variance risk premia

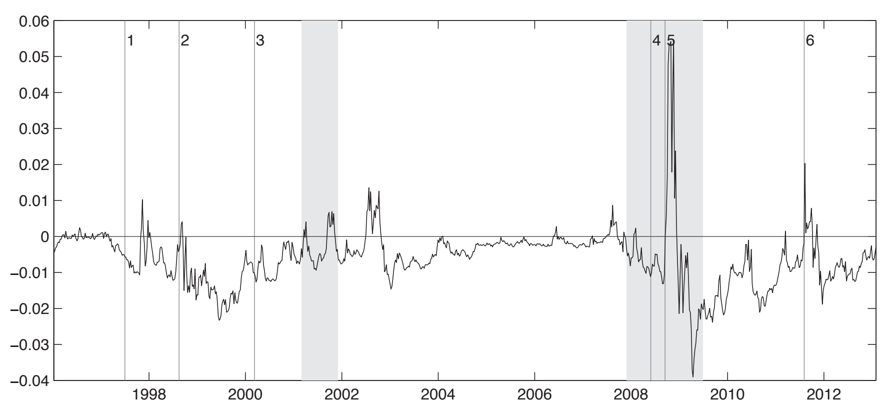
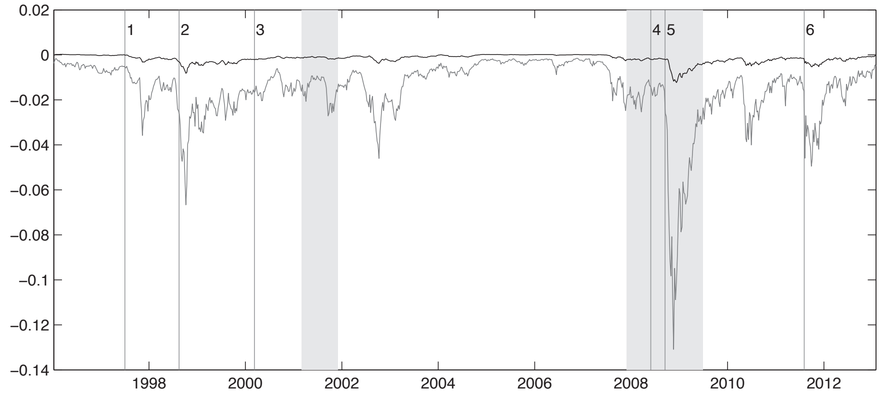
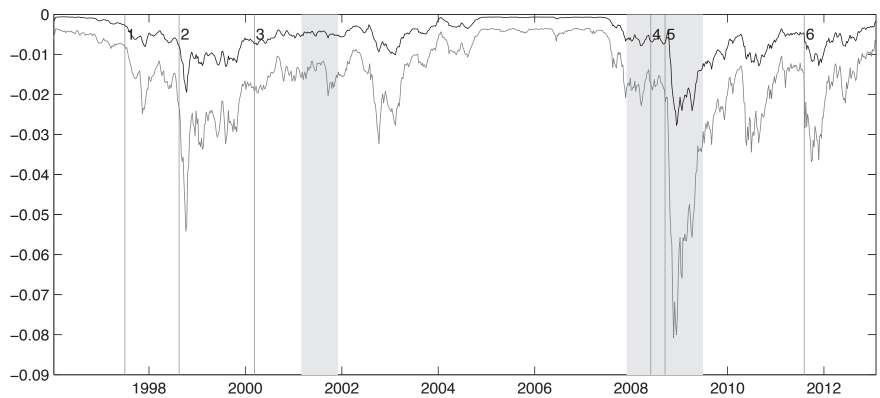


Figure 5: Variance risk premium and slope of the term structure of variance risk premia. In panel A (B), we plot the annualized model-implied 1 month (12 months) variance risk premium (black lines) and the payoffs of synthetic variance swaps (grey lines). In panel C, we plot the slope of the model-implied term structure of variance risk premia, computed as the difference of 12-months and 1-month variance risk premia. Grey areas highlight NBER recessions; vertical lines indicate important crisis events as listed in the caption of Figure 2.

Panel A: 3 months diffusive and jump variance risk premium



Panel B: 12 months diffusive and jump variance risk premium



Panel C: Term structure of diffusive and jump variance risk premia

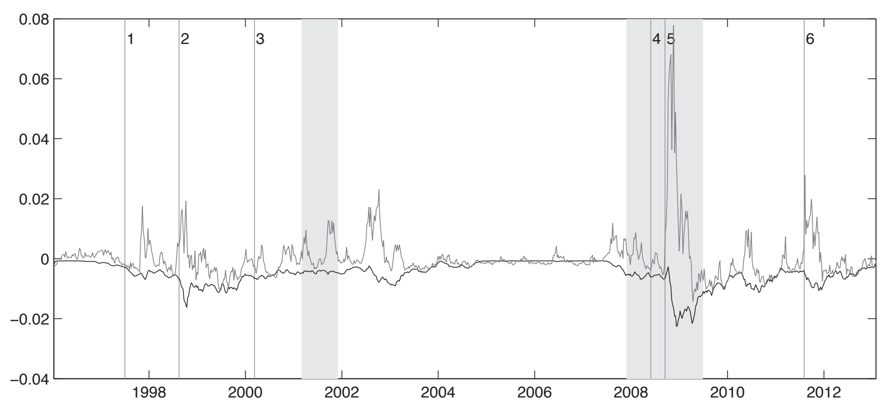
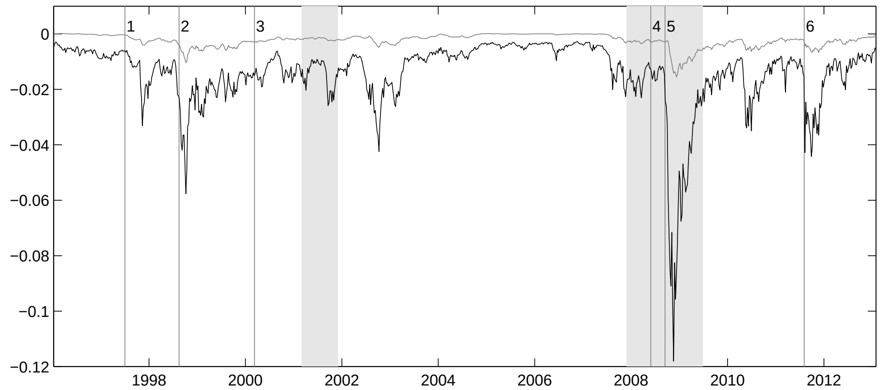
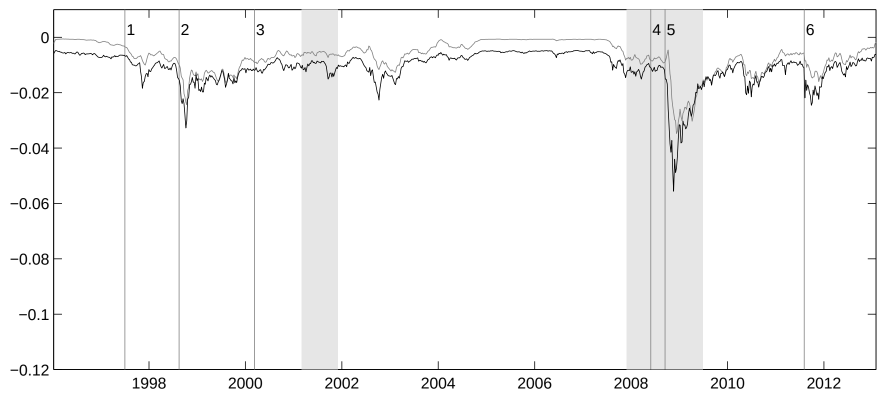


Figure 6: Diffusive and jump variance risk premia $VRP_t^c(\tau)$ (black line) and $VRP_t^d(\tau)$ (grey line). Panel A: 3 months horizon; Panel B: 12 months horizon. In panel C, we plot the slope of the model-implied term structure of diffusive and jump variance risk premia, computed as the difference of 12 months and 3 months risk premia. Grey areas highlight NBER recessions; vertical lines indicate important crisis events as listed in the caption of Figure 2.

Panel A: 3 months intensity and pure jump variance risk premium



Panel B: 12 months intensity and pure jump variance risk premium



Panel C: Term structure of intensity and pure jump variance risk premia

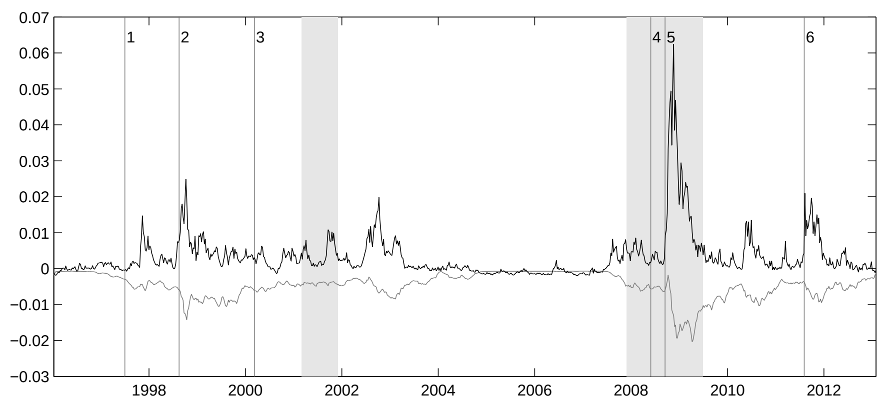


Figure 7: Intensity and pure jump variance risk premia $VRP_t^{dc}(\tau)$ (grey line) and $VRP_t^{dj}(\tau)$ (black line). Panel A: 3 months horizon, Panel B: 12 months horizon. In panel C, we plot the slope of the model-implied term structure of intensity and pure jump variance risk premia, computed as the difference of 12 months and 3 months risk premia. Grey areas highlight NBER recessions; vertical lines indicate important crisis events as listed in the caption of Figure 2.

Panel A: Diffusion parameters

	SV_{20}	SV_{30}	SV_{31}	SVJ_{20}	SVJ_{30}	SVJ_{31}
M_{11}	-0.3121	-0.0844	-1.0716	-0.3242	-0.1231	-0.0079
	(0.0063)	(0.0020)	(0.0185)	(0.0067)	(0.0023)	(0.0002)
M_{22}	-5.0719	-5.4283	-4.9213	-4.4564	-4.2041	-2.6808
	(0.1040)	(0.1254)	(0.0489)	(0.0895)	(0.0582)	(0.0261)
M_{33}		-1.4410			-0.5517	
		(0.0307)			(0.0104)	
M_{21}			14.3050			1.0265
			(0.2173)			(0.0120)
Q_{11}	0.2370	0.1957	0.0556	0.0903	0.0742	0.0698
	(0.0024)	(0.0026)	(0.0006)	(0.0015)	(0.0010)	(0.0009)
Q_{22}	0.4209	0.4498	0.5256	0.4204	0.2853	0.2924
	(0.0057)	(0.0062)	(0.0033)	(0.0054)	(0.0026)	(0.0024)
Q_{33}		0.0718			0.0738	
		(0.0019)			(0.0016)	
Q_{12}			-0.1440			-0.0770
			(0.0021)			(0.0012)
R_{11}	-1.0000	-1.0000	-0.0431	-1.0000	-0.9997	-0.2970
	(0.0131)	(0.0134)	(0.0008)	(0.0227)	(0.0189)	(0.0036)
R_{22}	-0.5348	-1.0000	-0.6405	-0.3823	-0.7111	-0.4057
	(0.0087)	(0.0192)	(0.0055)	(0.0069)	(0.0117)	(0.0048)
R_{33}		0.9633			-0.1178	
		(0.0255)			(0.0026)	
R_{12}			-0.7672			-0.8708
			(0.0110)			(0.0121)
β_{11}	1.0000	1.0031	1.0000	1.0006	1.0064	1.0012
	(0.0160)	(0.0169)	(0.0118)	(0.0191)	(0.0180)	(0.0116)
β_{22}	1.0000	1.0007		1.0000	1.0042	
	(0.0187)	(0.0219)		(0.0197)	(0.0153)	
β_{33}		1.0162			1.0146	
		(0.0235)			(0.0187)	
M^*_{11}	-1.4051	-1.2204	-0.6378	-0.7395	-0.8289	-0.5467
	(0.0266)	(0.0298)	(0.0091)	(0.0172)	(0.0134)	(0.0083)
M^*_{22}	-1.8593	-2.2558	-2.7528	-1.9462	-1.2661	-2.6808
	(0.0401)	(0.0584)	(0.0435)	(0.0477)	(0.0221)	(0.0334)
M^*_{33}		-0.4869			-0.5539	
		(0.0116)			(0.0093)	
M^*_{21}			1.9200			0.3982
			(0.0284)			(0.0051)
β^*_{11}	1.0000	1.0017	1.0000	1.0006	1.0064	1.0012
	(0.0203)	(0.0216)	(0.0162)	(0.0200)	(0.0190)	(0.0124)
β^*_{22}	1.0000	1.0046		1.0000	1.0042	
	(0.0201)	(0.0199)		(0.0251)	(0.0232)	
β^*_{33}		1.0693			1.0146	
		(0.0316)			(0.0208)	

Panel B: Jump parameters

	SVJ_{20}	SVJ_{30}	SVJ_{31}
λ_0	0.0000	0.0003	0.0000
	(0.0003)	(0.0002)	(0.0002)
Λ_{11}	43.8971	57.3248	25.6671
	(0.9240)	(0.9276)	(0.3193)
Λ_{22}	1.0566	11.9429	15.9795
	(0.0265)	(0.1899)	(0.1933)
Λ_{33}		0.0454	
		(0.0008)	
Λ_{12}			40.4278
			(0.6332)
$\bar{k}$	-0.1500	-0.1500	
	(0.0030)	(0.0019)	
δ	0.1500	0.1500	
	(0.0027)	(0.0020)	
λ^-			7.1518
			(0.0372)
λ^+			58.3547
			(0.7690)
β^*_Λ			0.3230
			(0.0553)

Table 1: In-sample (1996/01-2002/12) parameter estimates and standard errors. Panel A: diffusion parameters. Panel B: jump parameters. For consistency and for brevity, all parameter values are reported using a notation based on matrix AJD, i.e., by considering Bates- and Heston-type models as nested diagonal matrix AJD models.

References

Adrian, Tobias and Hyun Song Shin, "Liquidity and Leverage," *Journal of Financial Intermediation*, 2010, *19* (3), 418–437.

—— **and Joshua Rosenberg**, "Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk," *The Journal of Finance*, 2008, *63* (6), 2997–3030.

—— **, Emanuel Moench, and Hyun Song Shin**, "Leverage asset pricing," *Staff Reports 625, Federal Reserve Bank of New York*, 2013.

Ait-Sahalia, Yacine, Mustafa Karaman, and Loriano Mancini, "The Term Structure of Variance Swaps, Risk Premia and the Expectation Hypothesis," *SSRN Working Paper*, 2012.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, "Parametric inference and dynamic state recovery from option panels," *Econometrica, forthcoming*, 2015.

Bakshi, Gurdip and Nikunj Kapadia, "Delta-Hedged Gains and the Negative Market Volatility Risk Premium," *Review of Financial Studies*, 2003, *16(2)*, 527–566.

—— **, Charles Cao, and Zhiwu Chen**, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 1997, *52 (5)*, 2003–2049.

Barras, Laurent and Aytek Malkhozov, "Variance Risk Premium Dynamics in Equity and Option Markets," *Working Paper*, May 2014.

Bates, David, "Post-'87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics*, 2000, *94* (1), 181–238.

Bergomi, Lorenzo, "Smile dynamics," *Risk*, 2004, *17* (9), 117–123.

—— , "Smile dynamics II," *Risk*, 2005, *18* (10), 67–73.

—— , "Smile dynamics III," *Risk*, 2008, *21* (10), 90–96.

—— , "Smile dynamics IV," *Risk*, 2009, *22* (12), 94–100.

Bloom, Nicholas, "The Impact of Uncertainty Shocks," *Econometrica*, 2009, *77* (3), 623–685.

Bollerslev, Tim and Viktor Todorov, "Tails, Fears, and Risk Premia," *The Journal of Finance*, 2011, *66* (6), 2165–2211.

——, **George Tauchen, and Hao Zhou**, "Expected Stock Returns and Variance Risk Premia," *The Review of Financial Studies*, 2009, *22* (11), 4463–4492.

Bondarenko, Oleg, "Variance Trading and Market Price of Variance Risk," *Journal of Econometrics*, 2014.

Britten-Jones, Mark and Anthony Neuberger, "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 2000, *55*, 839–866.

Bru, Marie-France, "Wishart Processes," *Journal of Theoretical Probability*, 1991, *4*, 725–751.

Buraschi, Andrea and Jens C. Jackwerth, "The price of a smile: Hedging and spanning in option markets," *Review of Financial Studies*, 2001, *14*, 495–527.

——, **Paolo Porchia, and Fabio Trojani**, "Correlation Risk and Optimal Portfolio Choice," *Journal of Finance*, 2010, *65* (1), 393–420.

Calvet, Laurent E., Marcus Fearnley, Adlai J. Fisher, and Markus Leippold, "What's Beneath the Surface? Option Pricing with Multifrequency Latent States," *HEC Paris Research Paper No. 969/2013*, 2013.

Carr, Peter and Dilip B. Madan, "Towards a theory of volatility trading," *in R. Jarrow (ed) Risk Book on Volatility. New York: Risk, pp 417-27*, 1998.

—— **and** ——, "Option valuation using the fast Fourier transform," *Journal of Computational Finance*, 1999, *2*, 61–73.

—— **and Liuren Wu**, "Leverage Effect, Volatility Feedback, and Self-Exciting Market Disruptions: Disentangling the Multi-dimensional Variations in S&P500 Index Options," *Bloomberg Portfolio Research Paper No. 2009-03*, 2009.

—— **and** ——, "Variance Risk Premiums," *Review of Financial Studies*, 2009, *22* (3), 1311–1341.

CBOE, "The CBOE volatility index - VIX," *CBOE*, 2009.

Cheridito, Patrick, Damir Filipovic, and Robert L. Kimmel, "Market Price of Risk Specifications for Affine Models: Theory and Evidence," *Journal of Financial Economics*, 2007, *83*, 123–170.

Christoffersen, Peter, Steve Heston, and Kris Jacobs, "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well," *Management Science*, 2009, *55*, 1914–1932.

Constantinides, George M. and Lei Lian, "The Supply and Demand of S&P 500 Put Options," *NBER Working Paper No. 21161*, 2015.

da Fonseca, Jose, Martino Grasselli, and Claudio Tebaldi, "A Multifactor Volatility Heston Model," *Quantitative Finance*, 2008, *8 (6)*, 591–604.

Dai, Qiang and Kenneth D. Singleton, "Specification analysis of affine term structure modes," *Journal of Finance*, 2000, *50*, 1943–1978.

Demeterfi, Kresimir, Emanuel Derman, Michael Kamal, and Joseph Zou, "A Guide to Volatility and Variance Swaps," *The Journal of Derivatives*, 1999, *6(4)*, 9–32.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, "The price of variance risk," *Working Paper November 2014*, 2014.

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

Duffie, Darrell, Jun Pan, and Kenneth Singleton, "Transform analysis and asset pricing for affine jump-diffusions," *Econometrica*, 2000, *68 (6)*, 1343–1376.

Dupire, Bruno, "Model art," *Risk*, 1993, *September*, 118–120.

Fang, Fang and Kees Oosterlee, "A novel pricing method for European Options based on Fourier-Cosine Expansions," *SIAM Journal on Scientific Computing*, 2008, *31(2)*, 826–848.

Filipovic, Damir, Elise Gourier, and Loriano Mancini, "Quadratic Variance Swap Models," *Journal of Financial Economics, forthcoming*, 2015.

Gabaix, Xavier, "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *The Quarterly Journal of Economics*, 2012, *127* (2), pp. 645–700.

Gourieroux, Christian, "Continuous Time Wishart Process for Stochastic Risk," *Econometric Reviews*, 2006, *25*, 177–217.

Gourio, Francois, "Disaster Risk and Business Cycle," *American E*, 2012, *102* (6), 2734–66.

——, "Financial Distress and Endogenous Uncertainty," *Workig Paper*, 2014.

Gruber, Peter H., "Eliciting a Smile: Numerical Methods for Option Pricing with Matrix Affine Jump Diffusions," *Working paper, Universita della Svizzera Italiana, Lugano*, 2015.

——**, Claudio Tebaldi, and Fabio Trojani**, "Three Make a Smile – Dynamic Volatility, Skewness and Term Structure Components in Option Valuation," 2010. Working Paper. University of Lugano and Bocconi University.

Heston, Steven L., "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 1993, *6*, 327–343.

Huang, Jing-Zhi and Liuren Wu, "Specification analysis of option pricing models based on time-changed Levy processes," *Journal of Finance*, 2004, *59*, 1405–39.

Kozhan, Roman, Anthony Neuberger, and Paul Schneider, "Understanding Risk Premia in Index Option Prices," *SSRN Working paper August 12, 2010*, 2010.

Leippold, Markus and Fabio Trojani, "Asset Pricing with Matrix Jump Diffusions," *SSRN eLibrary*, 2008.

Martin, Ian, "Simple Variance Swaps," Working Paper 16884, National Bureau of Economic Research March 2012.

Mayerhofer, Eberhard, "Wishart processes and Wishart Distributios: An affine Process Point of View," *Lecture Notes*, 2014.

Muir, Tyler, "Financial Crises, Risk Premia, and the Term Structure of Risky Assets," *Job Market Paper*, 2013.

Neuberger, Anthony, "The Log Contract," *Journal of Protfolio Management*, 1994, *20* (2), 74–80.

——, "Realized Skewness," *SSRN eLibrary*, 2012.

Schneider, Paul and Fabio Trojani, "Divergence And The Price Of Uncertainty," *Working Paper*, 2014.

—— **and** ——, "Fear Trading," *Working Paper*, 2014.

Storn, Rainer and Kenneth Price, "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous," *Journal of Global Optimization*, 1997, *11*, 341–359.

Todorov, Viktor, "Variance Risk-Premium Dynamics: The Role of Jumps," *Review of Financial Studies*, 2010, *23* (1), 345–383.

Wachter, Jessica A., "Can time-varying risk of rare disasters explain aggregate stock market volatility?," *Journal of Finance*, 2013, *68*, 987–1035.

Wu, Liuren, "Variance dynamics: Joint evidence from options and high-frequency returns," *Journal of Econometrics*, 2011, *160* (1), 280–287.

Supplemental Appendix to:
The Price of the Smile and Variance Risk Premia

PETER H. GRUBER, CLAUDIO TEBALDI, and FABIO TROJANI

This version: November 11, 2015

I. Additional Results in the Matrix AJD Model

A. Pricing Transform in the Matrix AJD Model

Under Assumption 2 and Assumption 7, the closed-form exponentially affine risk-neutral transform for $Y_T := \log(S_T)$ is given by:

$$\Psi(\tau;\gamma) := E_t\left[\exp\left(\gamma Y_T\right)\right] = \exp\Big(\gamma Y_t + tr\big[A(\tau)X_t\big] + B(\tau)\Big)\ , \tag{A-1}$$

where $\tau = T - t$, $A(\tau) = C_{22}(\tau)^{-1}C_{21}(\tau)$ and the 2×2 matrices $C_{ij}(\tau)$ are the $ij-$th blocks of the matrix exponential:

$$\begin{pmatrix} C_{11}(\tau) & C_{12}(\tau) \\ C_{21}(\tau) & C_{22}(\tau) \end{pmatrix} = \exp\left[\tau\begin{pmatrix} M+\gamma Q'R & -2Q'Q \\ C_0(\gamma) & -(M'+\gamma R'Q) \end{pmatrix}\right]. \tag{A-2}$$

The explicit expressions for the 2×2 matrix C_0 is:

$$C_0(\gamma) = \frac{\gamma(\gamma-1)}{2}I_2 + \Lambda\left[\Theta^Y(\gamma) - 1 - \gamma\Theta^Y(1)\right]\ , \tag{A-3}$$

and real-valued function $B(\tau)$ is given by:

$$\begin{aligned} = \ & \tau\Big\{(\gamma-1)r + \lambda_0\big[\Theta^Y(\gamma) - 1 - \gamma\Theta^Y(1)\big]\Big\} \\ & -\frac{\beta}{2}tr[\ln(C_{22}(\tau)) + \tau(M'+\gamma R'Q)] \end{aligned} \tag{A-4}$$

where $\ln(\cdot)$ is the matrix logarithm and $\Theta^Y(\gamma)$ is the univariate Laplace transform of the return jump size distribution. In the case of the double exponential distribution,

$$\Theta^Y_{DX}(\gamma) = \frac{\lambda^+\lambda^-}{\lambda^+\lambda^- + \gamma(\lambda^+ - \lambda^-) - \gamma^2}\ .$$

In the case of the lognormal distribution

$$\Theta^Y_{LN}(\gamma) = (1+\overline{k})^\gamma \exp\left(\gamma(\gamma-1)\frac{\delta^2}{2}\right)\ ,$$

see, e.g., Leippold and Trojani (2008).

B. Variance Risk Premium in the Matrix AJD Model

The affine expression for the variance risk premium in Proposition 1 is obtained by recalling the relations:

$$
\begin{aligned}
VRP_t(\tau) &= tr\left((E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right]\right) + (E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{1}{\tau}\int_t^{t+\tau} (dS_s/S_{s-})^2\right] \\
&= tr\left((E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right]\right) \\
&\quad + E^{\mathbb{Q}}[\mathcal{E}(1+k)] tr\left(\Lambda(\beta_\Lambda^* E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right]\right) .
\end{aligned}
$$

This shows that $VRP_t(\tau)$ is the sum of two-affine functions of state X_t. To compute these functions in closed-form, we need to compute the $\mathbb{P}$ and $\mathbb{Q}$ expectation of the average integrated state X in our model. These expectations are available in closed-form:

$$
E_t^{\mathbb{Q}}\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right] = X_\infty^{\mathbb{Q}} + \frac{1}{\tau}\int_0^\tau e^{Mu}(X_t - X_\infty^{\mathbb{Q}})e^{M'u} du \ , \tag{A-5}
$$

where the long-run mean $X_\infty^{\mathbb{Q}}$ is the unique solution of the Lyapunov equation $MX_\infty^{\mathbb{Q}} + X_\infty^{\mathbb{Q}} M' = \beta Q'Q$. Similarly,

$$
E_t^{\mathbb{P}}\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right] = X_\infty^{\mathbb{P}} + \frac{1}{\tau}\int_0^\tau e^{M^*u}(X_t - X_\infty^{\mathbb{P}})e^{M^{*\prime}u} du \ , \tag{A-6}
$$

where $X_\infty^{\mathbb{P}}$ is such that $M^* X_\infty^{\mathbb{P}} + X_\infty^{\mathbb{P}} M^{*\prime} = \beta^* Q'Q$. This implies, for any 2×2 matrix D:

$$
\begin{aligned}
tr\left(D E_t^{\mathbb{Q}}\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right]\right) &= tr\left(D(X_\infty^{\mathbb{Q}} + A_\tau^{\mathbb{Q}}(X_t - X_\infty^{\mathbb{Q}}))\right) \ , \\
tr\left(D E_t^{\mathbb{P}}\left[\frac{1}{\tau}\int_t^{t+\tau} X_s ds\right]\right) &= tr\left(D(X_\infty^{\mathbb{P}} + A_\tau^{\mathbb{P}}(X_t - X_\infty^{\mathbb{P}}))\right) \ ,
\end{aligned}
$$

where, for any 2×2 matrix H:

$$
A_\tau^{\mathbb{Q}}(H) := \frac{1}{\tau}\int_0^\tau e^{Mu} H e^{M'u} du \ ; \ A_\tau^{\mathbb{P}}(H) := \frac{1}{\tau}\int_0^\tau e^{M^*u} H e^{M^{*\prime}u} du \ .
$$

Since these two functions are linear in H, the variance risk premium is affine in X_t. This conclude the proof. $\square$

C. Stochastic Discount Factor in the Matrix AJD Model

Existence of a well-defined stochastic discount factor to price all shocks in our model is ensured by a proper density for an equivalent change of measure, from the physical to the risk neutral probability. To this end, we specify matrix processes $\{\Gamma_{1t}\}$, $\{\Gamma_{2t}\}$ for the market prices of Brownian shocks dW_t^*, dB_t^*, and an appropriate distribution for return jumps. Following Assumption 2, we specify a double exponential distribution for log return jumps, with parameters $\lambda^{+*}, \lambda^{-*}$ and λ^{+}, λ^{-}, respectively, under the physical and the risk neutral probabilities. We show that, under Assumption 5, a proper density process consistent with these properties is defined for any $T \geq 0$ by:

$$\begin{aligned} \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} &= \exp\left\{tr\left(-\int_0^T \Gamma_{1t} dW_t^* + \frac{1}{2}\int_0^T \Gamma_{1t}'\Gamma_{1t} dt - \int_0^T \Gamma_{2t} dB_t^* + \frac{1}{2}\int_0^T \Gamma_{2t}'\Gamma_{2t} dt\right)\right\} dt \\ &\quad \times \prod_{i=1}^{N_T^*} \exp\left\{-(\lambda^- - \lambda^{*-})J_i^{*-} - (\lambda^+ - \lambda^{*+})J_i^{*+} + \ln\left(\frac{1/\lambda^{*-} + 1/\lambda^{*+}}{1/\lambda^- + 1/\lambda^+}\right)\right\}, \end{aligned} \tag{A-7}$$

where

$$\Gamma_{1t} = \sqrt{X_t}\Gamma + \frac{1}{2\sqrt{X_t}}(\beta^* - \beta)Q' , \tag{A-8}$$

and

$$\Gamma_{2t} = \sqrt{X_t}\Delta + \frac{\mu_0 - (r - q)}{\sqrt{X_t}} , \tag{A-9}$$

with $\mu_0 - (r - q) \geq 0$ and Δ a 2×2 parameter matrix. The first (second) line of equality (A-7) defines a possible change of measure for diffusive (jump) shocks in our model.

Under Assumption 5, the stochastic exponential in the first line of (A-7) is a well-defined positive local martingale, and hence a supermartingale. Therefore, to show that this term is a martingale, it is enough to show that it has a constant expectation:

$$1 = E_0^{\mathbb{P}}\left[\exp\left\{tr\left(-\int_0^T \Gamma_{1t} dW_t^* + \frac{1}{2}\int_0^T \Gamma_{1t}'\Gamma_{1t} dt - \int_0^T \Gamma_{2t} dB_t^* + \frac{1}{2}\int_0^T \Gamma_{2t}'\Gamma_{2t} dt\right)\right\} dt\right] .$$

In our matrix AJD setting, this property does not follow from a standard Novikov-type condition. However, it follows from a localization argument; see, e.g., Mayerhofer (2014). We now show that the second line of (A-7) also defines a martingale process. Using the independence between IID log jump sizes J^* and counting process N^* under the physical

probability, it is enough to show that:

$$1 = E_0^{\mathbb{P}} \left[\exp\left\{-(\lambda^- - \lambda^{*-})J^{*-} - (\lambda^+ - \lambda^{*+})J^{*+}\right\} \frac{1/\lambda^{*-} + 1/\lambda^{*+}}{1/\lambda^- + 1/\lambda^+} \right] . \tag{A-10}$$

Explicit calculations of the expectation on the right hand side yield:

$$\frac{\lambda^{*-}\lambda^{*+}}{\lambda^{*-} + \lambda^{*+}} \cdot \frac{1/\lambda^{*-} + 1/\lambda^{*+}}{1/\lambda^- + 1/\lambda^+} \int_{-\infty}^{\infty} \exp\left(-\lambda^- J^{*-} - \lambda^+ J^{*+}\right) dJ^* = 1 .$$

With respect to the risk-neutral probability $\mathbb{Q}$, log return jumps follows a double exponential distribution with parameters λ^- , λ^+. Indeed, for any $u \in \mathbb{R}$ it follows:

$$E^{\mathbb{Q}}\left[\exp(uJ)\right] = \frac{\lambda^-\lambda^+}{\lambda^- + \lambda^+} \int_{-\infty}^{\infty} e^{uJ} e^{-\lambda^- J^- - \lambda^+ J^+} dJ ,$$

which is the Laplace transform of a double exponential distribution with parameter λ^-, λ^+. This concludes the proof.

□

II. Estimation Procedure

A. First Step: Kalman filter

The first estimation step is performed using a Kalman filter of the linearized process, using exclusively options in the observation equation. Thus we can estimate all risk-neutral parameters via the observation equation and the physical parameters of the state dynamics via the transition equation. We denote the set of all parameters estimated in the first step by $\theta := (M, Q, R, \beta, \lambda_0, \Lambda;\ M^*, \beta^*)$.

The physical dynamics of our state variable is given in (14):

$$dX_t = \left[\beta^* Q'Q + M^* X_t' + X_t M^{*\prime}\right] dt + \sqrt{X_t} dB_t^* Q + Q' dB_t^{*\prime} \sqrt{X_t}$$

We discretize this process on a weekly grid with $\Delta_k = 7$ calendar days. When there is no data for a given Wednesday, we skip the respective week and set $\Delta_k = 14$.

We initialize the filtered state $\widehat{X}_t$ to be the steady state $X_\infty^{\mathbb{P}}$, which can be computed by solving the Lyapunov equation $M^* X_\infty^{\mathbb{P}} + X_\infty^{\mathbb{P}} (M^*)' = Q'Q$. We initialize the variance matrix of $\widehat{X}_t$ as $\widehat{\Sigma}_0 = 0$. At each step, we compute exact expectations of mean and variance of $X_{t+\Delta}$

given X_t from the Laplace transform (A-1)

$$\begin{aligned}
\overline{X}_{t+\Delta} &= \beta\,\overline{\mu} + \Phi\widehat{X}_t\Phi' && \text{(A-11)}\\
\overline{V}_{t+\Delta} &= (I_4 + K_4)\left(\Phi\widehat{X}_t\Phi' \otimes \overline{\mu} + \beta\,\overline{\mu}\otimes\overline{\mu} + \overline{\mu}\otimes\Phi\widehat{X}_t\Phi'\right) && \text{(A-12)}\\
\text{with} & \\
\overline{\mu} &= -\frac{1}{2}C_{12}C_{11}' \\
\Phi &= e^{\Delta M^*} \\
C &= \exp\left[\Delta\begin{pmatrix} M^* & -2Q'Q \\ 0 & -(M^*)' \end{pmatrix}\right] = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}
\end{aligned}$$

where $C_{11}, C_{12}, C_{21}, C_{22}$ are all square 2×2 matrices and K_4 is the 4×4 commutation matrix. These calculations are used in the transition equation:

$$\widetilde{X}_{t+\Delta} = \overline{X}_{t+\Delta} \tag{A-13}$$

The predicted state $\widetilde{X}_{t+\Delta}$ is then used to compute the observation equations:

$$\widehat{O}_{t+\Delta,i} = O_{t+\Delta,i}(\widetilde{X}_{t+\Delta};\theta) + \varepsilon_{t+\Delta,i}, \qquad i = 1,\ldots,N_{t+\Delta} \tag{A-14}$$

where $\widehat{O}_{t+\Delta,i}$ denotes the Black-Scholes implied volatility of the $i-$th option on day $t+\Delta$, $N_{t+\Delta}$ the total number of options observed on that day, $O_{t+\Delta,i}(\widetilde{X}_{t+\Delta};\theta)$ the model-implied option prices and $\varepsilon_{t+\Delta,i}$ is an iid noise with zero mean and variance σ_r. We also allow for autocorrelation in the noise:

$$corr(\overline{\varepsilon}_{t+\Delta}, \overline{\varepsilon}_t) = \rho_r$$

where $\overline{\varepsilon}_t$ is the mean error over all options on day t.

We finally *update* the state using a linearization of the dynamics. We first linearize the transition equation and the observation equations by computing the Jacobian matrices:

$$\begin{aligned}
F &= \frac{\partial\overline{X}_{t+\Delta}}{\partial\widehat{X}_t} = \Phi\otimes\Phi \\
G_t &= \frac{\partial O_{t+\Delta}}{\partial\widetilde{X}_t}
\end{aligned}$$

where we applied the identity $\frac{\partial}{\partial X}BXC = C'\otimes B$ to obtain F, while G is calculated via numerical differentiation. The variance matrix of the state is:

$$\widetilde{\Sigma}_{t+\Delta} = F\widehat{\Sigma}_t F' + \overline{V}_{t+\Delta} \tag{A-15}$$

Finally we update the state and the variance matrix to be used in the next step:

$$\begin{aligned} S_t &= G_t \widetilde{\Sigma}_{t+\Delta} G_t' + \sigma_r^2 I_2 \\ H_t &= \widetilde{\Sigma}_{t+\Delta} G_t' S_t^{-1} \\ \widehat{X}_{t+\Delta} &= \widetilde{X}_{t+\Delta} + H_t \left(\widehat{O}_{t+\Delta,i} - O_{t+\Delta,i}(\widetilde{X}_{t+\Delta}, \theta) \right) \\ \widehat{\Sigma}_{t+\Delta} &= (I_2 - H_t G_t) \widetilde{\Sigma}_{t+\Delta} \end{aligned}$$

For every parameter set θ, we compute the time-series of the predicted state $\{\widetilde{X}_t\}$ and the log-likelihood function

$$\mathcal{L}(\theta) = \sum_{i=1}^{N} \left[\log det(S) + \left(\widehat{O}_{t+\Delta,i} - O_{t+\Delta,i}(\widetilde{X}_{t+\Delta}, \theta) \right)' S_t^{-1} \left(\widehat{O}_{t+\Delta,i} - O_{t+\Delta,i}(\widetilde{X}_{t+\Delta}, \theta) \right) \right] \tag{A-16}$$

The estimated parameter $\widehat{\theta}$ is the maximizer of $\mathcal{L}(\theta)$. The maximization itself is performed using differential evolution of Storn and Price (1997).

B. *Model Identification*

Our model allows for several parameter combinations that are observationally equivalent. Parameter identification requires that the option pricing model be unique under invariant transformations. We borrow from Dai and Singleton (2000) and study invariant transformations that change the state and parameter matrices without changing the joint distribution of option prices and thus the spot variance $V_t := Tr[X_t] + E(k^2)\,(Tr[\Lambda X_t] + \lambda_0)$.

In the first step of our estimation process, we jointly estimate the state, all risk-neutral parameters, and the physical parameters of the state dynamics $\theta = (X_t;\ M, Q, R, \beta, \lambda_0, \Lambda;\ M^*, \beta^*)$. To identify these parameters, we first focus on the risk-neutral, diffusive part.

The diffusive spot volatility is $Tr[X_t]$, therefore the only class of transformations that needs to be considered are trace invariant transformations. These are first the similarity transformation $\mathcal{T}_S = \mathcal{D} X_t \mathcal{D}^{-1}$ and second the permutation $\mathcal{T}_P$ that reorders the rows (or columns) of X_t.

Applying $\mathcal{T}_S$ to (6) results in a transformed model with state and parameter matrices

$$\mathcal{T}_S \theta = (\mathcal{D} X_t \mathcal{D}^{-1};\ \mathcal{D} M \mathcal{D}^{-1}, \mathcal{D} Q \mathcal{D}^{-1}, \mathcal{D} R \mathcal{D}^{-1};\ \beta\ ).$$

In order to identify our model, we apply parameter restrictions that only admit $\mathcal{D} = I_2$.

Without loss of generality, we can assume $|det(\mathcal{D})| = 1$.[1] Next we observe that the state matrix X_t is symmetric by construction, thus $\mathcal{D}X_t\mathcal{D}^{-1}$ also needs to be symmetric. This requires $\mathcal{D}$ to be orthogonal ($\mathcal{D}' = \mathcal{D}^{-1}$), thus $\mathcal{D}$ must be a rotation or mirror matrix.

We choose the following restrictions: M is lower triangular and the sign of M_{21} is positive. Choosing M to be lower triangular requires $\mathcal{D}$ to be lower triangular, in order to ensure $\mathcal{D}M\mathcal{D}^{-1}$ lower triangular. If $\mathcal{D}$ is both orthogonal and lower triangular, it must be a diagonal matrix $\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$ with elements $d_i = \pm 1$. We now have $\mathcal{D}M\mathcal{D}^{-1} = \begin{pmatrix} M_{11} & 0 \\ d_2/d_1\, M_{21} & M_{22} \end{pmatrix}$. By choosing the sign of M_{21} we exclude the case $d_1 \neq d_2$.

Our choices for M implicitly identify the state and select the order of the mean reversion speeds of the eigenvalues and thus of the components of X_t. Thus we also achieve identification with respect to $\mathcal{T}_P$. A direct consequence of our identification choices is the result that $X_{22,t}$ is the leading volatility factor and the identification of the sign of $X_{12,t}$.

We now discuss the identification of Q and R. To do so, we inspect the infinitesimal generator of the joint process for stock returns $Y_t := dS_t/S_t$ and state X_t (see Leippold and Trojani (2008)):

$$\begin{aligned} \mathcal{L}_{Y,X} &= \left(r - q - \frac{1}{2}Tr[X]\right)\frac{\partial}{\partial Y} + \frac{1}{2}Tr[X]\frac{\partial^2}{\partial Y^2} + 2Tr[XR'QD]\frac{\partial}{\partial Y} + \\ &\quad + Tr\left[\left(\beta Q'Q + MX + XM'\right)D + 2XDQ'QD\right] \end{aligned} \tag{A-17}$$

where $(D)_{ij} = \frac{\partial}{\partial X_{ij}}$ is the matrix differential operator.

The matrices Q and R only appear in the expressions $Q'Q$ and $R'Q$, i.e. only seven of their eight elements are identified. We choose Q to be the unique Choleski decomposition of $Q'Q$, i.e. Q upper triangular and positive definite. In order to reduce the number of parameters, we add the ad-hoc restriction for R to be also upper triangular.

Next, we focus on the spot jump variance $E^{\mathbb{Q}}(k^2)\left(Tr[\Lambda X_t] + \lambda_0\right)$ with $Tr[\Lambda X_t] = \Lambda_{11}X_{11,t} + (\Lambda_{12} + \Lambda_{21})X_{12,t} + \Lambda_{22}X_{22,t}$. Only the sum of the out-of diagonal elements of Λ are identified and we choose Λ upper triangular.

The physical parameter M^* enters our estimation via the transition equation of the Kalman filter (A-13) in the two expressions $\bar{\mu}$ and $\Psi\hat{X}_t\Psi'$. By construction, both expressions are symmetric and therefore only three elements are identified. We choose M^* to be lower triangular, to allow for an easy comparison to M. With this step, we indirectly identify the the price of risk Γ in (15).

[1] We can always construct a $\widetilde{\mathcal{D}} = \frac{1}{\sqrt{|det(\mathcal{D})|}}\mathcal{D}$ with $|det(\widetilde{\mathcal{D}})| = 1$ such that $\mathcal{D}X_t\mathcal{D}^{-1} = \widetilde{\mathcal{D}}X_t\widetilde{\mathcal{D}}^{-1}$.

The remaining parameter β^*_Λ in (28), which is estimated via OLS in the second estimation step, is fully identified.

C. Admissible Parameter Set

In order to ensure the existence and non-explosivity of our latent process (14), we have to apply the following additional restrictions to the feasible parameter set. First, $M^{*\prime}M^*$ must be negative definite to ensure the non-explosivity. Second, R must satisfy $RR' < I_2$ in order to ensure the existence of $Z_t = B_t R + W_t\sqrt{I_2 - RR'}$ in (8). Third, to ensure the existence of processes (6, 14) and of the change of measure (13), we require either $\beta > 1$, $\beta^* > 1$ and $\beta = \beta^*$ or $\beta > 3$ and $\beta^* > 3$. Finally, we require $\Lambda'\Lambda$ to be positive semi-definite and $\lambda_0 \geq 0$ in order to satisfy the positivity condition of the jump intensity $\lambda_t = \lambda_0 + Tr[\Lambda X_t]$. See Gruber (2015) for the details of the implementation of the constraints.

D. Definition of Level $\mathcal{L}_t$, Skew $\mathcal{S}_t$ and Term Structure $\mathcal{M}_t$ Proxies

To analyze our results in terms of observable properties of the implied volatility surface, such as in Figure 4, we define the following proxies[2]

$$
\begin{array}{lcll}
\text{level} & \mathcal{L}_t & := & IV(\tau = \frac{1}{12}, \Delta = 0.5) \\
\text{short term skew} & \mathcal{S}_t & := & \frac{1}{0.6-0.4}\left[IV(\tau = \frac{1}{12}, \Delta = 0.6) - IV(\tau = \frac{1}{12}, \Delta = 0.4)\right] \\
\text{long term skew} & \mathcal{S}_t^{long} & := & \frac{1}{0.6-0.4}\left[IV(\tau = \frac{3}{12}, \Delta = 0.6) - IV(\tau = \frac{3}{12}, \Delta = 0.4)\right] \\
\text{term structure} & \mathcal{M}_t & := & \frac{1}{\frac{3}{12}-\frac{1}{12}}\left[IV(\tau = \frac{3}{12}, \Delta = 0.5) - IV(\tau = \frac{1}{12}, \Delta = 0.5)\right] \\
\text{skew term structure} & \mathcal{M}_t^{skew} & := & \frac{1}{\frac{3}{12}-\frac{1}{12}}\left[\mathcal{S}_t^{long} - \mathcal{S}_t\right]
\end{array}
$$

where IV and Δ stand for the Black-Scholes implied volatility and delta. The time to maturity τ is measured in years. In the data, we obtain the required implied volatilities through two-dimensional interpolation of the volatility surface. In the model, we calculate these quantities exactly.

[2] We have evaluated the regression $IV(\tau, K)_t = \mathcal{L}_t + \mathcal{S}_t \cdot K + \mathcal{M}_t \cdot \tau$ as an alternative specification. We have found similar, but more noisy results. We have also performed robustness checks with respect to our definition. The alternative term structure measure $\mathcal{M}_t^6 := \frac{1}{\frac{6}{12}-\frac{1}{12}}\left[IV(\tau = \frac{6}{12}, \Delta = 0.5) - IV(\tau = \frac{1}{12}, \Delta = 0.5)\right]$ is, for example, 92% correlated with our term structure measure.

III. Unspanned Risks as Risk Premium Factors

Intuitively, the information encoded by the unspanned risks traded in option markets might help to predict future unspanned risks and the excess returns for trading these risks using option portfolios. Moreover, since risks X_{11}, X_{12} and X_{22} are computed virtually in real time, using the fix set of in-sample estimated model parameters, these predictions can be computed virtually in real time. In this section, we study the predictive power of unspanned risks for future realized variance, for the payoffs of variance swaps and for S&P 500 index excess returns. As a robustness check, we compare the model-implied predictive power to the one implied by nonparametric proxies of the smile for (i) the level of the one-month implied volatility, (ii) the slope of the term structure of at-the-money implied volatilities and (iii) the one-month implied skew.

A. Predictability of Realized Variance

Figure 5 of this Appendix collects the predictive regression results for realized variance $RV_{t+\tau}(\tau)$, over forecasting horizons between 1 and 12 months. As expected, realized variance is highly predictable over short horizons, with predictive regression R^2s up to 55% using model-implied unspanned risks X_{11}, X_{22} and X_{12}. At longer horizons, i.e., between 9 and 12 months, the predictive R^2s drop to about 15%. The predictive power using nonparametric option-implied proxies is virtually identical and the degree of predictability is very similar in- and out-of-sample. The largest contribution to the predictive power derives from high-frequency risk X_{22}, which accounts for about 80% (66%) to the predictive R^2 at horizons of 1 month (12 months), while the residual contribution to the predictive R^2s is virtually exhausted by unspanned risk X_{12}. Given the high speed of mean reversion of these two risks, the term structure of predictive R^2s for future realized variance is downward sloping.

B. Predictability of Variance Swap Payoffs

We now address the predictability of variance swap payoffs $RV_{t+\tau}(\tau) - E_t^{\mathbb{Q}}[RV_{t+\tau}(\tau)]$. As the expected value of such payoffs is the variance risk premium, this is equivalent to studying the existence of a particular set of affine risk premium factors for variance risk premia. Figure 6 of this Appendix summarizes the predictive regression results. We obtain significant in-sample R^2s for all predictive regressions, where the largest predictive power is generated by the unconstrained predictive regression with model-implied risks X_{11}, X_{12} and X_{22}. The R^2s of the unconstrained predictive regression range between 25% and 60%, with a peak at the 12 month horizon. Those generated by the model-implied variance risk premium range between 22% and 45%, with a peak at the 6 month horizon. We find that the predictive

power of the unconstrained regression is almost exclusively generated by unspanned risks X_{11} and X_{12}, suggesting these two risks as natural variance risk premium factors. The larger predictive power of the unconstrained regression relative to the model-implied variance risk premium is completely explained by a larger estimated loading for low-frequency risk X_{11}.

The out-of-sample predictive regression analysis confirms the information content of unspanned risks as variance risk premium factors. We obtain the largest predictive R^2s, increasing monotonically with the horizon from 12% to 24%, for the model-implied variance risk premium predictions. The predictive R^2s of unconstrained regressions are clearly lower, while the degree of predictability implied by nonparametric proxies of the smile is negligible. Based on these findings, we conclude that the arbitrage-free constraints embedded into the model-implied variance risk premia are supported by the in-sample and the out-of-sample evidence. Compared to unconstrained predictive regressions, the model-implied arbitrage-free constraints allow us to isolate the high-frequency variance risk premium factor X_{22}, which otherwise would be very difficult to identify. As shown in the main text, identifying such high-frequency unspanned risk is important to understand the dynamics of the term structure of variance risk premia.

C. Predictability of S&P 500 index Excess Returns

We finally address the predictability of S&P 500 index excess returns $r^e_{t+\tau}$. As we did not assume any particular specification the equity premium for our identification of unspanned risks, we study exclusively unconstrained predictive regressions with unspanned risks X_{11}, X_{12} and X_{22}. In order to account parsimoniously for a possible nonlinearity of the predictive relation with respect to the frequency composition of the volatility, we estimate the following threshold-linear regression with endogenous threshold $T \in (0,1)$:

$$r^e_{t+\tau} = [\alpha_l + tr(\beta_l X_t)]\mathbb{I}_{\{[X_{11t}/tr(X_t)]<T\}} + [\alpha_h + tr(\beta_h X_t)]\mathbb{I}_{\{[X_{11t}/tr(X_t)]\geq T\}} + \epsilon_{t+\tau} \ . \quad \text{(A-18)}$$

In this equation, $\mathbb{I}_{\{\cdot\}}$ is an indicator function, while α_u and symmetric 2×2 matrix β_u are regime-dependent regression intercepts and slope parameters in states $u = l$ and $u = h$ of less and more persistent volatility, respectively. States of less (more) persistent volatility are defined as states in which condition $[X_{11t}/tr(X_t)] < T$ (condition $[X_{11t}/tr(X_t)] \geq T$) holds, for a threshold $T \in (0,1)$ estimated jointly with the (regime-dependent) predictive regression parameters. By making the equity compensation for low-frequency volatility risk possibly dependent on the frequency composition of the volatility, we incorporate an economically plausible long-run risk channel for index equity premia, suggested by our identification of unspanned risks. At the same time, we allow for a more flexible specification of the predictive

system, in which we can directly test the specification of an affine relation between S&P 500 index equity premia and unspanned risks. The predictive regression results are summarized in Figure 7 of this Appendix.

We estimate a significant threshold $T = 0.15$ ($T = 0.18$) for the in-sample (the full) sample period, which is quite stable across horizons. This finding is direct evidence of a nonlinearity of the predictive relation in the frequency composition of the volatility. All (regime-dependent) parameters in the predictive system are significant and the resulting degree of predictability is not negligible, with full-sample predictive R^2s ranging between 5% and 16% and peaking at an horizon of 5 months. We also find that this predictive power is almost entirely generated by the more persistent unspanned risks X_{11} and X_{12}, suggesting short-term implied skewness and long-term at-the-money implied volatility as useful risk premium factors for modeling time-varying S&P 500 index equity premia. In a different perspective, given the interpretation of X_{11} as the low frequency component of the volatility and $-X_{12}$ as the negative skewness of index returns in our model, these findings are suggestive of an extended risk-return tradeoff consistent with an intertemporal CAPM in which stochastic skewness is priced.

IV. Additional Figures

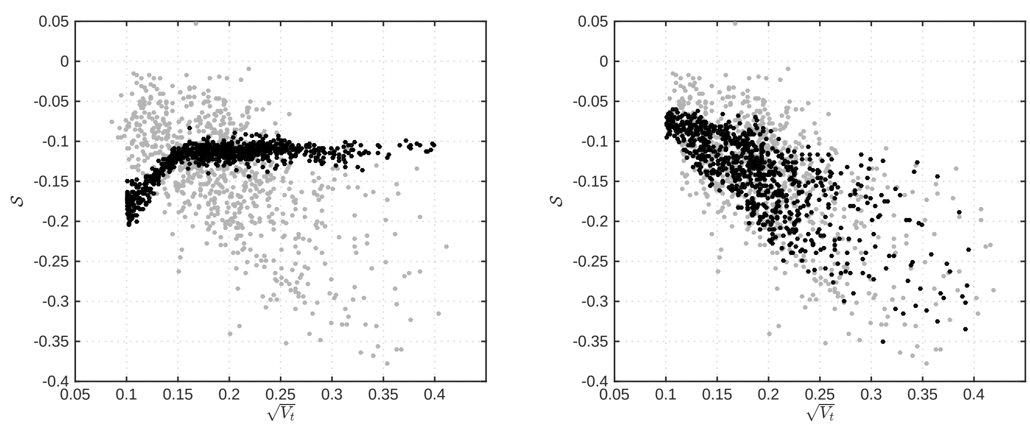


Figure 1: Relation between implied volatility level and implied volatility skew in Bates-type models. In each panel, we scatter plot two proxies of option-implied skewness ($\mathcal{S}$) and volatility ($\sqrt{V_t}$). Every dot in each panel corresponds to the implied-volatility smile observed on a single trading day. Grey dots: data-implied values. Black dots – left panel: model-implied values for a one factor Bates model (SVJ_{10}). Black dots – right panel: model-implied values for a two-factor Bates model (SVJ_{20}). The exact calculation method for the option-implied skewness $\mathcal{S}$ is explained in Section II.$D.$ of this Appendix.

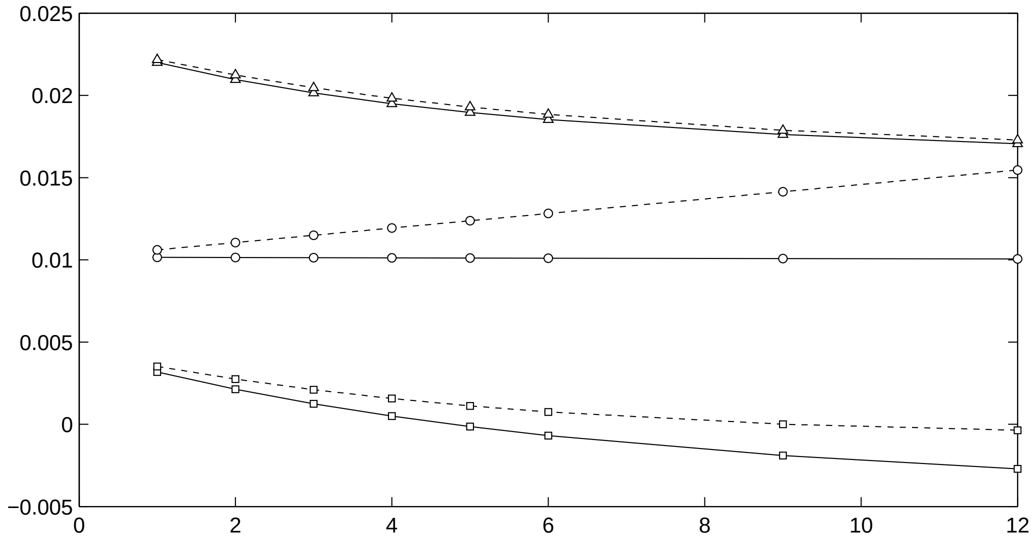


Figure 2: The term structure of the price of the smile. We plot the sample average of the model-implied expectations $\frac{1}{\tau}E_t^{\mathbb{P}}[\int_t^{t+\tau} X_s ds]$ and $\frac{1}{\tau}E_t^{\mathbb{Q}}[\int_t^{t+\tau} X_s ds]$, component-wise, for horizons τ from 1 to 12 months. Full lines report expectations under $\mathbb{P}$, dashed lines expectations under $\mathbb{Q}$. Circles: X_{11}; squares: X_{12}; triangles: X_{22}.

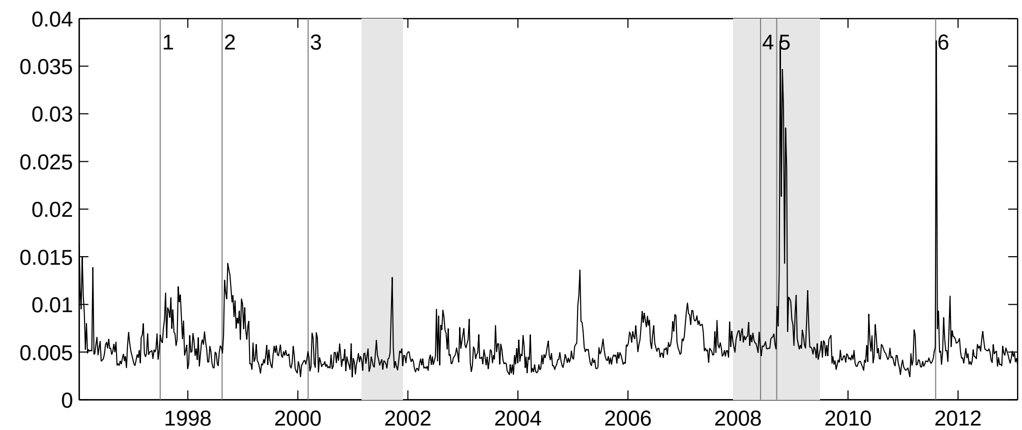


Figure 3: Time series of mean absolute implied volatility errors ($MAIVE$) for our model (SVJ_{31}). For every day t in our sample, we plot the $MAIVE$ on that day, defined by $MAIVE_t := \frac{1}{N_t}\sum_{i=1}^{N_t} |IV_i - \widehat{IV_i}|$, where N_t is the number of available options on that day. Grey areas highlight NBER recessions; vertical lines indicate important crisis events as listed in the caption of Figure 2 of the main paper.

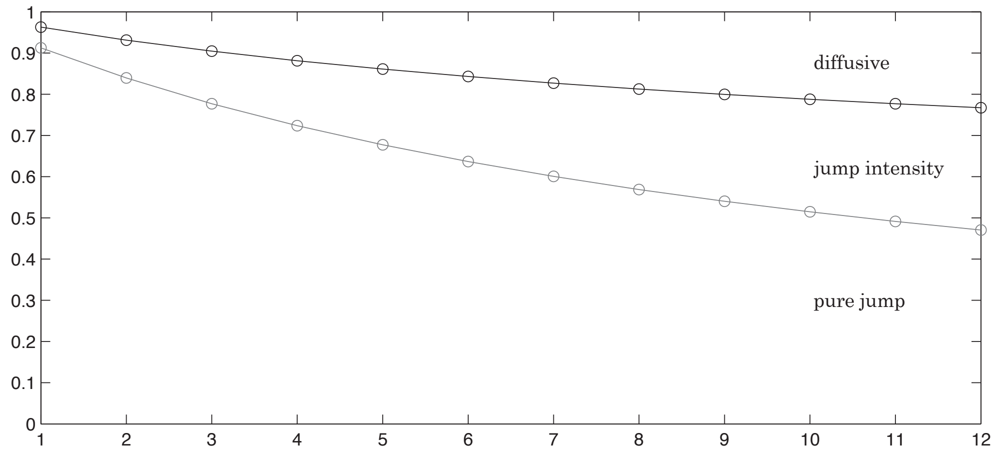


Figure 4: Unconditional decomposition of the variance risk premium for horizons 1 to 12 months. We plot from the bottom to the top the fractions of variance risk premium due to pure jump risk, jump intenstiy risk and diffusive volatility risk.

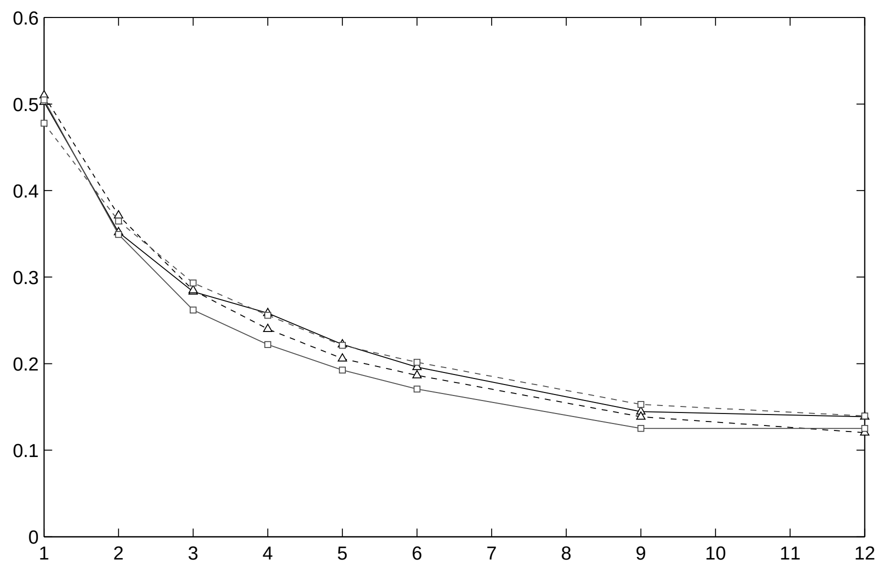


Figure 5: Predictive regression for realized variance $RV_{t+\tau}(\tau)$. We regress the future realized variance for horizons between 1 month and 12 months (reported on the-axis) on (i) the three unspanned risks $X_{11,t}, X_{12,t}, X_{22,t}$ (triangles) and (ii) three standard nonparametric measures of option-implied volatility level, skew and term structure (squares). Full lines correspond to in-sample (1996/01-2002/12) R^2s; dashed lines correspond to out-of-sample (2003/01-2013/01) R^2s. For both R^2 computations, model-implied and predictive regression parameters are fixed to the in-sample point estimates. The exact calculation method for the option-implied skewness and term structure is explained in Section II.*D.* of this Appendix.

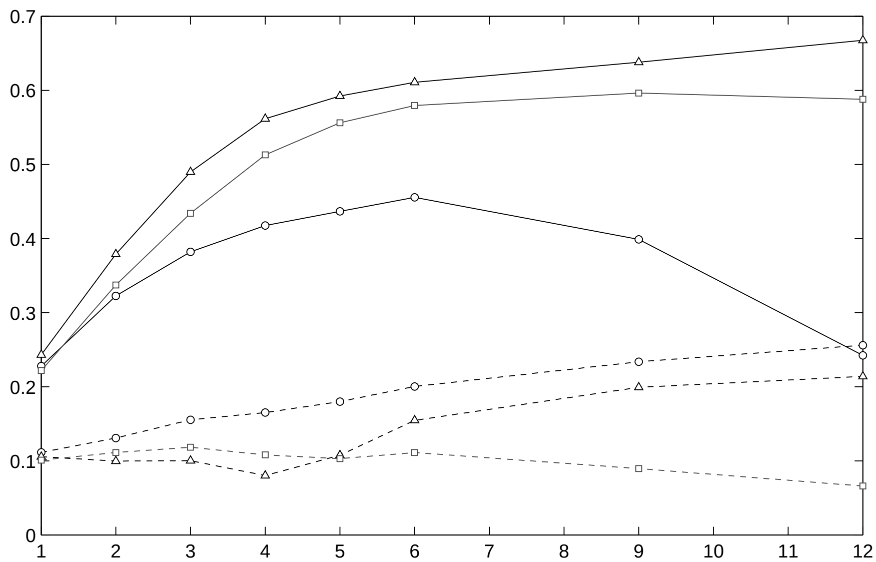


Figure 6: Predictive regression for synthetic variance swap positions. We regress future variance swap payoffs for horizons between 1 and 12 months (reported on the x-axis) on (i) the model-implied variance risk premium (circles), (ii) the three unspanned risks $X_{11,t}, X_{22,t}, X_{12,t}$ (triangles) and (iii) three standard nonparametric measures of option-implied volatility level, skew and term structure (squares). Full lines correspond to in-sample (1996/01-2002/12) R^2s; dashed lines correspond to out-of-sample (2003/01-2013/01) R^2s. For both R^2 computations, model-implied and predictive regression parameters are fixed to the in-sample point estimates. The exact calculation method for the option-implied skewness and term structure is explained in Section II.D. of this Appendix.

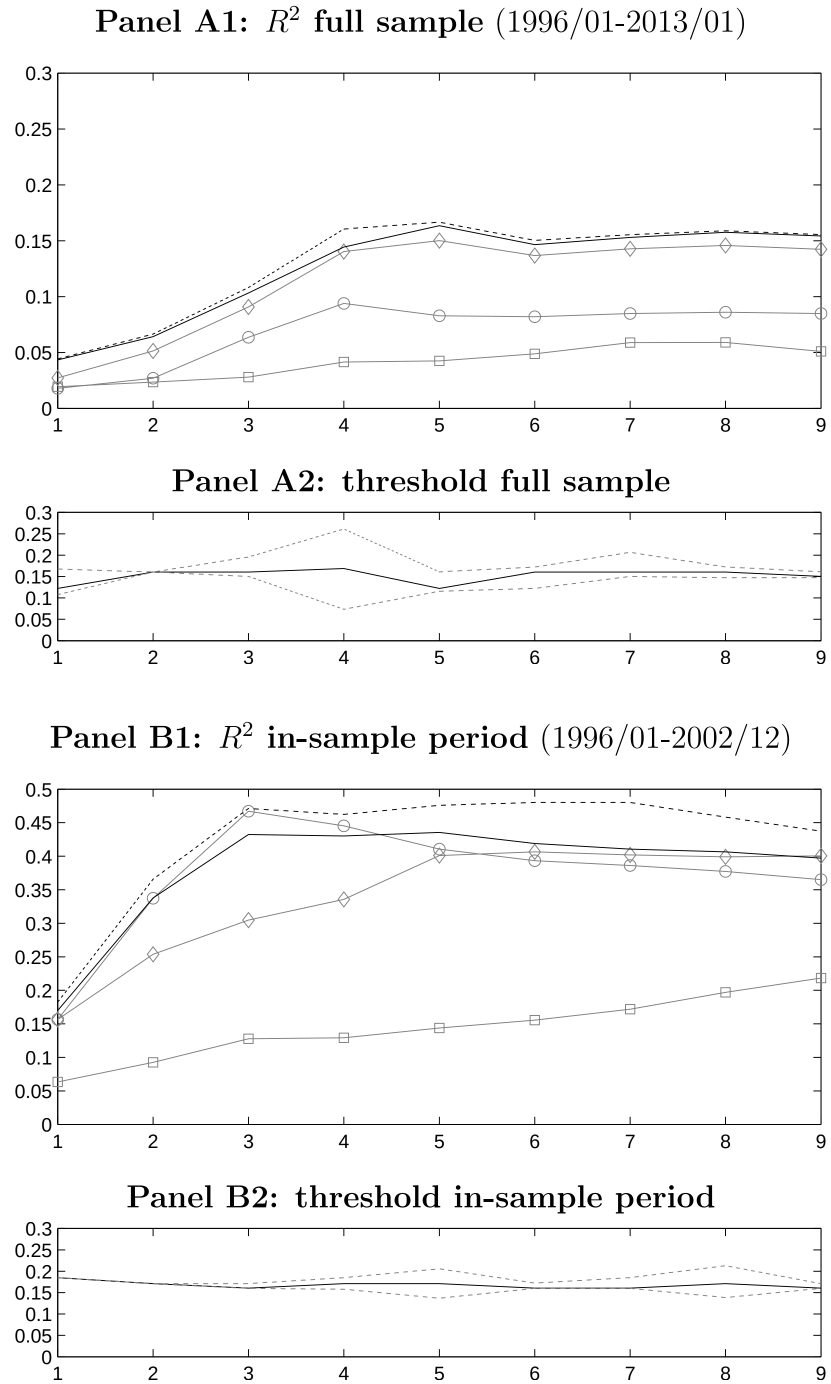


Figure 7: Predictive regression for future index excess returns. We perform a threshold predictive regression for future S&P 500 index excess returns, as defined in equation (A-18). We perform a threshold regression of future realized excess returns for horizons between 1 month and 9 months (reported on the x-axis) on (i) the individual unspanned risks X_{11} (diamonds), X_{12} (circles), X_{22} (squares); (ii) our preferred model using the risk factors $\tilde{X}_{11}$ and $\tilde{X}_{12}$, where $\tilde{X}_{ij}$ is the orthogonal projection of risk X_{ij} on X_{22} (black line) and (iii) jointly all unspanned risks (dashed line). In all regressions, we use the fraction of long-run risk in the diffusive variance $X_{11t}/Tr[X_t]$ as threshold variable. We report the R^2 of the predictive regressions in panels A1, B1, as well as estimates and 95% confidence intervals for the threshold in panels A2, B2.

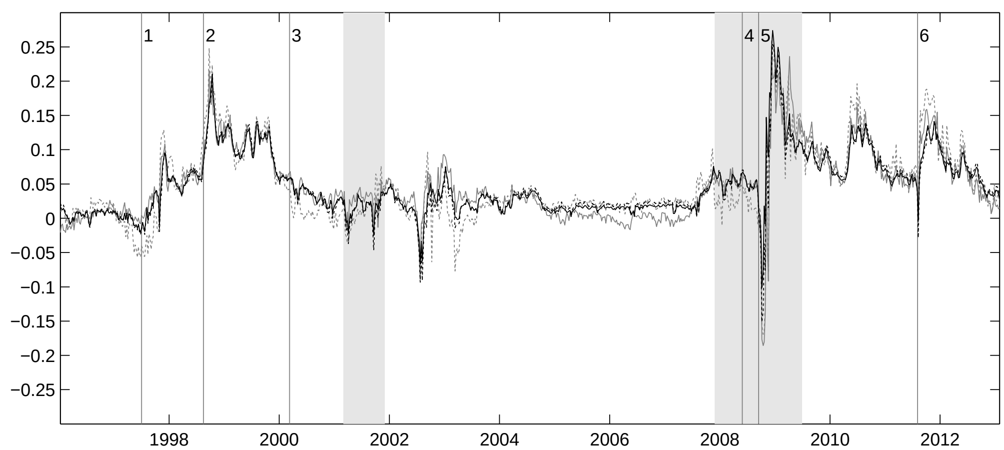
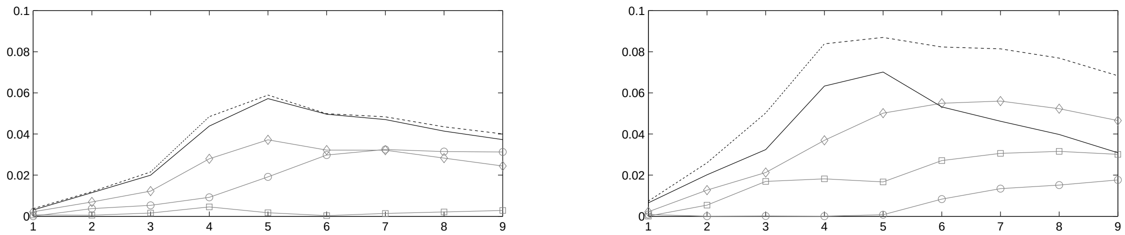


Figure 8: Conditional equity risk premium ERP at the optimal prediction horizon of five months. We perform a predictive regression of index excess returns $r_{t,t+\tau} = \alpha + \beta' x_t + \epsilon_t$, where x_t is a set of model-implied risk factors. We then calculate the annualized conditional ERP as $ERP_t = \frac{1}{\tau} E[r_{t,t+\tau}]$. Grey line: ERP implied by a predictive regression using the $\tilde{U}$ factor of Andersen, Fusari, Todorov (2013), where $\tilde{U}$ is an orthogoal projection of their $U-$factor on their factors V_1 and V_2. Grey dashed line: ERP implied by their full model. Black line: ERP implied by our model's $\tilde{X}_{11}$ and $\tilde{X}_{12}$ factors, where $\tilde{X}_{ij}$ is the orthogonal projection of risk X_{ij} on X_{22}. Black dashed line: ERP implied by our full model. Grey areas highlight NBER recessions; vertical lines indicate important crisis events as listed in the caption of Figure 2 of the main paper.

Panel A: full sample (1996/01-2013/01)



Panel B: in-sample period (1996/01-2002/12)

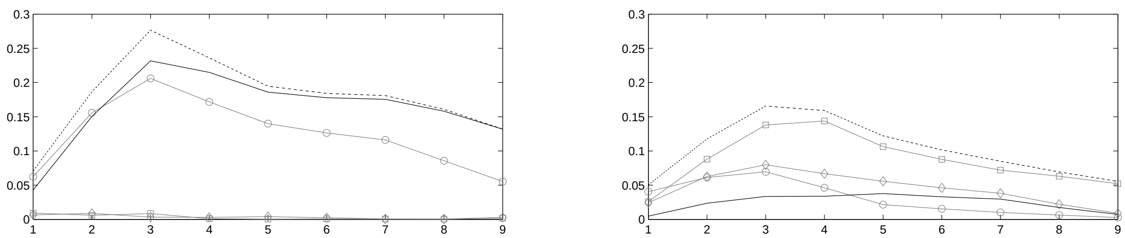


Figure 9: R^2 of the predictive regressions of future index excess returns on the unspanned risks traded in option markets, as a function of the horizon in months. We regress excess returns $r_{t,t+\tau}$ of the S&P 500 index for horizons $\tau = 1$ to 9 months on a set of option risks x_t. Left panels: Our model SVJ_{31}; grey: univariate regressions with individual risks X_{11} (diamonds), X_{12} (circles), X_{22} (squares); black: preferred model using the risk factors $\tilde{X}_{11}$ and $\tilde{X}_{12}$, where $\tilde{X}_{ij}$ is the orthogonal projection of risk X_{ij} on X_{22}; black dashed: all three unspanned risks X_{11}, X_{12} and X_{22}. Right panels: model of Andersen, Fusari, Todorov (2013); grey: univariate regressions with individual factors U (diamonds), V_1 (circles), V_2 (squares); black: preferred model using the risk factor $\tilde{U}$, where $\tilde{U}$ is an orthogoal projection of their $U-$factor on their factors V_1 and V_2; black dashed line: all risk factors V_1, V_2 and U.

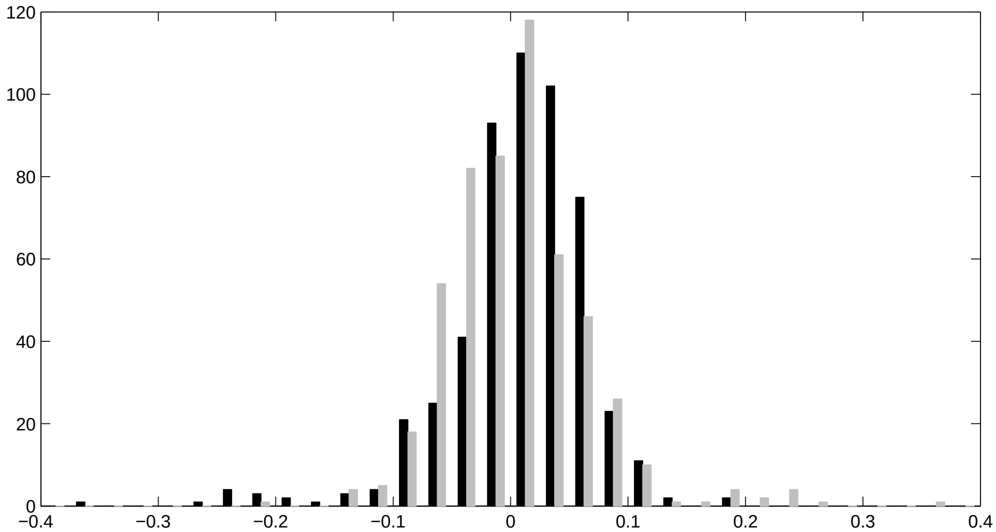


Figure 10: Distribution of out-of sample returns of the S&P 500 index (black bars) and a simple trading strategy that goes long (short) the index when the predicted equity premium is positive (negative), depicted in grey bars, over a forecasting horizon of 5 months.

V. Additional Tables

r	q	**Pure diffusion models**	**Jump-diffusion models**
1	0	SV_{10} $(N = 6)$ Heston (1993)	SVJ_{10} $(N = 8)$ Bates (1996)
2	0	SV_{20} $(N = 12)$ Christoffersen et al. (2009)	SVJ_{20} $(N = 18)$ Bates (2000)
3	0	SV_{30} $(N = 18)$ (this paper)	SVJ_{30} $(N = 25)$ (this paper)
3	1	SV_{31} $(N = 14)$ da Fonseca et al. (2008)	SVJ_{31} $(N = 21)$ Leippold and Trojani (2008)

Table 1: Models related to Assumption 2. r is the number of model state variables and q the number of skewness components disconnected from volatility. N is the number of model parameters.

Panel A: Summary statistics of the data

	In-sample	Out-of sample	Total
Time frame	1996-2002	2003-01/2013	1996-01/2013
Sampling frequency		weekly	
Trading days T	359	524	883
Number of observations	37'499	85'237	122'736
Average time to maturity (days)	141.5	124.9	130.0
Average moneyness (S/K)	0.99	0.98	0.99

Panel B: Number of observations by duration and delta

	$\tau < 30$	$30 < \tau < 75$	$75 < \tau < 180$	$180 < \tau$	all
$\|\Delta\| < 0.2$	1'761	4'647	3'679	3'858	13'945
$0.2 < \|\Delta\| < 0.4$	2'576	7'460	6'369	6'769	2'3174
$0.4 < \|\Delta\| < 0.6$	2'575	8'258	7'303	7'586	25'722
$0.6 < \|\Delta\| < 0.8$	3'479	10'808	9'399	10'446	34'132
$0.8 < \|\Delta\|$	2'981	8'651	6'947	7'184	25'763
all	13'372	39'824	33'697	35'843	122'736

Table 2: Main characteristics of our S&P500 option panel. We use out-of the money calls and puts.

	SV_{20}	SV_{30}	SV_{31}	SVJ_{20}	SVJ_{30}	SVJ_{31}
RMSIVE						
in-sample	1.323	1.237	0.941	0.858	0.718	0.678
out-of sample	1.672	1.552	1.203	1.093	0.826	0.769
MAIVE						
in-sample	1.023	0.957	0.731	0.680	0.565	0.549
out-of sample	1.325	1.226	0.948	0.854	0.640	0.610
Average log-likelihood						
in-sample	7.288	7.359	8.001	8.100	8.315	8.491
out-of sample	6.667	6.878	7.298	7.265	7.955	8.005

Table 3: Indicators of pricing performance and statistical fit. We report indicators of in- and out-of-sample pricing performance and fit for our model (SVJ_{31}) and for the benchmark models in Table 1. The in-sample period for estimation is January 1996 to December 2002. The out-of-sample period is from January 2003 to January 2013. For each model, we report the daily root-mean-squared implied volatility error (*RMSIVE*) and the daily mean absolute implied volatility error ($MAIVE$). These quantities are computed using the filtered states implied by the in-sample weekly parameter estimates for each day of our in- and out-of-sample periods. As a measure of statistical model fit and predictive ability, we also report the in- and the out-of-sample average value of the weekly likelihood function, evaluated at the in-sample parameter estimates.

Panel A1: $MAIVE$ for SVJ_{20} model, in-sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\lvert\Delta\rvert < 0.2$	0.938	0.607	0.655	0.593	0.664
$0.2 \leq \lvert\Delta\rvert < 0.4$	0.786	0.480	0.637	0.634	0.610
$0.4 \leq \lvert\Delta\rvert < 0.6$	0.699	0.510	0.537	0.517	0.539
$0.6 \leq \lvert\Delta\rvert < 0.8$	0.818	0.654	0.482	0.548	0.589
$0.8 \leq \lvert\Delta\rvert$	1.284	1.049	0.809	1.023	1.002
all	0.894	0.662	0.606	0.652	0.670

Panel A2: $MAIVE$ for SVJ_{31} model, in-sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\lvert\Delta\rvert < 0.2$	0.897	0.552	0.496	0.506	0.576
$0.2 \leq \lvert\Delta\rvert < 0.4$	0.842	0.419	0.480	0.468	0.502
$0.4 \leq \lvert\Delta\rvert < 0.6$	0.714	0.465	0.554	0.387	0.486
$0.6 \leq \lvert\Delta\rvert < 0.8$	0.735	0.497	0.486	0.378	0.476
$0.8 \leq \lvert\Delta\rvert$	0.887	0.567	0.628	0.871	0.721
all	0.804	0.495	0.531	0.503	0.542

Panel B1: $MAIVE$ for SVJ_{20} model, out of sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\lvert\Delta\rvert < 0.2$	0.804	0.976	1.067	1.393	1.089
$0.2 \leq \lvert\Delta\rvert < 0.4$	0.675	0.531	0.719	1.018	0.730
$0.4 \leq \lvert\Delta\rvert < 0.6$	0.955	0.501	0.469	0.759	0.605
$0.6 \leq \lvert\Delta\rvert < 0.8$	1.228	0.780	0.471	0.846	0.757
$0.8 \leq \lvert\Delta\rvert$	1.403	1.207	0.753	1.026	1.062
all	1.061	0.797	0.645	0.961	0.826

Panel B2: $MAIVE$ for SVJ_{31} model, out of sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\lvert\Delta\rvert < 0.2$	0.819	0.690	0.655	0.957	0.765
$0.2 \leq \lvert\Delta\rvert < 0.4$	0.753	0.430	0.506	0.566	0.522
$0.4 \leq \lvert\Delta\rvert < 0.6$	0.871	0.483	0.549	0.414	0.523
$0.6 \leq \lvert\Delta\rvert < 0.8$	0.954	0.519	0.503	0.544	0.566
$0.8 \leq \lvert\Delta\rvert$	0.964	0.653	0.487	0.881	0.704
all	0.886	0.546	0.528	0.639	0.602

Table 4: $MAIVE$ stratified by maturity and moneyness. We report the mean absolute implied volatility error across maturity and moneyness bins for our model (SVJ_{31}) and for the benchmark Bates (2000) model (SVJ_{20}), for the in-sample period (1996/01-2002/12) and the out of sample period (2003/01-2013/01).

Panel A1: Fraction of prices within bid/ask spread SVJ_{20} model, in-sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\|\Delta\| < 0.2$	0.363	0.405	0.322	0.371	0.368
$0.2 \leq \|\Delta\| < 0.4$	0.439	0.535	0.348	0.316	0.400
$0.4 \leq \|\Delta\| < 0.6$	0.602	0.574	0.413	0.318	0.445
$0.6 \leq \|\Delta\| < 0.8$	0.722	0.607	0.554	0.360	0.516
$0.8 \leq \|\Delta\|$	0.709	0.555	0.492	0.222	0.443
all	0.588	0.555	0.449	0.317	0.449

Panel A2: Fraction of prices within bid/ask spread SVJ_{31} model, in-sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\|\Delta\| < 0.2$	0.372	0.441	0.453	0.409	0.423
$0.2 \leq \|\Delta\| < 0.4$	0.379	0.617	0.451	0.404	0.472
$0.4 \leq \|\Delta\| < 0.6$	0.604	0.596	0.315	0.396	0.453
$0.6 \leq \|\Delta\| < 0.8$	0.766	0.713	0.489	0.488	0.580
$0.8 \leq \|\Delta\|$	0.844	0.852	0.598	0.276	0.589
all	0.616	0.668	0.461	0.403	0.517

Panel B1: Fraction of prices within bid/ask spread SVJ_{20} model, out of sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\|\Delta\| < 0.2$	0.471	0.351	0.273	0.176	0.298
$0.2 \leq \|\Delta\| < 0.4$	0.574	0.615	0.395	0.302	0.464
$0.4 \leq \|\Delta\| < 0.6$	0.590	0.727	0.607	0.383	0.586
$0.6 \leq \|\Delta\| < 0.8$	0.684	0.717	0.742	0.374	0.624
$0.8 \leq \|\Delta\|$	0.815	0.693	0.742	0.474	0.665
all	0.649	0.650	0.593	0.360	0.556

Panel B2: Fraction of prices within bid/ask spread SVJ_{31} model, out of sample

	$\tau < 30$	$30 \leq \tau < 75$	$75 \leq \tau < 180$	$\tau \geq 180$	all
$\|\Delta\| < 0.2$	0.452	0.498	0.512	0.253	0.433
$0.2 \leq \|\Delta\| < 0.4$	0.521	0.757	0.595	0.487	0.613
$0.4 \leq \|\Delta\| < 0.6$	0.625	0.683	0.503	0.557	0.590
$0.6 \leq \|\Delta\| < 0.8$	0.803	0.819	0.668	0.518	0.690
$0.8 \leq \|\Delta\|$	0.905	0.915	0.875	0.489	0.794
all	0.696	0.763	0.643	0.483	0.647

Table 5: Fraction of model-implied option prices within bid-ask spread for the benchmark SVJ_{20} model and our model (SVJ_{31}), across maturity and moneyness bins for the in-sample period (1996/01-2002/12) and the out of sample period (2003/01-2013/01).

	X_{11}	X_{12}	X_{22}
Min	0.0000	-0.0096	0.0001
Max	0.0516	0.0893	0.2610
Mean	0.0102	0.0044	0.0233
Median	0.0091	0.0014	0.0171
Positive	1.0000	0.6659	1.0000
Stdv	0.0091	0.0104	0.0259
Skewness	1.3460	3.6063	4.6427
Kurtosis	5.6376	20.9171	35.1886
AR(1)	0.9896	0.9529	0.8842
Half life	1.2753	0.2766	0.1083

Table 6: Summary statistics of weekly filtered unspanned risks X_{11}, X_{12} and X_{22} for our model for the whole sample (1996/01-2013/01). "Positive" denotes the fraction of positive realizations. Half lives are given in years.

Price of	$(E^{\mathbb{P}} - E_t^{\mathbb{Q}})\frac{1}{\tau}\left[\int_t^{t+\tau} X_{11s}ds\right]$			$(E^{\mathbb{P}} - E_t^{\mathbb{Q}})\frac{1}{\tau}\left[\int_t^{t+\tau} X_{12s}ds\right]$			$(E^{\mathbb{P}} - E_t^{\mathbb{Q}})\frac{1}{\tau}\left[\int_t^{t+\tau} X_{22s}ds\right]$		
Loading	X_{11}	X_{12}	X_{22}	X_{11}	X_{12}	X_{22}	X_{11}	X_{12}	X_{22}
1	-0.043	0.000	0.000	-0.025	-0.037	0.000	-0.002	-0.043	-0.000
2	-0.083	0.000	0.000	-0.048	-0.064	0.000	-0.006	-0.071	-0.000
3	-0.120	0.000	0.000	-0.069	-0.082	0.000	-0.012	-0.087	-0.000
4	-0.156	0.000	0.000	-0.088	-0.094	0.000	-0.018	-0.096	-0.000
5	-0.189	0.000	0.000	-0.106	-0.101	0.000	-0.025	-0.101	-0.000
6	-0.221	0.000	0.000	-0.121	-0.105	0.000	-0.032	-0.102	-0.000
7	-0.251	0.000	0.000	-0.136	-0.107	0.000	-0.038	-0.101	-0.000
8	-0.279	0.000	0.000	-0.149	-0.106	0.000	-0.044	-0.098	-0.000
9	-0.306	0.000	0.000	-0.161	-0.105	0.000	-0.050	-0.095	-0.000
10	-0.331	0.000	0.000	-0.172	-0.103	0.000	-0.055	-0.091	-0.000
11	-0.355	0.000	0.000	-0.182	-0.100	0.000	-0.060	-0.087	-0.000
12	-0.377	0.000	0.000	-0.191	-0.096	0.000	-0.065	-0.083	-0.000
12-1	-0.335	0.000	0.000	-0.166	-0.059	0.000	-0.063	-0.040	-0.000
12-3	-0.257	0.000	0.000	-0.122	-0.015	0.000	-0.053	0.004	0.000
3-1	-0.078	0.000	0.000	-0.044	-0.044	0.000	-0.010	-0.044	-0.000

Table 7: Market price of the smile. Loadings of risk factors X_{11}, X_{12}, X_{22} on the market price of the smile of integrated risk factor $\frac{1}{\tau}\int_t^{t+\tau} X_{ijs}ds$, i.e., $(E^{\mathbb{P}} - E_t^{\mathbb{Q}})\left[\frac{1}{\tau}\int_t^{t+\tau} X_{ijs}ds\right]$ $(1 \le i \le j \le 2)$, for horizons τ from 1 to 12 months. The last three rows compute the contributions to the term structure of the market price of the smile, measured price of the smile at a longer horizon minus price of the smile at a shorter horizon.

τ	$VRP_t^c(\tau)$			$VRP_t^{dc}(\tau)$			$VRP_t^{dj}(\tau)$			$VRP_t(\tau)$		
months	X_{11}	X_{12}	X_{22}	X_{11}	X_{12}	X_{22}	X_{11}	X_{12}	X_{22}	X_{11}	X_{12}	X_{22}
1	-0.044	-0.043	-0.000	-0.068	-0.045	-0.000	-0.537	-0.762	-0.272	-0.649	-0.851	-0.272
2	-0.089	-0.071	-0.000	-0.131	-0.076	-0.000	-0.525	-0.679	-0.223	-0.745	-0.825	-0.223
3	-0.132	-0.087	-0.000	-0.191	-0.095	-0.000	-0.511	-0.607	-0.186	-0.835	-0.790	-0.186
4	-0.174	-0.096	-0.000	-0.248	-0.108	-0.000	-0.497	-0.546	-0.158	-0.919	-0.750	-0.158
5	-0.214	-0.101	-0.000	-0.300	-0.114	-0.000	-0.483	-0.492	-0.135	-0.998	-0.708	-0.135
6	-0.253	-0.102	-0.000	-0.349	-0.117	-0.000	-0.469	-0.446	-0.118	-1.071	-0.665	-0.118
7	-0.289	-0.101	-0.000	-0.395	-0.118	-0.000	-0.455	-0.406	-0.103	-1.139	-0.625	-0.103
8	-0.323	-0.098	-0.000	-0.437	-0.117	-0.000	-0.441	-0.371	-0.092	-1.202	-0.586	-0.092
9	-0.356	-0.095	-0.000	-0.477	-0.114	-0.000	-0.427	-0.341	-0.083	-1.260	-0.550	-0.083
10	-0.386	-0.091	-0.000	-0.514	-0.111	-0.000	-0.414	-0.315	-0.075	-1.314	-0.517	-0.075
11	-0.415	-0.087	-0.000	-0.548	-0.107	-0.000	-0.401	-0.291	-0.068	-1.364	-0.486	-0.068
12	-0.442	-0.083	-0.000	-0.581	-0.103	-0.000	-0.388	-0.271	-0.063	-1.411	-0.457	-0.063
12-1	-0.398	-0.040	-0.000	-0.513	-0.058	-0.000	0.149	0.492	0.210	-0.762	0.394	0.210
12-3	-0.310	0.004	0.000	-0.389	-0.007	0.000	0.123	0.337	0.123	-0.576	0.333	0.123
3-1	-0.088	-0.044	-0.000	-0.124	-0.050	-0.000	0.026	0.155	0.086	-0.186	0.061	0.086

Table 8: Loadings of option-implied components X_{11}, X_{12}, X_{22} on diffusive, intensity and pure jump variance risk premia $VRP_t^c(\tau)$, $VRP_t^{dc}(\tau)$ and $VRP_t^{dj}(\tau)$, respectively. For horizons τ from 1 to 12 months, we compute the model implied loading of state variables X_{11}, X_{12} and X_{22} in diffusive, intensity and pure-jump variance risk premia (columns 2 to 4). The last column reports the state variables loadings in the total model-implied variance risk premium $VRP_t(\tau)$. The last three rows compute the contribution of each option-implied component to three proxies for the slope of the term structures of variance risk premia, measured as $VRP_t^u(12) - VRP_t^u(\tau)$, for $u = c, dc, dj$ and $\tau = 1, 3$ months, respectively.

	1mo	2mo	3mo	4mo	5mo	6mo	7mo	8mo	9mo
Out of sample	0.53	0.54	0.52	0.52	0.58	0.57	0.55	0.57	0.57
Conundrum	0.52	0.50	0.51	0.46	0.47	0.49	0.48	0.51	0.53
Crisis	0.63	0.60	0.50	0.63	0.85	0.85	0.86	0.87	0.88
Post crisis	0.49	0.57	0.56	0.55	0.58	0.54	0.47	0.47	0.42

Table 9: Out of sample sign correlations between realized excess returns of the S&P 500 index over horizons from 1 to 9 months and predicted returns from our threshold regression. We perform a predictive threshold regression as defined in equation (A-18) for the in-sample period 1996/01-2002/12 and evaluate the signs of the predicted returns for the out of sample period 2003-2013/01 and three sub-periods: Conundrum (2003/01-2007/12), Financial Crisis (2008/01-2009/12) and Post-Crisis (2010/01-2013/01).

Panel A: 1 month horizon

	Index			Strategy		
	mean	SR	skewness	mean	SR	skewness
Out of sample	0.031	0.184	-1.597	0.046	0.229	-1.064
Conundrum	0.056	0.511	-0.669	0.022	0.133	-0.398
Crisis	-0.117	-0.438	-1.167	-0.042	-0.134	-1.067
Post crisis	0.088	0.584	-1.040	0.145	0.960	0.022

Panel B: 2 month horizon

	Index			Strategy		
	mean	SR	skewness	mean	SR	skewness
Out of sample	0.033	0.190	-1.807	0.064	0.323	-1.828
Conundrum	0.060	0.556	-0.197	0.061	0.388	0.116
Crisis	-0.118	-0.401	-1.187	-0.108	-0.333	-1.570
Post crisis	0.091	0.654	-0.704	0.189	1.533	-0.481

Panel C: 3 month horizon

	Index			Strategy		
	mean	SR	skewness	mean	SR	skewness
Out of sample	0.032	0.181	-1.593	0.069	0.357	-0.343
Conundrum	0.058	0.542	-0.124	0.015	0.099	0.037
Crisis	-0.104	-0.347	-0.892	0.170	0.549	-0.775
Post crisis	0.083	0.609	-0.555	0.092	0.672	-0.141

Panel D: 4 month horizon

	Index			Strategy		
	mean	SR	skewness	mean	SR	skewness
Out of sample	0.031	0.173	-1.531	0.046	0.230	0.081
Conundrum	0.056	0.535	-0.157	0.003	0.022	0.162
Crisis	-0.098	-0.313	-0.760	0.167	0.512	-0.450
Post crisis	0.079	0.589	-0.314	0.034	0.237	-0.137

Table 10: Out-of sample statistics of a simple trading strategy that goes long (short) the index when the predicted equity premium is positive (negative). We compare mean return, Sharpe ratio (SR) and skewness of the returns of the S&P 500 index to our trading strategy over horizons from 1 to 4 months. Returns and Sharpe ratios are annualized. We break down the out of sample period (2003/01-2013/01) into three sub periods: the Conundrum (2003/01-2007/12), the Financial Crisis (2008/01-2009/12) and the Post-Crisis period (2010/01-2013/12).