The Price of the Smile and Variance Risk Premia

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

First version: February 2014. This version: April 15, 2020

*Peter H. Gruber is from the Università della Svizzera Italiana, Lugano; Claudio Tebaldi is from Università L. Bocconi, Department of Finance and is a fewllow of IGIER, Baffi-Carefin and LTI@UniTO; Fabio Trojani is from the University of Geneva and a Senior Chair of the Swiss Finance Institute. We thank the Editor (Gustavo Manso) and two anonymous referees for many insightful comments that helped us to improve our paper. All authors gratefully acknowledge the financial support of the CAREFIN foundation. We wish to thank Nicola Fusari for making his option pricing toolbox available to us, Kai Wang for research support and the CINECA consortium for access to their computing infrastructure. 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"). Claudio Tebaldi acknowledges funding from MIUR - PRIN Bando 2017 - prot. 2017TA7TYC and LTI@UniTo. We thank participants to the 2010 and 2014 World Congress of the Bachelier Society, the European Summer Symposia in Financial Markets, the 2010 and 2016 EFA meetings, the 2014 SoFIE Conference on Skewness, Heavy Tails, Market Crashes, and Dynamics, the 2015 AFFI conference, the 2016 Annual SoFIE conference, the 2016 Workshop on Forecasting and Financial Markets at Erasmus University and the finance seminars at Bocconi University, the University of Geneva, the EPFL Lausanne, the University of Bonn, the University of Manchester, the Luxembourg School of Finance, WU Wien and Banque de France, for valuable comments. The usual disclaimer applies.

Autors' contacts: peter.gruber@usi.ch, claudio.tebaldi@unibocconi.it, fabio.trojani@unige.ch
Corresponding author: Fabio Trojani, University of Geneva and Swiss Finance Institute, Bd du Pont d'Arve 40, 1205 Geneve, Switzerland

The Price of the Smile and Variance Risk Premia

ABSTRACT

Using a new specification of multi-factor volatility, we estimate the hidden risk factors spanning SPX implied volatility surfaces and the risk premia of volatility-sensitive payoffs. SPX implied volatility surfaces are well-explained by three dependent state variables reflecting (i) short- and long-term implied volatility risks and (ii) short-term implied skewness risk. The more persistent volatility factor and the skewness factor support a downward sloping term structure of variance risk premia in normal times, while the most transient volatility factor accounts for an upward sloping term structure in periods of distress. Our volatility specification based on a matrix state process is instrumental to obtain a tractable and flexible model for the joint dynamics of returns and volatilities, which improves pricing performance and risk premium modeling with respect to recent three-factor specifications based on standard state spaces.

JEL classification: G10, G12, G13

Keywords: Price of the Smile, Price of Volatility, Factor Models, Matrix Jump Diffusions, Option Pricing, Stochastic Volatility, Unspanned Skewness, Financial Constraints, Financial Intermediation, Financial Crisis, Variance Swaps, Skew Swaps.

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 financial literature has reached a consensus that aggregate volatility shocks are priced in modern financial markets, i.e. that there exists a (typically negative) variance risk premium. However, far less is known about:

- The relation between the characteristics of the option-implied volatility smile and variance risk premia,
- The time-variation and term structure of variance risk premia,
- The relation between volatility risk factors and market equity premia.

Studying these features in a coherent arbitrage-free framework is not an easy task. First, volatility risk is driven by multiple sources of risk with distinct persistence properties, which may comove dynamically and contribute differently to the price of volatility; see, e.g., Bates (2000), Gruber, Tebaldi and Trojani (2010), Andersen, Fusari and Todorov (2015a) and Bardgett, Gourier and Leippold (2019). Therefore, multi-factor models are essential for identifying the distinct components of volatility and their market prices of risk. Second, volatility risk is tradable in liquidity option markets by means of suitable replicating portfolios, similar to other sources of risk such as skewness risk; see, e.g., Bergomi (2004, 2005, 2008, 2009), Kozhan, Neuberger and Schneider (2013), and Schneider and Trojani (2014, 2018a). Therefore, volatility and skewness risk need to be studied jointly with the risk premia of tradable option portfolios in arbitrage-free markets. Third, many models, such as Bates (2000)-type models, induce a counterfactually tight link between volatility and skewness; see, e.g., Gruber et al. (2010), Andersen et al. (2015a) and Constantinides and Lian (2015). Therefore, they may artificially compress the distinct information contents of

option-implied skewness and volatility.[1] Fourth, the latent factor dynamics estimated by these models is often difficult to interpret economically in terms of directly observable key properties of the implied volatility surface.

To jointly address all the main economic properties of volatility risk, we adopt a new multi-factor volatility specification that allows us to study in a coherent, arbitrage-free framework the multivariate implied volatility dynamics, the term structure of variance risk premia and their link to equity premia. We study the price of volatility risk consistently with the prices of all tradable risks in arbitrage-free option markets using a suitable three-factor specification of stochastic volatility in the class of matrix affine jump diffusions (AJD) introduced by Leippold and Trojani (2008). This approach has a number of useful properties for identifying the risk premia of the various sources of risk in a multi-factor volatility dynamics. First, it yields a well-defined joint dynamics for the volatility, the volatility of volatility and the feedback effects between returns and volatility, under similarly simple parameter constraints as in standard affine models. Second, it naturally embeds a stochastic skewness that is only weakly linked to the level of the volatility. Third, it easily accommodates flexible functional forms for factor risk premia, which can depend on all volatility risk factors. Fourth, it is based on dynamically correlated volatility factors that are more directly interpretable in terms of the observable dynamics of implied volatility and skewness.

Using exclusively option information, we first estimate with a tractable filtering approach the hidden transition dynamics of the volatility, together with all risk neutral parameters in the model. This allows us to identify the volatility factors driving S&P500 option *implied* volatilities and their factor risk premia. In a second step, we estimate variance risk premia, i.e., the risk premia for trading *realized* variance, by means of a consistent arbitrage-free regression of synthetic variance swap payoffs on model-implied variance risk premia. Both

[1]They may also imply too tight restrictions for factor risk premia, which are usually specified as affine in individual factor levels to preserve tractability.

steps do not require a specification of the equity premium dynamics, making this part of our approach robust to a potential misspecification of equity premia. In a third step, we assess the economic reliability of the estimated variance risk premium dynamics, by studying the model-implied equity premium properties under a matrix-affine specification also for equity premia.

Our volatility dynamics is based on a matrix AJD. This is instrumental for estimating three stochastically dependent volatility factors X_{11}, X_{12} and X_{22} with different persistence properties, which naturally summarize distinct observable characteristics of the volatility surface. With a half-life of about 5 weeks, factor X_{22} reflects the high-frequency component of the volatility and closely follows the 30-day at-the-money implied volatility. Factor X_{12} has a longer half-life of about three months and closely targets the 30-day implied skew. With a clearly longer half-life of more than one year, factor X_{11} reflects the low-frequency component of the volatility and closely traces the 12 month at-the-money implied volatility.

We find that our model produces an excellent fit of the cross-sectional and time series properties of the SPX volatility smile, relative to benchmark models in the literature. For instance, it clearly improves on the fit of two- and three-factor Bates (2000)-type models. Similarly, it delivers a pricing performance comparable to the affine double-jump diffusion model in Andersen et al. (2015a), despite the latter model having five more parameters and being estimated only under the risk neutral measure. Importantly we obtain a very stable fit in- and out-of sample, which is known to be a challenge for models consistently estimated under both the physical and pricing probabilities.[2] This evidence supports the usefulness of our matrix state space approach for capturing in a robust way the intrinsic volatility dynamics without the need for a complex jump specification in volatility.

[2]For example, the pricing performance ($RMSRE$) of the double-jump diffusion model of Bardgett et al. (2019) implies an out-of sample pricing error that can be up to 40% larger than in-sample.

The estimated physical and risk-neutral volatility dynamics in our model give rise to interesting risk premium properties. We obtain highly time-varying and unambiguously negative variance risk premia, ranging between zero and -16 percent squared (-11 percent squared), on an annualized basis, for monthly (annual) investment horizons. While variance risk premia are negative, we find that the variance term premium may change sign. The term structure of variance risk premia is usually downward sloping, but can become strongly upward sloping in conditions of market distress, when short-term downside risk is extraordinarily high. The downward sloping term structure of variance risk premia in normal times is spanned by the lower-frequency volatility factors X_{12} and X_{11}, which act as variance risk premium factors in such periods. In contrast, the upward sloping term structure in periods of distress is dominated by the high-frequency factor X_{22}, which acts as an additional variance risk premium factor for the short end of the term structure, when the short-term volatility and the volatility mean-reversion are unusually high.

Interestingly, we find that the lower-frequency factors X_{12} and X_{11} completely span the risk premia of all systematic shocks to implied volatilities, which are condensed in a vector of three factor risk premia. We call this vector the price of the smile. The structure of the price of the smile is striking. We find that the risk premium of low-frequency volatility factor X_{11} is proportional to its level, similar to standard affine models. Hence, long-horizon implied volatility fully explains the level of its risk premium. In contrast, the risk premium of skewness factor X_{12} depends on the level of both X_{11} and X_{12}. Therefore, implied skewness only partly explains the risk premium of implied skewness shocks. Finally, the risk premium of transient volatility factor X_{22} is proportional to the level of X_{12}. Thus, implied skewness fully explains the risk premium of short-horizon implied volatility shocks.

Finally, we show that the estimated variance risk premium dynamics can be supported with a matrix-affine specification of equity premia that is fully model-consistent and which

produces an economically plausible counter-cyclical dynamics. Under this specification, the estimated market equity premium varies between 0.2% and 16.5%, on an annual basis for a quarterly horizon, and it never violates the model-free equity premium lower bound in Schneider and Trojani (2019a).

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. It also explains the motivation for modelling the stochastic volatility of univariate returns using a matrix state space. Section **II.** presents our empirical findings, while Section **III.** concludes and highlights avenues for future developments.

I. Model

Our model is characterized by three mutually exciting risk factors defined on a matrix state space. It features interdependent factor prices as well as a compensation for variance risk and a time-varying skewness that are both partly disconnected from the diffusive spot volatility. In the sequel, we first discuss the motivation for using factor models defined on a matrix state space when specifying stochastic volatility features. In a second step, we introduce our model and the corresponding solutions for option prices and variance risk premia.

A. Multi-factor Volatility Models for Univariate Returns

Besides the standard positivity of the volatility, multi-factor models have to incorporate further natural constraints in order to induce a well-defined return dynamics. Indeed, while these models can give rise to flexible joint dynamics for the volatility, the volatility of volatility and the co-variation between returns and volatility, in doing so they need to imply a well-defined joint variance-covariance matrix for returns and volatilities. This matrix is also key to specifying well-defined dynamics for the risk premia of derivative assets in a market

with stochastic volatility.

A.1. The Drivers of Risk Premia in Derivatives Markets with Stochastic Volatility

Let dS_t/S_{t-} be a univariate continuous-time asset return and $V_{rrt} := \frac{1}{dt} Var_t^{\mathbb{P}}[dS_t/S_{t-}]$ be the conditional return variance under the physical probability. The joint covariance matrix of returns and the volatility is defined by

$$\boldsymbol{V}_t := \begin{bmatrix} V_{rrt} & V_{rvt} \\ V_{vrt} & V_{vvt} \end{bmatrix} , \tag{1}$$

where $V_{rvt} := \frac{1}{dt} Cov_t^{\mathbb{P}}[dS_t/S_{t-}, dV_{rrt}]$ and $V_{vvt} = \frac{1}{dt} Var_t^{\mathbb{P}}[dV_{rrt}]$. To understand the importance of matrix (1) for modelling the excess returns of derivatives in a market with stochastic volatility, consider for illustration a volatility derivative with price process $P_t := P(S_t, V_{rrt})$. Itô's Lemma yields the following expression for the dynamics of the derivative's value, assuming for brevity that it is delta- and theta-hedged:[3]

$$\frac{dP_t}{P_{t-}} = \frac{1}{P_{t-}} \left[\frac{\partial P(S_t, V_{rrt})}{\partial V_{rr}} dV_{rrt} + \frac{1}{2} tr(\boldsymbol{V}_t \boldsymbol{H}_t) dt \right] , \tag{2}$$

with $tr(\cdot)$ the trace operator and the Hessian matrix

$$\boldsymbol{H}_t := \begin{bmatrix} \frac{\partial^2 P(S_t, V_{rrt})}{\partial^2 S} & \frac{\partial^2 P(S_t, V_{rrt})}{\partial S \partial V_{rr}} \\ \frac{\partial^2 P(S_t, V_{rrt})}{\partial S \partial V_{rr}} & \frac{\partial^2 P(S_t, V_{rrt})}{\partial^2 V_{rr}} \end{bmatrix} . \tag{3}$$

In equation (2), the first term on the right hand side captures changes in the value of the derivative that are due to first-order variations in the level of the volatility. It induces a risk

[3]The assumption $P_t := P(S_t, V_{rrt})$ is only for illustration purposes. For volatility derivatives with an asset value depending on all volatility factors, the resulting risk premium description is more involved, but the main intuition remains the same.

premium contribution to the derivative position that is proportional to the risk premium for instantaneous volatility shocks, where the proportionality coefficient is captured by the derivative's vega. In contrast, the second term captures changes in the value of the portfolio that are due to second-order variations in returns and the level of the volatility. This term has finite variation and therefore directly identifies a particular time-varying risk premium component in the excess return of the derivative position. It consists of a weighted sum of second order sensitivities collected in matrix (3), with weights that are the time-varying volatility, volatility of volatility and volatility feedback coefficients V_{rrt}, V_{vvt} and V_{rvt} in matrix (1). In particular, the risk premium of the derivative depends on the level of the volatility through the Gamma-sensitivity of the derivative price to shocks in the underlying asset's valuation, which gives rise to a Gamma-Vega relationship in derivative risk premia. In contrast, all other components of the derivative risk premium depend either on the volatility of volatility or the covariation of returns and volatility, which are empirically only partly correlated with the level of the volatility.

In summary, in addition to a precise specification of instantaneous volatility risk premia, an accurate model for the risk premia of volatility derivatives needs to rely on an empirically plausible state process for the symmetric positive definite matrix (1). We discuss in the next section how an approach based on matrix state processes can give rise to natural alternative specifications for matrix (1) having distinct theoretical and empirical properties.

A.2. Main Issue

$K-$factor affine models specify every element of matrix (1) as a linear combination of underlying factors with affine state dynamics, i.e.,

$$\boldsymbol{V}_t := \sum_{k=1}^{K} X_{kt} \boldsymbol{V}_k \ , \tag{4}$$

with scalar stochastic factors X_{kt} and constant parameter matrices $\boldsymbol{V}_k$. The key feature of benchmark affine models defined on standard state spaces, such as multi-factor Bates (2000)-type models, is that in order to ensure positivity of matrix (1) they typically specify all factors X_{kt} as positive (and typically independent). In parallel, they directly ensure positive definiteness of each matrix $\boldsymbol{V}_k$ with suitable parameter constraints.[4] This choice de facto constraints matrix (1) to be a convex combination of matrices $\boldsymbol{V}_k$ once the level of the volatility is fixed.

In this paper, we follow a different approach based on a matrix state process X_t defined on a state space of $n \times n$ symmetric and positive semi-definite matrices. Therefore, we specify volatility matrix (1) as affine in the matrix state X_t,

$$\boldsymbol{V}_t := \begin{bmatrix} tr(X_t \tilde{D}' \tilde{D}) & tr(X_t \tilde{R}' \tilde{Q}) \\ tr(X_t \tilde{R}' \tilde{Q}) & tr(X_t \tilde{Q}' \tilde{Q}) \end{bmatrix}, \tag{5}$$

where $\tilde{D}$, $\tilde{R}$, $\tilde{Q}$ are corresponding $n \times n$ matrices of parameters. By construction, the diagonal elements of matrix (5) are non-negative. Therefore, to ensure positive definiteness of this matrix, one just has to ensure that parameter matrix $\tilde{D}'\tilde{D} - \tilde{R}'\tilde{R}$ is positive definite.[5] Clearly, given the matrix affine form (5), it is always possible in these models to write:

$$\boldsymbol{V}_t = \sum_{i=1}^{n} \sum_{j=1}^{n} X_{ijt} \begin{bmatrix} (\tilde{D}'\tilde{D})_{ij} & (\tilde{R}'\tilde{Q})_{ij} \\ (\tilde{R}'\tilde{Q})_{ij} & (\tilde{Q}'\tilde{Q})_{ij} \end{bmatrix}, \tag{6}$$

where M_{ij} denotes the $ij-$th component of a generic matrix M. While this additive structure seems superficially identical to the one in equation (4), it is actually very different in terms of

[4]For instance, each factor X_{kt} can be specified as an independent Heston (1993)-type volatility model, which may as well be extended by an additional independent component modelling volatility jumps, while matrix $\boldsymbol{V}_k$ can be specified as the (positive definite) covariance matrix of returns and factor k shocks.

[5]From Cauchy-Schwarz inequality, one directly has $(tr(X_t \tilde{R}' \tilde{Q}))^2 \leq tr(X_t \tilde{Q}' \tilde{Q}) tr(X_t \tilde{R}' \tilde{R})$. Therefore, positivity of matrix (5) always follows if $tr(X_t(\tilde{D}'\tilde{D} - \tilde{R}'\tilde{R})) \geq 0$.

the states for matrix $\boldsymbol{V}_t$ that are attainable. Indeed, note first that not all components X_{ijt} in equation (6) need to be positive, since the positivity condition need only be fulfilled at the matrix level. Second, the various matrices in the sum on the right hand side of equation (6) are not in general all positive definite, either. Lastly, we relax the assumption of conditional volatility factor independence in many benchmark models. Indeed, the different components X_{ijt} of the matrix state process have to be conditionally mutually dependent to ensure positive semi-definiteness of matrix X_t.

Figure 1 illustrates graphically the admissible combinations of variance of variance V_{vvt} and leverage effect V_{rvt} in three-factor Bates (2000)-type models and in a our 2×2 matrix AJD setting of Section I.*B.1.*, for four distinct levels of the returns variance V_{rrt}. For any given level, the admissible regions for matrix (1) in three-factor Bates (2000)-type models form a three-dimensional simplex, a result of the convex combinations in (4). In contrast, the corresponding admissible regions in our matrix AJD are bounded by ellipses, which almost completely contain the corresponding regions of three-factor Bates (2000)-type models.

[Insert Figure 1 about here.]

In summary, matrix state processes are a convenient instrument to specify multi-factor volatility dynamics that imply a well-defined covariance matrix (1). We follow this insight and introduce in the next sections a tractable three-factor volatility model using a 2×2 matrix-valued diffusion of symmetric and positive semi-definite matrices of the form

$$X_t := \begin{pmatrix} X_{11t} & X_{12t} \\ X_{12t} & X_{22t} \end{pmatrix} . \tag{7}$$

B. A Three-Factor Matrix AJD Volatility Model

Positivity of the diagonal elements X_{11t} and X_{22t} in matrix (7) allows us to naturally specify these state variables as (diffusive) variance risk factors. In parallel, we largely make use of out-of-diagonal element X_{12t} to model shocks in return jump intensities and volatility feedbacks that are not perfectly correlated with the diffusive volatility.

B.1. Matrix State Dynamics

We specify a tractable state dynamics in our model, starting from the Wishart diffusion of Bru (1991).

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

$$dX_t = [\beta^* Q'Q + M^* X_t + X_t (M^*)']dt + \sqrt{X_t} dB_t^* Q + Q'(dB_t^*)' \sqrt{X_t} \ , \tag{8}$$

where $\beta^* \geq 1$, M^*, Q are 2×2 parameter matrices and B^* is a 2×2 standard Brownian motion under physical measure $\mathbb{P}$. $\sqrt{X_t}$ denotes the symmetric square root of X_t.[6]

Process (8) is positive semi-definite (positive definite) if $\beta^* \geq 1$ ($\beta^* \geq 3$), ensuring that the volatility components are reflected (cannot reach) the zero boundary, see Bru (1991) and Mayerhofer (2014), among others. Note that in the general case 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 (8). When M^* and Q are diagonal, vector (X_{11t}, X_{22t}) is an autonomous Markov process with components distributed as independent Heston (1993) volatility process. In this way, one can nest the two-factor state process in Bates (2000) model within the state dynamics (8).

[6]Give a symmetric positive semi-definite matrix X_t, $\sqrt{X_t}$ is the uniquely defined symmetric and positive semi-definite matrix such that $\sqrt{X_t}\sqrt{X_t} = X_t$.

B.2. Return Dynamics

We specify the return dynamics using the following affine jump-diffusion.

Assumption 2 *Under the physical probability measure $\mathbb{P}$, the dynamics of price process S_t is given by*[7]*:*

$$\frac{dS_t}{S_{t-}} = \mu_t dt + tr(\sqrt{X_t} dZ_t^*) + k^* dN_t^* \ , \tag{9}$$

where X_t follows the dynamics (8),

$$Z_t^* = B_t^* R + W_t^* \sqrt{I_2 - R'R} \ , \tag{10}$$

with W^* another 2×2 standard Brownian motion, independent of B^*, and R a 2×2 matrix such that $I_2 - R'R$ is positive semi-definite. Return jumps follow a compound Poisson process $k^* dN_t^*$ with jump intensity $\lambda_t^* = tr(\Lambda^* X_t)$, for a 2×2 matrix Λ^* and an iid jump size k^*. The distribution of log return jumps $J^* := ln(1 + k^*)$ is a double exponential with parameter $\lambda^{*+}, \lambda^{*-} > 0$ and density:

$$f(J^*) = \frac{\lambda^{*+}\lambda^{*-}}{\lambda^{*+} + \lambda^{-*}} \left[e^{-\lambda^{*-} J^{*-} - \lambda^{*+} J^{*+}} \right] \ , \tag{11}$$

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

Under Assumption 2, the covariance matrix (1) of returns and the volatility reads

$$\boldsymbol{V}_t = \begin{bmatrix} tr(X_t) & 2tr(X_t R' \tilde{Q}^*) \\ 2tr(X_t R' \tilde{Q}^*) & 4tr(X_t \tilde{Q}^{*'} \tilde{Q}^*) \end{bmatrix} + E^{\mathbb{P}}(k^{*2}) \begin{bmatrix} tr(\Lambda^* X_t) & 0 \\ 0 & 0 \end{bmatrix} \ , \tag{12}$$

[7] In our specification, the jump process is not compensated. The total expected return under the physical measure is $E^{\mathbb{P}}[\frac{dS_t}{S_{t-}}] = \mu_t + E^{\mathbb{P}}[k^* dN^*]$.

with the 2×2 matrix $\tilde{Q}^* := Q(I_2 + E^{\mathbb{P}}[k^{*2}]\Lambda^*)$.[8] Here, $tr(X_t)$ is the diffusive variance of returns and $E^{\mathbb{P}}(k^{*2})tr(\Lambda^* X_t)$ the jump variance of returns. Note that while matrix component X_{12} does not affect the diffusive variance of returns, it impacts in general on the dynamics of both the jump intensity $tr(\Lambda^* X_t)$ and the correlation between returns and volatility. Hence, whenever matrices $R'\tilde{Q}^*$ and Λ^* are not diagonal, component X_{12} jointly impacts the jump-driven volatility, the jump-driven skewness and the diffusive skewness. At the same time, it does not appear in the diffusive volatility.

Remark 3 *Diagonal matrices $R'\tilde{Q}^*$ and Λ^* give rise to Bates (2000) specification of volatility feedbacks and stochastic intensities. Thus, a diagonal model in Assumption 2 based on diagonal matrices $M^*, \tilde{Q}, R$ and Λ^* gives rise to Bates (2000) model.*[9] *All such diagonal models have independent volatility risks, as well as jump intensities, volatility feedbacks and factor prices that are completely spanned by the diffusive variance risk factors.*

The three-factor return specification in Assumption 2 is only slightly less parsimonious than Bates (2000) model, with three additional parameters. In contrast, a three-factor Bates (2000)-type model implies seven additional parameters. Parsimony of our three-factor specification helps the identification of parameters and risk factors in our two-step estimation procedure.[10]

[8] Matrix (12) is positive semi-definite because correlation matrix R in Assumption 2 is such that $I_2 - R'R$ is positive semi-definite.

[9] In order to formally nest two-factor Bates (2000)- and Heston (1993)-models in our setting, we specify β^* as a diagonal matrix B^* when both $\tilde{Q}^*$ and M^* are diagonal.

[10] Table 1 of the Online Appendix gives a summary of benchmark models related to Assumption 2. We denote by SV_{rq} pure diffusion and by SVJ_{rq} jump diffusion models, according to the numbers r and q of state variables and skewness factors 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.

C. *Stochastic Discount Factor*

In our model, three types of shocks can be priced: (i) diffusive shocks in index returns, (ii) diffusive shocks in the $X-$driven volatility dynamics and (iii) jump-type shocks in index returns. According to Assumption 2, these shocks correspond to the matrix Brownian shocks dW_t^*, 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 probability $\mathbb{P}$ to the risk neutral probability $\mathbb{Q}$. For suitable 2×2 matrix processes $\{\Gamma_{1t}\}$, $\{\Gamma_{2t}\}$ and our double-exponential specification for log return jumps, such a density can take the form:

$$\begin{aligned}\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} &= \exp\left\{tr\left(-\int_0^T \Gamma_{1t}dW_t^* + \frac{1}{2}\int_0^T \Gamma_{1t}'\Gamma_{1t}dt - \int_0^T \Gamma_{2t}dB_t^* + \frac{1}{2}\int_0^T \Gamma_{2t}'\Gamma_{2t}dt\right)\right\} \\ &\quad \times \prod_{i=1}^{N_T^*} \exp\left\{-(\lambda^- - \lambda^{*-})J_i^{*-} - (\lambda^+ - \lambda^{*+})J_i^{*+} + \ln\left(\frac{1/\lambda^{*-}+1/\lambda^{*+}}{1/\lambda^- + 1/\lambda^+}\right)\right\}, \quad (13)\end{aligned}$$

where the second line defines the change of measure for return jump size distributions and we have already assumed for identification purposes identical functional forms of physical and risk-neutral intensities, consistently with identification Assumption 5 in Section II. *C.2.* below.[11] This choice implies a double exponential distribution with density (11) for return jumps, having parameters λ^{*+}, λ^{*-} and λ^+, λ^- under the physical and the risk neutral distribution, respectively.

Consistently with Leippold and Trojani (2008) and Mayerhofer (2014), we can price the volatility shocks in Assumption 2, driving the risk factors in our model, with a stochastic discount factor that preserves an affine dynamics under the risk neutral probability measure.

[11] Filtration $(\mathcal{F}_T)_{T\geq 0}$ is generated by all Brownian and Poisson-type shocks in the joint dynamics of returns and volatility factors.

A convenient specification preserving the affine structure under the pricing measure is:

$$\Gamma_{1t} = \sqrt{X_t}\Delta + \delta\frac{1}{\sqrt{X_t}} \ , \tag{14}$$

$$\Gamma_{2t} = \sqrt{X_t}\Gamma + \frac{1}{2\sqrt{X_t}}(\beta^* - \beta)Q' \ , \tag{15}$$

with δ a scalar and Γ, Δ being 2×2 parameter matrices. However, it is important to realize that knowledge of the specific form of Γ_{1t} is not necessary to identify the full state dynamics of X_t under our estimation approach. This feature allows us to estimate the time series of volatility states X_t (and the variance risk premium) separately from Γ_{1t} (and the equity premium). In the following sections, we discuss in more detail the resulting change of measure, admissibility conditions and model-implied risk premia.

C.1. Affine Market Price of Volatility Risk

Assumption 4 *The change of measure from the physical probability $\mathbb{P}$ to the risk neutral probability $\mathbb{Q}$ is such that:*[12]

$$dB^* = dB - \Gamma_{2t}dt, \tag{16}$$

with B is a 2×2 standard Brownian motion under the risk neutral probability measure, matrix Γ_{2t} in (15) and parameter constraints implying either $\min(\beta, \beta^*) \geq 3$ or $\beta^* = \beta$.

When $\min(\beta^*, \beta) \geq 3$, process X is almost surely positive definite under both probabilities $\mathbb{P}$ and $\mathbb{Q}$. Such positive definiteness is necessary if one wants to specify a well-defined extended affine market price of risk in the matrix AJD setting.[13] Therefore, for all parameter choices

[12]The inverse of a positive definite matrix X is the matrix $\frac{1}{X}$ such that $I_n = X\frac{1}{X} = \frac{1}{X}X$.

[13]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 defined on standard state spaces.

such that $\min(\beta^*, \beta) < 3$, we need to impose the additional constraint $\beta^* = \beta$ to obtain a well-defined pricing kernel. This constraint gives rise to a completely affine market price of risk in the matrix AJD setting. Under Assumption 4 the $X-$dynamics under the risk neutral probability is again affine and given by:[14]

$$dX_t = [\beta Q'Q + MX_t + X_t M']dt + \sqrt{X_t} dB_t Q + Q' dB_t' \sqrt{X_t} \ , \tag{17}$$

where

$$M^* = M + \Gamma Q \ . \tag{18}$$

An important feature of Assumption 4 is that the price of a shock in any of the risk factors can depend on all factors X_{11}, X_{22} and X_{12}. Specifically, using the notation $D := \Gamma Q$, we obtain the instantaneous factor risk premia:

$$\frac{1}{dt}(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})[dX_{11t}] = (\beta^* - \beta)(Q'Q)_{11} + 2D_{11}X_{11t} + 2D_{12}X_{12t} \ , \tag{19a}$$

$$\frac{1}{dt}(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})[dX_{12t}] = (\beta^* - \beta)(Q'Q)_{12} + (D_{11} + D_{22})X_{12t} + D_{21}X_{11t} + D_{12}X_{22t} \ , \tag{19b}$$

$$\frac{1}{dt}(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})[dX_{22t}] = (\beta^* - \beta)(Q'Q)_{22} + 2D_{22}X_{22t} + 2D_{21}X_{12t} \ . \tag{19c}$$

When matrix ΓQ is not diagonal, the risk premium of diagonal factor X_{iit} is an affine function only of X_{iit} and X_{12t}, while the risk premium of out-of-diagonal factor X_{12t} is affine in all components of state matrix X_t. In contrast, when ΓQ is diagonal all risk premia are disconnected and each factor premium is linear in the factor level alone, a situation emerging, e.g., also in three-factor Bates (2000)-type models. Note that all risk premia can contain a constant component whenever $\min(\beta, \beta^*) \geq 3$, which is the setting where an extended affine

[14]In our empirical analysis, we also explored more general physical dynamics for X_t having an unconstrained constant drift component $\Omega\Omega'$ in equation (17). Such an unconstrained estimation of the model produced an estimated matrix $\Omega\Omega'$ that is statistically indistinguishable from $\beta^* Q'Q$. We thank an anonymous referee for suggesting the study of this possible extension of our state dynamics.

market price of risk can be supported using a matrix state space. In all other cases, $\beta = \beta^*$ and the underlying market price of risk is completely affine.

Identification. So far we have specified all parameters relevant for the first step of our estimation approach. For identification, we require matrices M, M^* to be non-singular and lower triangular. Matrix Q is required to be non-singular and upper triangular. Further, matrix R is the uniquely defined upper triangular Cholesky decomposition of $R'R$ and is constrained to ensure that matrix $I_2 - R'R$ is positive semi-definite. Finally, Λ must be a non singular upper triangular matrix; see Section II.*B.* of the Online Appendix for a derivation of these identification conditions.

C.2. Variance Swap Payoffs and Variance Risk Premia

We characterize the risk premia of variance swap contracts that can be synthesized 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:

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

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.[15] Since the variance risk premium $VRP_t(\tau)$ is the difference of the $\mathbb{P}$ and $\mathbb{Q}$ expectations of $RV_{t+\tau}(\tau)$,

[15] See Schneider and Trojani (2019b) for details.

Assumptions 1 and 4 give:

$$VRP_t(\tau) = (E_t^{\mathbb{P}} - E_t^{\mathbb{Q}}) \left[\frac{1}{\tau} \int_t^{t+\tau} tr(X_s) ds \right] + (E_t^{\mathbb{P}} - E_t^{\mathbb{Q}}) \left[\frac{2}{\tau} \sum_{t \leq s \leq t+\tau} \mathcal{E}\left(\frac{S_s}{S_{s-}} \right) \right] . \quad (21)$$

The first term on the right hand side is the diffusive variance risk premium, i.e., the premium contribution from continuous return shocks. The second term is the jump variance risk premium, i.e., the premium contribution from return jumps. Under our specification (13) of the risk-neutral density, both terms are affine in state X_t. While the diffusive variance risk premium is completely inferable from the risk premia of the diagonal components of state process X_t, the jump variance risk premium depends on the chosen specification of jump intensities and jump distributions under the physical and risk neutral probabilities.

In our empirical analysis, we identify risk-neutral jump intensities and jump size distributions from the first-step estimation of the physical and risk-neutral matrix state dynamics, while relying exclusively on option data information. In the second step of our empirical approach, we identify the physical jump expected variance in the variance risk premium (21) using the variance swap payoffs in equation (20). Given the non identifiability of λ_t^* and $E^{\mathbb{P}}[\mathcal{E}(S_t/S_{t-})]$ from the term structure of variance swap payoffs, we assume in the sequel identical functional forms of physical and risk-neutral intensities for identification.[16]

Assumption 5 *Physical and risk neutral jump intensities are such that* $\Lambda^* = \Lambda$.

Assumption 5 is a fairly standard identification assumption in the literature, see for example Bates (2000), Carr and Wu (2017) or Ait-Sahalia, Karaman and Mancini (2012), and is consistent with the risk-neutral density specification in Section I. *C.*[17] With the parameterization

[16] We investigated different affine specifications of physical jump intensities and found that a specification with physical intensities proportional to risk-neutral intensities is preferred by the out-of-sample analysis of variance swap payoffs.

[17] It is important to realize that under Assumption 5 the physical and risk-neutral intensity dynamics

$\mathbb{E}^{\mathbb{P}}[\mathcal{E}(1+k)] = \beta_\lambda^* \mathbb{E}^{\mathbb{Q}}[\mathcal{E}(1+k)]$, Assumption 5 yields:

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

The resulting closed-form (affine) expression for the variance risk premium is detailed below.

Proposition 1 *Given Assumptions 1, 2, 4 and 5, the variance risk premium for time to maturity $\tau > 0$ is given by:*

$$VRP_t(\tau) = VRP_t^c(\tau) + VRP_t^d(\tau) \ , \tag{22}$$

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

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

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

$$X_\infty^{\mathbb{Q}} M + M' X_\infty^{\mathbb{Q}} = \beta Q' Q \ ; \qquad X_\infty^{\mathbb{P}} M^* + M^{*\prime} X_\infty^{\mathbb{P}} = \beta^* Q' Q \ , \tag{23}$$

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

Remark 6 *(i) In non-diagonal models, the term structure of variance risk premia depends on shocks to factor X_{12}, which are partially separated from shocks to the spot volatility. (ii) It is easy to see that when matrices Q, M, M^* and Λ are diagonal, both $VRP_t^c(\tau)$ and $VRP_t^d(\tau)$ depend only on X_{11} and X_{22}, which induces a perfect correlation between shocks*

---

are still different, because the physical and risk-neutral dynamics of state process X are different. As a consequence, jump intensity risk is priced in our model and contributes to variance risk premia.

in variance risk premia and the spot (diffusive) variance. This situation emerges in Bates (2000)-type models.

C.3. Model-Implied Equity Premium

The instantaneous equity premium in our model is given by

$$E_t^{\mathbb{P}}[dS_t/S_{t-}] - (r-q) = \mu_t + E_t^{\mathbb{P}}[k^* dN_t^*] - (r-q) \ , \tag{24}$$

with r the risk-free rate and q the index dividend yield. Given the risk-neutral density specification in equation (13) with a market price of volatility (15), the jump-related component $E_t^{\mathbb{P}}[k^* dN_t^*]$ of the equity premium is affine in state X_t. The concrete specification (14) for the market price of W^*-shocks further supports an affine dependence also for the diffusive equity premium component $\delta + tr(\Delta X_t)$:[18]

$$\mu_t - \delta - tr(\Delta X_t) = r - q - E_t^{\mathbb{Q}}[k dN_t] \ , \tag{25}$$

and

$$E_t^{\mathbb{P}}[dS_t/S_{t-}] - (r-q) = \delta + tr(\Delta X_t) + E_t^{\mathbb{P}}[k^* dN_t^*] - E_t^{\mathbb{Q}}[k dN_t] \ . \tag{26}$$

We test this affine specification of the equity premium in our model in a separate third estimation step, detailed in Section I.*C.6.* below. In Section II.*G.* of our empirical analysis we test the statistical and economic implications of this specification with respect to those implied by unconstrained predictive regressions of future index returns on state X_t.

[18]Section **I.** *C.* of the Online Appendix provides a formal proof of the fact that under Assumptions 1 and 4 these choices of market price of risk specifications Γ_{1t} and Γ_{2t} imply a well defined density process $\{\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_T}\}_{T\geq 0}$.

C.4. Option Valuation

Assumption 2 and Assumption 5 yield closed-form risk-neutral transforms, 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:

$$\Psi(\tau;\gamma) := E_t\left[\exp\left(\gamma Y_T\right)\right] = \exp\left(\gamma Y_t + tr\left[A(\tau)X_t\right] + B(\tau)\right) \ , \tag{27}$$

where $\tau = T - t$, with the 2×2 matrix $A(\tau)$ and the scalar $B(\tau)$ given in closed form in Section **I.***A.* of the Online Appendix.[19]

C.5. Link between parameter β_λ^* and variance swap payoffs

With the exception of parameter β_λ^*, all states and parameters in Proposition 1 are identifiable from the panel of index option prices. This feature motivates a simple estimation procedure for parameter β_λ^*, based on the model-free payoffs of variance swaps.

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:

$$\begin{aligned} RV^e_{t+\tau}(\tau) := RV_{t+\tau}(\tau) - E_t^{\mathbb{Q}}[RV_{t+\tau}(\tau)] \ &= \ \frac{2}{\tau}\left[\int_0^\infty \frac{O_{t+\tau}(K)}{K^2}dK + \int_t^{t+\tau}\left(\frac{1}{F_{s-}} - \frac{1}{F_t}\right)dF_s\right] \\ &\quad -\frac{2}{\tau}\int_0^\infty \frac{E_t^{\mathbb{Q}}[O_{t+\tau}(K)]}{K^2}dK \ , \end{aligned} \tag{28}$$

[19] In contrast to Bates (2000)-type models, the computation of the risk neutral transform cannot be reduced to calculations that involve only scalar exponential and logarithmic functions, because $A(\tau)$ and $B(\tau)$ depend on a matrix exponential and a matrix logarithm. To overcome these computational challenges and make the estimation feasible, we employ a fast evaluation scheme based on the efficient Cosine-Transform method in Fang and Oosterlee (2008) and perform the calculations on a cluster of 80 processors. To solve the problem of the multi-valued complex logarithm, which has been documented in Lord and Kahl (2010) for the scalar case, we employ the matrix rotation count algorithm. See Gruber (2015) for a detailed discussion of the computational aspects of our model.

where $O_{t+\tau}(K)$ is the payoff of an out-of-the-money European option on index futures, with maturity τ and strike price K. We compute payoff $RV^e_{t+\tau}(\tau)$ in a model-free way, i.e. using a panel of SPX options and a time of series high-frequency S&P 500 index futures prices to approximate the integrals in (28) at a weekly frequency. Note that by definition the expectation of this payoff is the variance risk premium in Proposition 1, which is itself linear in parameter β^*_λ. Therefore, we can subtract from payoff (28) the relevant terms in Proposition 1, in order to obtain an arbitrage-free regression setting that allows us to directly estimate parameter β^*_λ. The explicit form of this regression is detailed in the next result.

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

$$\begin{aligned} Y_{t+\tau}(\tau) &:= RV^e_{t+\tau}(\tau) - VRP^c_t(\tau) - 2E^{\mathbb{Q}}[\mathcal{E}(1+k)]tr[\Lambda(X^{\mathbb{Q}}_\infty + A^{\mathbb{Q}}_\tau(X_t - X^{\mathbb{Q}}_\infty))] \ , \quad (29) \\ U_t(\tau) &:= 2E^{\mathbb{Q}}[\mathcal{E}(1+k)]tr[\Lambda(X^{\mathbb{P}}_\infty + A^{\mathbb{P}}_\tau(X_t - X^{\mathbb{P}}_\infty))] \ . \quad (30) \end{aligned}$$

Given Assumptions 2-5 and maturities $\tau_1 < \ldots < \tau_n$, the following is an arbitrage-free linear regression model,

$$\begin{pmatrix} Y_{t+\tau_1}(\tau_1) \\ \vdots \\ Y_{t+\tau_n}(\tau_n) \end{pmatrix} = \beta^*_\lambda \begin{pmatrix} U_t(\tau_1) \\ \vdots \\ U_t(\tau_n) \end{pmatrix} + \begin{pmatrix} \eta_{t+\tau_1}(\tau_1) \\ \vdots \\ \eta_{t+\tau_n}(\tau_n) \end{pmatrix} \ , \quad (31)$$

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 $U_t(\tau)$ are computable from variance swap payoffs $RV^e_{t+\tau}(\tau)$ and a first-step estimation of parameters $\beta, \beta^*, M, M^*, Q, \Lambda, \lambda^+, \lambda^-$ 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.

C.6. Identification of the Equity Premium Parameters

As discussed in Section *C.3.*, we obtain the full state dynamics, all parameters necessary for option pricing and the variance risk premium without requiring a complete specification of the price of risk Γ_{1t}. This helps us to estimate and test the concrete equity premium specification (26) in an independent third estimation step, without impacting the estimated volatility dynamics.

We denote for brevity by $\beta^*_\Delta := E^{\mathbb{P}}[k^*]/E^{\mathbb{Q}}[k]$ the parameter for the expected physical jump size. Using equations (A-5)–(A-7) from the Online Appendix, together with equation (26), we then obtain the following explicit equity premium identity, which is the basis of our estimation approach:

$$\begin{aligned}\frac{1}{\tau}E^{\mathbb{P}}\left[\ln\left(\frac{S_{t+\tau}}{S_t}\right)\right] - \underbrace{(r-q) + E^{\mathbb{Q}}[k]\, tr\left(\Lambda\cdot A^{\mathbb{Q}}_\tau(X_t - X^{\mathbb{Q}}_\infty)\right)}_{Z_t(\tau)}\\ = \ \delta + \beta^*_\Delta \underbrace{E^{\mathbb{Q}}[k]\, tr\left(\Lambda\cdot A^{\mathbb{P}}_\tau(X_t - X^{\mathbb{P}}_\infty)\right)}_{W_t(\tau)} + tr\big(\Delta\cdot \underbrace{A^{\mathbb{P}}_\tau(X_t - X^{\mathbb{P}}_\infty)}_{V_t(\tau)}\big),\end{aligned} \tag{32}$$

Identification and positivity of the diffusive equity premium component (25) require Δ to be a symmetric, positive semi-definite 2×2 matrix, and scalar δ to be non-negative. To estimate $\theta := (\delta, \beta^*_\Delta, \Delta^{11}, \Delta^{12}, \Delta^{22})$ using identity (32), an identification assumption is needed, as the three components of the symmetric matrix $V_t(\tau)$ and scalar $W_t(\tau)$ are linearly dependent. In order to allow for the possibility that all volatility factors impact the diffusive equity premium, we thus require for identification Δ to have rank equal to one. Overall, we obtain a linear regression model of the form

$$\ln\left(\frac{S_{t+\tau}}{S_t}\right) - Z_t(\tau) = \delta + \beta^*_\Delta W_t(\tau) + tr(\Delta V_t(\tau)) + \zeta_{t+\tau}(\tau) \tag{33}$$

where error term $\zeta_{t,\tau}(\tau)$ is such that $E_t^{\mathbb{P}}[\zeta_{t+\tau}(\tau)] = 0$. In our empirical analysis, we estimate $\delta, \beta_\Delta^*, \Delta$ with constrained least squares under the constraints $0 \leq \beta_\Delta^* \leq 1$ and $rank(\Delta) = 1$, using index returns for horizons $\tau = 1, \ldots, 6, 9, 12$.

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 CBOE, for the sample period from January 1996 to August 2015 and maturities up to one year.[20] The sample consists of 4948 trading days, which we reduce to 1015 weekly observations (each Wednesday). In order to allow for an out-of sample evaluation of our model, we split these 1015 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 August 2015 with 656 observations. We apply a number of standard filtering procedures as outlined, e.g., in Bakshi, Cao and Chen (1997).

For our first-step estimation of model parameters and volatility factors, we make use of all options with a time to maturity of at least ten days and an absolute Black-Scholes delta between 0.1 and 0.9.[21] On average, this gives about 175 option prices per trading day. Table 2 of the Online Appendix presents a summary of the characteristics of our data set. To calculate weekly model-free variance payoffs as defined in (28), we use of all available options. The delta hedging component in the variance payoff is computed using tick data for the S&P 500 index future traded at the CBOE, obtained from tickdata.com, sampled at

[20] We obtain end-of-day midquotes as simple averages of end-of-day bid and ask call prices. We obtain the risk-free rate from OptionMetrics and force the put-call parity to hold when calculating implied dividend yields. We base our estimation on synthetic forward prices.

[21] Given the challenges in jointly fitting weekly options and longer maturity options with models having time-invariant jump distributions (see Andersen, Fusari and Todorov (2017)), we follow the standard approach in the literature and leave the study of weekly options with matrix AJD for future research. We thank an anonymous referee for raising this issue.

60 second intervals and aggregated over horizons $\tau = 1 \dots 12$ months to realized volatilities.

In the first step of the estimation, we use the panel of SPX in-sample observations to estimate the model parameters Q, M, M^*, R, Λ, λ^+, λ^-, β, β^*, together with the filtered time series of risk factors X_{11t}, X_{12t} and X_{22t}. We estimate the model parameters by maximizing the likelihood defined on the option-implied volatility forecasting errors in an extended Kalman filter, using an exact second-order state dynamics and a linearized observation equation for implied-volatilities.[22] For the observation equation, we assume Gaussian errors and account for a potential autocorrelation of implied volatility errors. Details on the estimation procedure are provided in Section II.*A.* of the Online Appendix.

In the second step, we estimate the jump size proportionality parameter β^*_λ from an arbitrage-free linear regression of realized variance swap payoffs on the time series of model-implied variance risk premia, as outlined in Proposition 2. Precisely, we first compute synthetic variance swap payoffs for maturities $\tau_1, \tau_2, \dots, \tau_n = 1, 2, 3, 4, 5, 6, 9, 12$ months and construct a time series of in-sample weekly observations for variables $(Y_{t+\tau_i}(\tau_i), U_t(\tau_i))_{i=1,\dots,n}$ in linear model (31), where $t = 1, \dots, N$ and the in-sample sample size is $N = 359$. We then estimate the single unknown parameter β^*_λ with a pooled linear regression.[23]

In a third step, we estimate with constrained least squares parameters $\delta, \beta^*_\Delta, \Delta$ in equity premium identity (33), based on constrained pooled predictive regression for horizons $\tau =$

[22]While this filtering approach is not exact and may imply a discretization bias, we feel it is a reasonable compromise between computational complexity and accuracy in our setting with more than 150 options per day. In principle, one could try to develop an unscented Kalman Filter or a particle filter for our matrix AJD, in order to propagate the underlying matrix state dynamics. As such approaches would considerably increase the computational burden for the model estimation in our current setting, we leave it as an interesting direction for future research. We thank an anonymous referee for raising this point.

[23]This approach allows us to estimate parameter β^*_λ consistently with a simple pseudo Maximum Likelihood approach, under the assumption of a correctly specified conditional variance risk premium. Clearly, this approach is less efficient than a Maximum Likelihood estimation incorporating the information from the full conditional distribution of the realized variance excess payoffs. As correctly specifying a full likelihood for realized variance would require additional not obvious assumptions, we feel that a pseudo-maximum likelihood approach is naturally motivated also from a robustness perspective. We thank an anonymous referee for raising this point.

$1, 2, 3, 4, 5, 6, 9, 12$ months, using our full sample January 1996 to August 2015.

B. Option Pricing Performance and Model Fit with Respect to Bates-Type Models

We quantify the option pricing performance and the statistical fit of our matrix AJD model (GT^2) in relation to the benchmark models in Table 1 of the Online Appendix. These benchmark models include two- and three-factor Heston (1993)- and Bates (2000)-type models, as well as the purely diffusive version of our model, which is obtained for $\Lambda = 0$. For all these models, we produce a joint estimation of physical and risk-neutral volatility dynamics.[24] Since the models have different numbers of parameters and state space dimensions (see in Table 1 of the Online Appendix), we control for overfitting using our in-sample (1996-2002) and out-of-sample (2003-2015) periods. All our parameter estimates are exclusively based on in-sample weekly data. We then use the full data set to compute the filtered states, in- and out-of sample option prices, implied volatilities and finally proxies of pricing accuracy, such as the mean weekly absolute implied volatility error (*MAIVE*). We furthermore compare the statistical fit of different models using the in- and the out-of-sample value of the average likelihood function and information criteria. Table 2 summarizes the pricing performance and statistical fit across models.

[Insert Table 2 about here.]

The results indicate that our model GT^2 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 19.4% in-sample and 31.3% out-of-sample, using the *MAIVE* metric. The improvement of the in-sample (out-of-sample) value of the likelihood function is 4.7% (10.4%) and is statistically significant at conventional levels.

[24] In Section II.*F.* below, we compare the performance of our model also with respect to recent flexible specifications of risk-neutral probabilities. There, we estimate risk neutral parameters and hidden volatility states with a penalized nonlinear least squares approach to allow for comparison.

Our model also improves with respect to a three-factor Bates (2000)-type model (SVJ_{30}), despite having four parameters less.[25]

The in- and out-of sample performances of our GT^2 model are virtually identical, with *MAIVE* s of 53.7 and 54.8 volatility basis points. In contrast, the out-of-sample *MAIVE* of the SVJ_{20} (SVJ_{30}) model is 19.8% (4.6%) higher. In summary, our model improves on benchmark two- and three-factor Bates (2000)-type models, both in- and out-of-sample, in a way that clearly indicates no relevant degree of overfitting.

C. Estimated Parameters

The estimated model parameters directly capture the dynamic interactions between volatility risk factors and their relation with the price of the smile. Table 1 presents the parameter estimates for our model, while Table 3 of the Online Appendix presents our estimates in the context of different benchmark models.

[Insert Table 1 about here.]

All estimated parameters are significant at standard significance levels. Since we cannot reject the null hypothesis $\beta^* = \beta$, the data support a completely affine specification of the market price of risk in (16). The out-of-diagonal element is strongly significant in all parameter matrices M, M^*, Λ, R and Q, 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 $\lambda^- \ll \lambda^+$ reflect the negative risk-neutral skewness of the distribution of log return jumps.

The large negative coefficient M^*_{22} indicates that factor X_{22} has the strongest autonomous mean reversion and the lowest persistence of all factors in our model. Since $Q_{22} \gg Q_{11}$, factor

[25]Our model is preferred to benchmark two- and three-factor Bates models also according to standard model selection criteria, such as Akaike Information Criterion (AIC).

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). As X_{12} is positive most of the time, with stronger excursions during phases of market distress, this feature induces a mutually-exciting behaviour of factors 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 factor 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 also 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. X_{11} is therefore a low-frequency component of the total variance, featuring mutually-exciting behaviour with X_{12} in periods of distress.

The negative coefficients M^*_{11} and M^*_{22} indicate that factor X_{12} has an autonomous mean-reversion between the one of the high- and low-frequency factors 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. In combination with its dominating role in the leverage effect, it has the interpretation of a risk factor for skewness risk.

D. Dynamics of Volatility Factors

The times series of factors X_{11}, X_{22} and X_{12} are presented in Figure 2 and imply half-lives of 1.212, 0.25 and 0.108 years, see also Table 5 of the Online Appendix. These factors are dynamically correlated and mutually exciting, with unconditional correlations of about 0.25 between X_{11} and X_{22} and of about 0.5 for the pairs (X_{11}, X_{12}) and (X_{12}, X_{22}).

[Insert Figure 2 about here.]

The diffusive variance $tr(X_t)$ is the sum of two positive components with significantly different levels of persistence and volatilities of volatility. Here, X_{22} has on average a larger contribution to the diffusive variance than X_{11}, besides being more volatile and less persistent. Factor X_{12} is positive most of the time and takes – as X_{22} – the largest values in periods of significant turmoil, as during the 2008-2009 financial crisis. Interestingly, while the most persistent factor X_{11} also spikes substantially in periods of distress, it often does so with a lag with respect to factors X_{12} and X_{22}. Figure 3 of the Online Appendix also documents that the relative importance of each factor in our multivariate state dynamics can vary substantially. For instance, Panel A shows that the fraction of persistent diffusive variance generated by factor X_{11} can vary essentially from zero to one.[26]

Panel B similarly shows that the relative importance of factor X_{12} also varies a lot. Given the predominant role of factor X_{12} in periods of market distress, it appears that this factor is a key driver of the variance dynamics in these periods.

[26]Indeed, this fraction is basically zero during Greenspan's 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. X_{12t} can be as large as 50% of the diffusive volatility during phases of market turmoil, e.g, shortly after the devaluation of the Thai Bhat, the beginning of the Russian crisis, the Lehman default and the US downgrade.

D.1. Interpretation of Factors X_{11}, X_{12} and X_{22} as Option-Implied Risks

Even though X_{11}, X_{12} and X_{22} are nominally latent factors, we are able to link them to directly observable characteristics of the option-implied volatility smile, in order to gain additional economic interpretation. We find that the high-frequency factor X_{22} is closely related to the 30-day at-the-money implied variance, with a correlation of 91% and virtually identical persistence properties. Factor X_{12} is closely related to the 30-day option-implied skew, with a correlation of -89%. Similarly, X_{11} closely targets the long end of the implied volatility curve and has a correlation of 91% with the 12 months at-the-money implied variance. Figure 3 presents the time series of model-implied factors 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 3 about here.]

D.2. The Price of the Smile

Our estimation results in Table 1 imply the following instantaneous risk premium for $X_{11}-$, $X_{22}-$ and $X_{12}-$shocks, as defined in equation (19):

$$\frac{1}{dt}(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})[dX_{11t}] = -1.0776X_{11t} \ , \tag{34a}$$

$$\frac{1}{dt}(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})[dX_{12t}] = -0.6283X_{11t} - 0.5388X_{12t} \ , \tag{34b}$$

$$\frac{1}{dt}(E_t^{\mathbb{P}} - E_t^{\mathbb{Q}})[dX_{22t}] = -1.2566X_{12t} \ . \tag{34c}$$

Therefore, the more persistent factors X_{11} and X_{12} are actually risk premium factors that completely span the price of the smile. Note that the risk premium of each factor is at least as persistent as the factor itself: The most persistent risk premium is the one for $X_{11}-$shocks, followed by the risk premium for $X_{12}-$shocks, while the most transient risk premium is the

risk premium for $X_{22}-$shocks. These instantaneous factor risk premia imply the form of the risk premium for a shock in the spot diffusive variance $V_t^c := Var_t[dS_t^c/S_t^c] = X_{11t} + X_{22t}$:

$$\frac{1}{dt}(E^{\mathbb{P}} - E^{\mathbb{Q}})\big[dV_t^c\big] = \frac{1}{dt}(E^{\mathbb{P}} - E^{\mathbb{Q}})\big[dX_{11t} + dX_{22t}\big] = -1.0776X_{11t} - 1.2566X_{12t} \ .$$

The negative and time-varying market price of spot variance risk is fully explained by the two more persistent risk premium factors X_{11} and X_{12} that span the price of the smile, where each of these factors loads with a similar weight on the price of the spot variance. Therefore, our estimated model embeds situations where high-frequency shocks in the spot volatility can materialize without immediately affecting the market price of the volatility, as this market price varies only when the lower frequency volatility and skewness factors X_{11} and X_{12} vary.

From this evidence, we conclude that the the price of the smile, which is a linear function of factors X_{11} and X_{12} alone, is directly related to observable proxies of long-term option-implied volatility and short term option-implied skewness. These proxies contain a substantial component that is orthogonal to the high-frequency volatility factor X_{22}. They are also naturally related to the price of particular option portfolios, such as calendar spreads or risk reversals, which are designed to profit from changes in long-term option-implied volatilities and short-term option-implied skewness. According to our findings, we can interpret the prices of these portfolios as observable risk premium factors that span the price of the smile and are naturally related to the term structure of variance risk premia.

E. Variance Risk Premia and their Term Structure

The model-implied variance risk premia for horizons $\tau = 1, 12$ months are plotted in Figure 4, together with their difference, as a proxy for the slope of their term structure. 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, variance risk premia are largest in absolute value.

The term structure of variance risk premia is usually downward-sloping. However, it can also be upward sloping during short periods of time that correspond to about 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.

[Insert Figure 4 about here.]

F. Option Pricing Performance and Volatility Fit with Respect to Three-Factor Double-Jump Affine Specifications of Risk-Neutral Dynamics

A natural question is how does our matrix AJD specification of the state dynamics perform with respect to recent flexible specifications of the risk neutral-dynamics for standard state spaces, which incorporate both an unspanned component for jump skewness and co-jumps between returns and volatility, such as the option pricing model in Andersen, Fusari and Todorov (2015b) (AFT). Therefore, we closely follow the estimation procedure in Andersen et al. (2015b) for estimating the model's hidden states and risk-neutral parameters from option panels, and we apply this procedure for the estimation of their model and our model using data from our in-sample period 1996-2002.[27]

[27]We are grateful to Nicola Fusari for making available his Matlab Toolbox for Option Pricing on which we based our estimation of the AFT-model. We also thank Kai Wang for excellent research assistance.

Their estimation approach relies on a regularized nonlinear least squares estimation that balances the tradeoff between the model's option pricing and spot volatility fits. Hence, conditional on a parameter vector θ and a time series of hidden states $\{X_t\}$, the objective function for this estimation problem is:

$$E_{WGHT}(\theta, \{X_t\}) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(E_{PRICE,t} + \lambda E_{RV,t}\right)}\ , \tag{35}$$

where $\lambda > 0$ is a penalization parameter and[28]

$$E_{PRICE,t} := \sum_{k=1}^{N_t}(IV_{k,t} - \widehat{IV}_{k,t})^2\ ,\ E_{RV,t} := \sqrt{(RV_t - \widehat{RV_t})^2}\ , \tag{36}$$

with N_t the number of option prices observed on day t. In these equations, $IV_{k,t}$ and $\widehat{IV}_{k,t}$ are Black-Scholes option-implied volatilities in the data and as implied by the given model for parameter θ and states $\{X_t\}$, respectively. Similarly, RV_t is the continuous part of the realized volatility on day t, as specified in Appendix B of Bollerslev and Todorov (2011), while $\widehat{RV_t}$ is the diffusive model-implied integrated variance. Table 3 summarizes the quality of the option pricing and the volatility fit resulting from the estimation of AFT model and our GT^2 model with our dataset.[29]

[Insert Table 3 about here.]

We find that both models perform similarly well in terms of the in-sample and out-of-sample weighted objective function (35), both with respect to its option pricing and realized volatility error components.[30] This evidence shows that our matrix AJD setting can produce

[28]Following Andersen et al. (2015b), we make use of a penalization parameter $\lambda = 0.2$.

[29]Table 4 of the Online Appendix also reports the parameters estimated by penalized nonlinear least squares for AFT model and for our matrix AJD model.

[30]Figure 2 of the Online Appendix also documents graphically the very similar fit of realized volatilities resulting from the estimation of these two models with our dataset.

a similarly accurate specification of the risk neutral dynamics and hidden states as AFT-type models, despite having four parameters less in its specification of risk neutral distributions. The hidden states estimated by penalized nonlinear least squares for our model GT^2 are also quite consistent with the filtered states estimated with our extended Kalman filter for matrix AJD.

When setting the volatility weight λ to zero in criterion (35), the NLLS approach also defines a natural benchmark for our extended Kalman filter estimation. We find that while the states estimated by nonlinear least squares are naturally more noisy and volatile, their correlation with the filtered states from the extended Kalman filter is 0.997 for state X_{11}, 0.972 for state X_{12} and 0.992 for state X_{22}; see also Figure 4 of the Online Appendix for more details.

G. Equity Premium Dynamics

The estimated model-implied equity premium dynamics resulting from specification (14)-(15) implies reasonable economic properties, such as a pronounced counter-cyclicality across maturities. Figure 5, Panel A, reports the time series of estimated index equity premia on an annual basis for quarterly horizons.[31] We find that the conditional equity premium variability is large, with a maximum of about 16.5% during the Great Financial Crisis and a minimum of about 0.2% in the Conundrum Period. Plausibility of the estimated equity premium dynamics can further be gauged using model-free lower bounds implied by the cross-section of option prices using the recovery approach in Schneider and Trojani (2019a). Importantly, we find that while the lower bound reported in Fig. 5 is quite volatile and counter-cyclical, the model-implied equity premium virtually never violates it.

[Insert Figure 5 about here.]

[31] Equity premium dynamics for other horizons highlight similar patterns.

The model-implied equity premium dynamics is naturally constrained by the consistency conditions induced on the model's parameters. A model-independent test of an affine equity premium specification can simply estimate unconstrained predictive regressions of index returns on model-implied volatility states. In doing so, we can naturally compare the dynamic equity premium implications of volatility risk factors of different models. Figure 5 reports the time series of equity premia obtained from predictive regressions of index returns on estimated volatility states in the AFT model.[32] While all volatility factors are statistically significant in explaining the equity premium for horizons of at least three months, we find that the resulting dynamics are economically counter-intuitive, with frequent strongly negative equity premia and frequent violations of the model-free lower bound.[33] When imposing ad-hoc exclusion or sign constraints on the coefficients of strictly positive volatility factors, a dynamics with uniformly positive equity premia is obtained, which is more consistent with the model-free lower bound. However, the economic foundation for such constraints is unclear.

III. Conclusions and Outlook

Motivated by the joint tradability of variance and skewness risk in option markets, we make use of a new specification of multi-factor volatility to estimate the hidden risk factors spanning SPX implied volatility surfaces and the risk premia of volatility-sensitive payoffs. We find that SPX implied volatility surfaces are well-explained by three dependent state variables reflecting (i) short- and long-term implied volatility risks and (ii) short-term implied skewness risk. The more persistent volatility factor and the skewness factor support a downward sloping term structure of variance risk premia in normal times, while the most transient

[32]Table 8 in the Online Appendix reports full details on estimated regression coefficients and predictive R^2s. Unconstrained predictive regressions using volatility factors in our model produce a qualitatively similar evidence.

[33]Estimated factor loading of strictly positive volatility factors are also always significantly negative.

volatility factor accounts for an upward sloping term structure in periods of distress. The volatility risk factors in our model also produce under a simple affine specification an economically plausible and counter-cyclical equity premium, which is compatible with model-free lower bounds extracted from option prices.

Our volatility specification based on a matrix state process is instrumental to obtaining a tractable and flexible model for the joint dynamics of returns and volatilities, which improves pricing performance and risk premium modelling with respect to recent three-factor specifications based on standard state spaces. In contrast to most existing arbitrage-free models, a key property of our specification is that it incorporates three dynamically correlated and differently persistent state variables, which (i) are linked to a stochastic skewness and factor prices not spanned by the spot volatility and (ii) produce a sharp identification of the sources of risk traded in option markets.

While we have focused in this paper on a three-factor specification without jumps in volatility, an interesting avenue for future research is the study of arbitrage-free option pricing models with a matrix state space that allows for a multivariate jump component in the joint covariance matrix of returns and volatilities. Such models can naturally incorporate multivariate bursts in volatility and volatility of volatility factors, consistently with the recent empirical evidence in, e.g., Christensen, Oomen and Podolskij (2014) that volatility bursts, rather than return jumps, are a key component of high-frequency return dynamics.

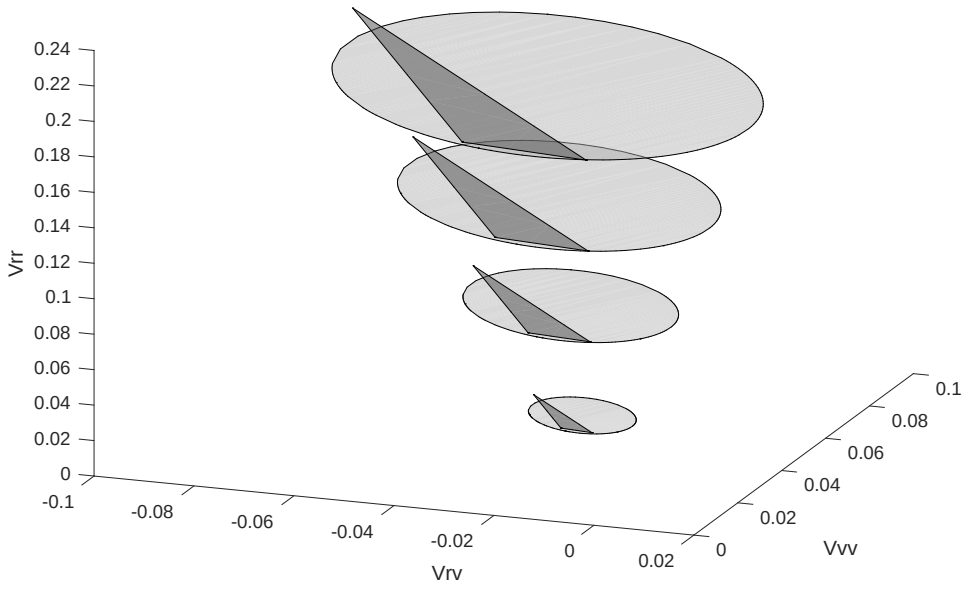
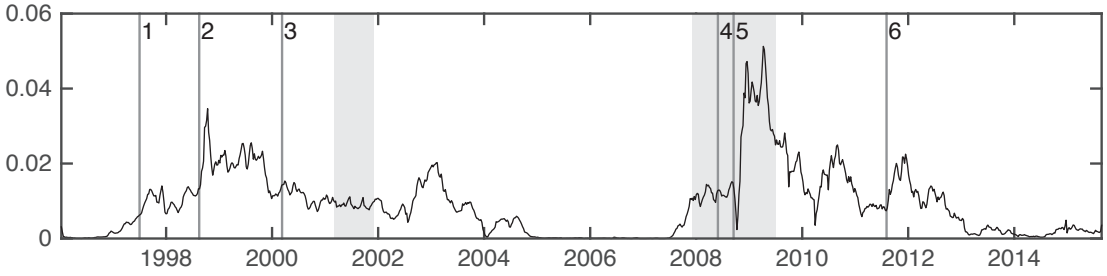
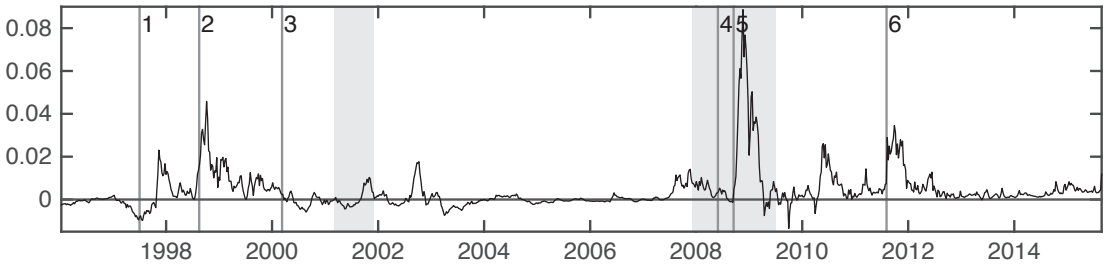


Figure 1: Attainable regions for variance-covariance matrix (1) in a three-factor Bates (2000)-type model and in our 2×2 matrix AJD model of Section I. *B.1.*. For four different levels of return variance V_{rr}, we plot the admissible combinations of variance of variance V_{vv} and leverage effect V_{rv} in a three-factor Bates (2000)-type affine model (dark grey triangles) and our 2×2 matrix AJD model of Section I. *B.1.* (light grey ellipses), making use of the estimated model parameters of Section **II.**.

Panel A: Volatility factor X_{11t}



Panel B: Volatility factor X_{12t}



Panel C: Volatility factor X_{22t}

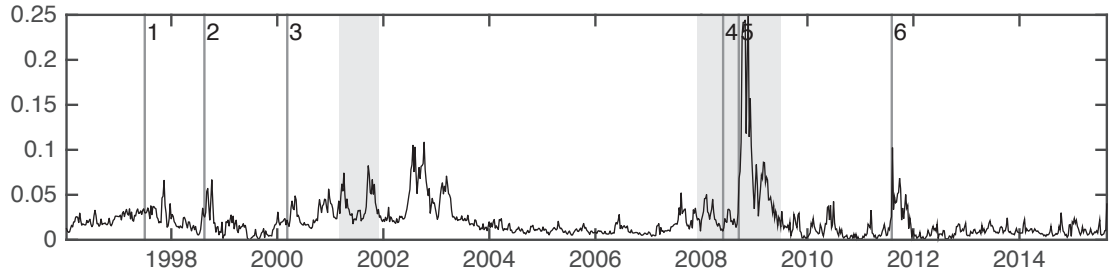
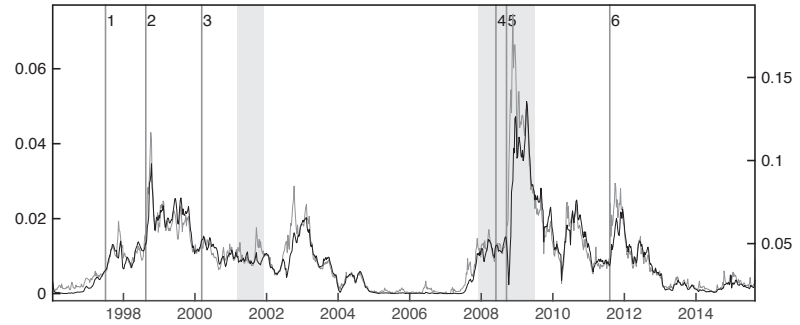


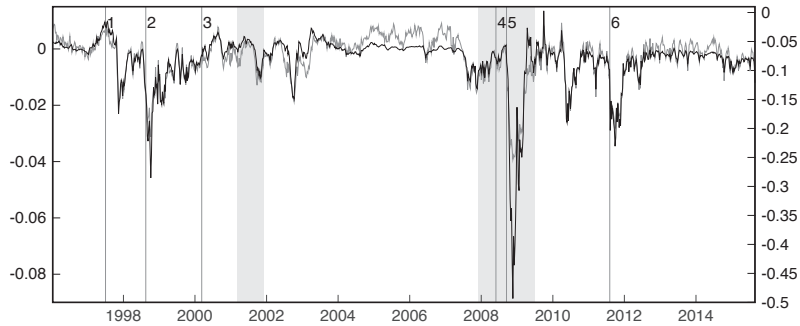
Figure 2: Filtered volatility factors $X_{11t}, X_{12t}, X_{22t}$. Grey areas highlight NBER recessions; vertical lines indicate the following crisis events:

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

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



Panel B: Negative value of filtered factor X_{12} as option-implied skew



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

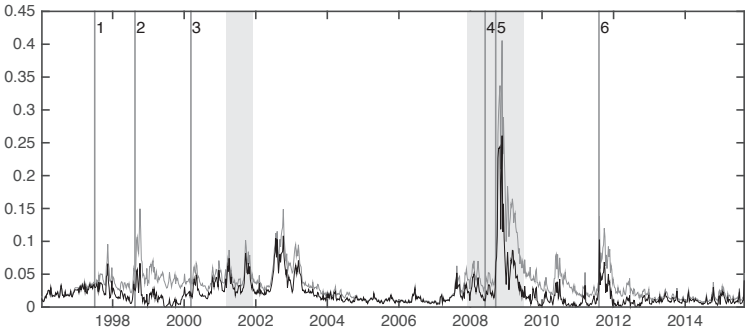
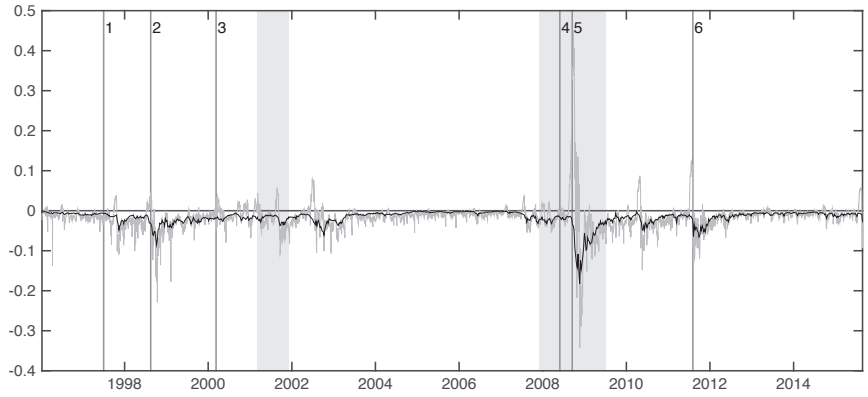
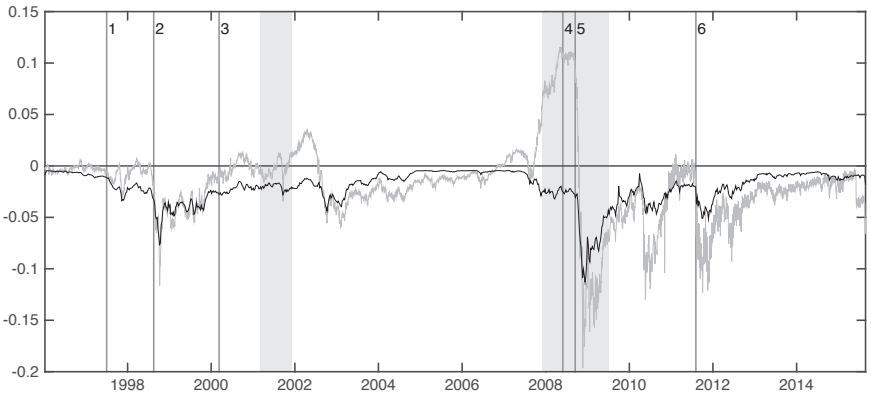


Figure 3: Filtered volatility factors as observable components of the volatility surface. Panel A: X_{11} (black line, left scale) and the 12-month at-the-money implied variance (grey line, right scale). Panel B: Negative value of X_{12} (black line, left scale) and option-implied skew $\mathcal{S}$ 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.*C.* of the Online Appendix for the definition of the option-implied 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

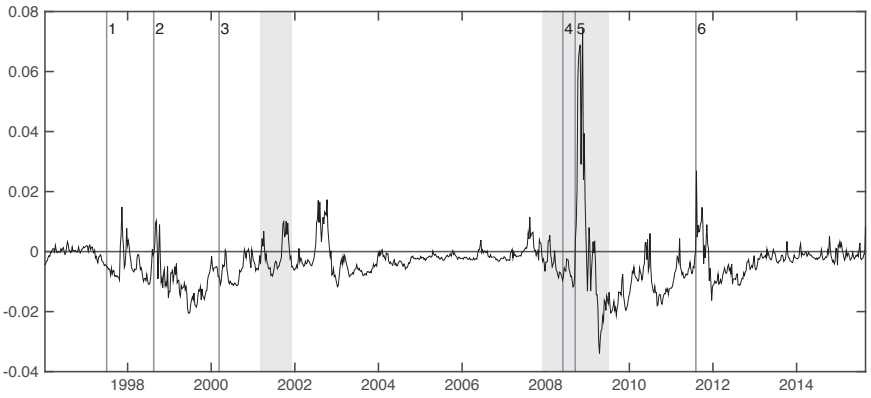
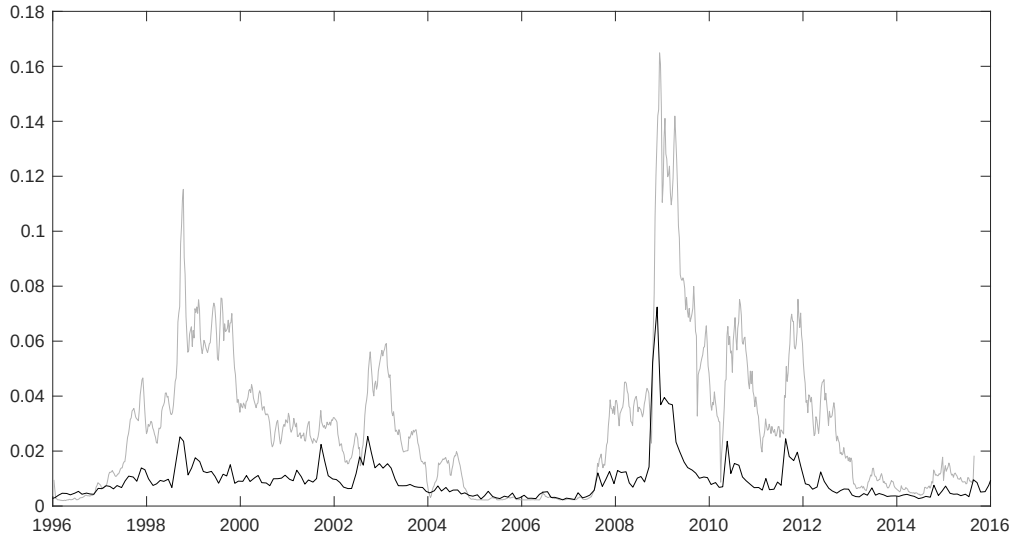


Figure 4: Variance risk premium and slope of the term structure of variance risk premia. In panel A (B), we plot the annualized model-implied 1 month (12 months) variance risk premium (black lines) and the payoffs of synthetic variance swaps (grey lines). In panel C, we plot the slope of the model-implied 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: Model GT^2



Panel B: Model AFT

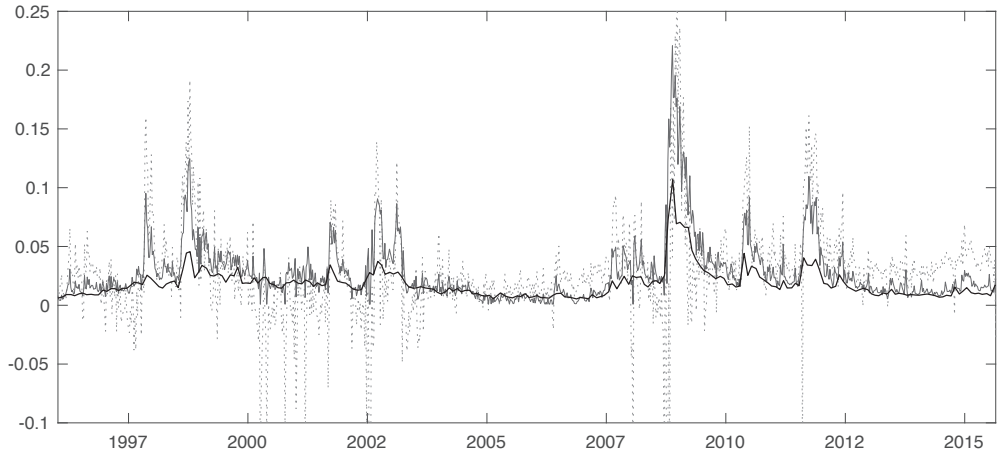


Figure 5: Time series of model-implied equity premia. Panel A: For our GT^2 model, we display the model-implied equity premium obtained using the third step estimation approach as described in Section I.*C.6.* (grey solid line). We contrast this quantity to the model-free lower bound of the equity premium in Schneider and Trojani (2019a) (black line). Panel B: For the AFT model, the estimated equity premium from a linear projection of future S&P 500 index excess returns on the hidden states obtained by nonlinear least squares (grey dotted line: all states, grey solid line: restricted regression using only state U_t). We contrast these quantities to the model-free lower bound of the equity premium in Schneider and Trojani (2019a) (black line).

Panel A: Diffusion parameters

M_{11}	-0.0079	M_{21}	1.0265	M_{22}	-2.6808
	(0.0000)		(0.0007)		(0.0023)
Q_{11}	0.0698	Q_{12}	-0.0770	Q_{22}	0.2924
	(0.0000)		(0.0000)		(0.0002)
R_{11}	-0.2970	R_{12}	-0.8708	R_{22}	-0.4057
	(0.0002)		(0.0006)		(0.0003)
M^*_{11}	-0.5467	M^*_{21}	0.3982	M^*_{22}	-2.6808
	(0.0004)		(0.0003)		(0.0021)
β	1.0012	β^*	1.0012		
	(0.0007)		(0.0007)		

Panel B: Jump parameters

Λ_{11}	25.6671	Λ_{12}	40.4278	Λ_{22}	15.9795
	(0.0200)		(0.0293)		(0.0114)
λ_0	0.0000	λ^-	7.1518	λ^+	58.3547
	(0.0000)		(0.0047)		(0.0413)
β^*_λ	0.3238				
	(0.0587)				

Panel C: Equity premium parameters

Δ_{11}	2.1879	Δ_{12}	0.1941	Δ_{22}	0.0172
δ	1.3×10^{-5}	β^*_Δ	0.9998		

Table 1: In-sample (1996/01-2002/12) parameter estimates and standard errors for our model GT^2. For the parameters of the benchmark models, see the Online Appendix, Tables 3 and 4.

	SV_{20}	SV_{30}	SV_{31}	SVJ_{20}	SVJ_{30}	GT^2
RMSIVE						
in-sample	1.322	1.225	0.897	0.841	0.703	0.685
out-of sample	1.586	1.395	1.049	1.019	0.736	0.705
MAIVE						
in-sample	1.009	0.935	0.700	0.666	0.538	0.537
out-of sample	1.246	1.093	0.835	0.798	0.563	0.548
Average log-likelihood						
in-sample	7.294	7.363	8.006	8.104	8.314	8.488
out-of sample	6.963	7.292	7.695	7.471	8.199	8.251
AIC						
in-sample	-5209	-5247	-5716	-5779	-5915	-6048
out-of sample	-9107	-9527	-10064	-9762	-10703	-10779

Table 2: Indicators of pricing performance and statistical fit. We report indicators of in- and out-of-sample pricing performance and fit for our model (GT^2) 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 August 2015. 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 report the in- and the out-of-sample average value of the weekly likelihood function, evaluated at the in-sample parameter estimates. Finally, we compute standard Akaike model-selection criteria (AIC). The in-sample period is 1996/01-2002/12 and the out of sample period is 2003/01-2015/08.

	Total Criterion			Option Pricing			Volatility Fit		
	IS	OS	All	IS	OS	All	IS	OS	All
GT^2	1.016	1.072	1.053	0.970	1.025	1.006	0.682	0.701	0.694
AFT	1.071	1.094	1.086	1.021	1.058	1.045	0.730	0.625	0.664

Table 3: Performance metrics under the nonlinear least squares estimation approach. For models AFT and GT^2, we report the optimized weighted nonlinear least squares criterion (35) for the in-sample (IS) period 1996-2002. We compute the values of the same criterion for the out-of-sample (OS) period 2003-2015/08 using estimated parameters and states from the in-sample period. For both periods, we also report the Option Pricing and the Volatility Fit components of the weighted objective function (35). All quantities are reported in percentage points as the square root of the average error in variance units.

References

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, Nicola Fusari, and Viktor Todorov, "Parametric Inference and Dynamic State Recovery from Option Panels," *Econometrica*, 2015, *83* (3), 1081–1145.

___ , ___ , and ___ , "The Risk Premia Embedded in Index Options," *Journal of Financial Economics*, 2015, *117* (3), 558–584.

___ , ___ , and ___ , "Short-Term Market Risks Implied by Weekly Options," *Journal of Finance*, 2017, *72* (3), 1335–1386.

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

Bardgett, C., E. Gourier, and M. Leippold, "Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets.," *Journal of Financial Economics*, 2019, *131* (3), 593–618.

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.

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

Carr, Peter and Dilip B. Madan, "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," *Journal of Financial and Quantitative Analysis*, 2017, *52* (5), 2119–2156.

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.

Christensen, Kim, Roel C.A. Oomen, and Mark Podolskij, "Fact or friction: Jumps at ultra high frequency," *Journal of Financial Economics*, 2014, *114*, 576–599.

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

Duffie, Darrell, Jun Pan, and Kenneth Singleton, "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 2000, *68 (6)*, 1343–1376.

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.

Gourio, Francois, "Disaster Risk and Business Cycle," *American Economic Review*, 2012, *102* (6), 2734–2766.

___ , "Financial Distress and Endogenous Uncertainty," *Society for Economic Dynamics Workig Paper*, 2014.

Gruber, Peter H., "Option Pricing with Matrix Affine Jump Diffusions," *Working paper, Università della Svizzera Italiana, Lugano*, 2015.

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

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.

Kozhan, Roman, Anthony Neuberger, and Paul G. Schneider, "The Skew Risk Premium in the Equity Index Market," *Review of Financial Studies*, 2013, *26*, 2174–2203.

Leippold, Markus and Fabio Trojani, "Asset Pricing with Matrix Jump Diffusions," *SSRN eLibrary paper no. 1274482*, 2008.

Lord, Roger and Christian Kahl, "Complex logarithms in Heston-like models," *Mathematical Finance*, 2010, *20* (4), 671–694.

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

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

Schneider, Paul and Fabio Trojani, "Fear Trading," *SSRN Working Paper 1994454*, 2014.

___ **and** ___ , "(Almost) Model-Free Recovery," *Journal of Finance*, 2019, *74* (1), 323–370.

___ **and** ___ , "Divergence And The Price Of Uncertainty," *Journal of Financial Econometrics*, 2019, *17* (3), 341396.