Resolution of Policy Uncertainty and Sudden Declines in Volatility

Dante Amengual
Centro de Estudios Monetarios y Financieros

Dacheng Xiu
University of Chicago Booth School of Business

This paper also can be downloaded without charge from the Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract=2348137
Resolution of Policy Uncertainty and Sudden Declines in Volatility

Dante Amengual† CEMFI
Dacheng Xiu‡ Chicago Booth

First Draft: May 7, 2013
This Version: Oct 28, 2013

Abstract

We introduce downward volatility jumps into a general framework of modeling the term structure of variance. With variance swap data alone, we find that downward volatility jumps are associated with a resolution of policy uncertainty, in particular through statements from Federal Open Market Committee meetings and speeches of Federal Reserve chairmen, and that such jumps are priced with positive risk premia, which reflect the premia for the “put protection” offered by the Federal Reserve. On the modeling side, we explore the structural differences and relative goodness-of-fits of factor specifications, and find that a log-volatility model with two Ornstein-Uhlenbeck factors and two-sided jumps is superior in capturing the volatility dynamics.

Keywords: Quadratic Volatility Models, Log Volatility Models, Downward Volatility Jumps, Variance Swaps

JEL Codes: G12, G13

†Centro de Estudios Monetarios y Financieros. Address: Casado del Alisal 5 Madrid, 28014 Madrid, Spain. E-mail address: amengual@cemfi.es.
‡Booth School of Business, University of Chicago. Address: 5807 S Woodlawn Avenue, Chicago IL 60637, USA. E-mail address: dacheng.xiu@chicagobooth.edu.
1 Introduction

Volatility responds to news. It rises dramatically and immediately following the occurrence of unexpected bad events.\(^1\) Moreover, volatility not only jumps upward but also moves downward rapidly. Sudden declines in volatility are sometimes related to stock market rallies stimulated by unexpected good news from economic indicators or earning announcements, yet are also very often triggered by the resolution of policy uncertainty that shift investors’ sentiment. Recent news headlines bring this indisputable fact into the spotlight. In particular, as can be seen from Figure 1, the VIX dropped 35% on May 10, 2010, as a result of Europe’s emergency loan plan, another 27% on Aug 9, 2011, due to Federal Reserve’s rate statement on keeping interest rates at a record low through mid-2013, and finally 23% on Dec 31, 2012, in anticipation of lawmakers making a deal to avert the “fiscal cliff.” \(^2\)

While the uncertainty of future fiscal and monetary policies may increase volatility of asset prices, the government and Federal Reserve often intervene in the midst of hard times, which effectively provides a put protection on asset prices.\(^3\) Our hypothesis is that many downward volatility jumps are market reactions to these policy measures, and that they are important sources of risk for volatility buyers. If the market price of risk for these sudden drops in volatility is positive, securities with larger positive sensitivities to such jumps should contain more jump risk premia. Many studies have estimated a negative price of risk for upward volatility jumps. Ignoring downward volatility jumps, if they are priced, may lead to an exaggeration of the negative jump risk premia and a biased-interpretation of the price of tail events. The goal of this paper is to provide a systematic investigation of where downward volatility jumps originate, how they affect asset prices, and whether they are priced risk factors.

These questions invite us to search for appropriate derivatives to investigate the asset pricing implications of volatility shocks. While the S&P 500 options offer a developed battlefield for volatility trading, volatility derivatives have thrived on the demand for hedging the variance risk since their inception. The over-the-counter index variance swap contract is one particular example of these popular derivatives. As with most swaps, the fixed leg of variance swaps pays a pre-determined amount at maturity in exchange for the realized variance the floating leg commits to offer. Despite the path-dependence of realized variance, the payoff structure of variance swaps is appealing for

\(^1\)For instance, the terrorist attack on September 11, 2001 sent the VIX near its historically high level prior to 2001.
\(^2\)Since we originally completed this paper, we have seen more events that reinforce this observation. For example, the VIX dropped 20.6% on Oct 10, 2013, following the news that House Republicans might agree to suspend the debt ceiling for six weeks.
\(^3\)We use “put protection” to refer to the monetary policy approaches that Alan Greenspan and Ben Bernanke, the former and current Chairmen of the Federal Reserve Board, exercised from 1987 to 2000 and during recent financial crisis.
studying the term structure of variance and variance risk premia, as opposed to the exchange-traded VIX derivatives, in that variance swaps directly reflect investors’ expectation on future volatility. Moreover, variance swaps are more sensitive to volatility jumps.

Despite their existence, whether and how these volatility jumps are priced by investors remains largely unknown, particularly in the case of the large downward ones. This is partially due to the absence of derivative pricing models that allow downward volatility jumps in the mainstream finance literature. Popular affine models such as the square-root volatility models can only incorporate upward jumps in order to ensure the positivity of variance. We incorporate downward volatility jumps and other potentially negative latent factors into a non-affine framework that guarantees the positivity of variance.

With this new and general non-affine framework, we price variance swaps in closed form, and identify downward volatility jumps along with two latent volatility factors from 17 years of variance swap data alone. We find that volatility jumps are often triggered by unexpected macro announcements. In particular, sudden declines in volatility are mostly associated with the resolution of policy uncertainty, such as monetary policy changes that are explicit or implicit from Federal Open Market Committee (FOMC) statements or the speeches of Federal Reserve’s chairmen, as well as fiscal policy decisions and compromises made by the Congress. Our analysis conforms with the existing findings that the total variance risk premia are negative. Yet, in addition we find that downward volatility jumps are priced with positive risk premia, evidence of the compensation for variance risk from the “Greenspan Put” or “Bernanke Put.” Our regression analysis shows that the latent volatility factors can be explained not only by excess market returns, but also by liquidity and credit factors, as well as policy news. In particular, policy news is important for the short-run factor, whereas the default risk is paramount for the long-term. More interestingly, we also find that most downward volatility jumps affect the short-term factor. Among several alternative specifications, we find that the log-volatility model with two Ornstein-Uhlenbeck factors and double exponential jumps yields the best in-sample fit and out-of-sample performance. Moreover, models that fail to incorporate such downward jumps are clearly misspecified, and lead to overestimates of total variance risk premia by 50% - 100%.

There is a growing amount of theoretical and empirical work relating political uncertainty to asset pricing. In particular, Pástor and Veronesi (2013) relate the stock market risk premia, volatility, and correlation to the policy uncertainty index constructed by Baker et al. (2013) which is based on the

4Since 2004 and 2006, the Chicago Board Options Exchange (CBOE) has introduced VIX futures and VIX options, respectively, offering investors additional instruments for volatility trading. These contracts are written on the VIX, a forward-looking implied volatility measure of the index, hence they are more complicated.

5While many macro announcements are pre-scheduled, their impact remains unexpected. As a result, the literature resorts to Poisson processes for modeling jumps, with a notable distinction by Piazzesi (2005).
frequency of newspaper references to economic policy uncertainty and other indicators. The regression results of Pástor and Veronesi (2013) agree with all the predictions of their learning model, see also Pástor and Veronesi (2012) for another related model of government policy choice. Bouchkova et al. (2012) investigate how local and global political risks affect industry return volatility. Kelly et al. (2012) find evidence for government guarantee premia by examining the basket-index spread from out-of-the-money put options. Bernanke and Kuttner (2005) study stock market reactions to Federal Reserve policy and find that the effects of unanticipated monetary policy actions on expected excess returns account for the largest part of the responses of stock prices. Recently, Lucca and Moench (2013) document large U.S. equity premia in anticipation of FOMC meetings, and discuss the challenges of explaining such premia with standard asset pricing theory. They also document an increase of intraday realized volatility shortly after the media releases that occur around 2:15 pm EST. Bekaert et al. (2013) decompose the VIX into a risk aversion component and an uncertainty component, using a structural vector autoregressive framework, and find that a lax monetary policy decreases both components. While these studies have shed light on the link between equity markets and political risk, we further point out that sudden decreases in volatility are particularly related to the resolution of monetary policy uncertainty, hence providing evidence of the variance risk premia associated with the Federal Reserve’s protection.

Our empirical findings on volatility jumps are also relevant to the large literature that investigates the unique role of jumps in asset pricing, which dates back as early as Merton (1976), who introduces jumps to model index returns. Since the seminal work by Duffie et al. (2000), positive volatility jumps, exponentially distributed, have been constantly added to model index volatility dynamics, so that volatility can jump upward but mean-revert back slowly. Eraker et al. (2003), in particular, point out the unique role played by such volatility jumps and compare them with the role of jumps in returns. However, there is rarely a model in the literature that discusses the existence and necessity of downward volatility jumps. An exception is Todorov and Tauchen (2011), who investigate the activity of volatility jumps using high-frequency returns of the VIX. In contrast, we focus on the asset pricing implications of volatility jumps and their price of risk. Recently, Chernov et al. (2012) discuss the impact of jumps on exchange rates and the impact of positive jumps on their variances, and relate these to macroeconomic and political news. They find few positive jumps in variance that respond to such news. Nonetheless, from variance swap data we find that downward volatility jumps are associated with resolutions reported in policy news.

Previous work in the literature on variance swaps is mostly based on fully specified parametric models using both variance swaps and index values with few exceptions. Bakshi and Kapadia (2003) estimate variance risk premia from delta-hedged gains. Carr and Wu (2009) study
Amengual (2008) find that single-factor volatility models are incapable of fitting the term structure of variance swap rates. They therefore suggest applying models with two-volatility factors to investigate the term structure of variance. None of their models have volatility jumps. Aït-Sahalia et al. (2012) propose a similar affine model without volatility jumps to estimate the variance risk premia and test the expectation hypothesis. They focus on the component of variance risk premia due to price jumps. Fusari and Gonzalez-Perez (2012) consider a log-affine model with two Ornstein-Uhlenbeck factors but without volatility jumps, in addition to an affine model. Carr et al. (2012) focus on the pricing and hedging of variance swaps and volatility derivatives in general using time-changed Lévy processes. Filipovic et al. (2013) independently propose a class of quadratic models without volatility jumps. All the aforementioned models are nested within our framework. Most importantly, while these two-factor volatility models without volatility jumps have been shown to yield accurate variance swap prices, we point out that their dynamics under the objective measure is critically misspecified, which in turn leads to an overestimate of total variance risk premia.

A notable distinction between our paper and the aforementioned ones lies in our econometric strategy, as we only use variance swap rates. The use of the S&P 500 index would help separate the volatility and price jump components from the total quadratic variation of the index returns, which is irrelevant for our purpose, since variance risk premia compensate the uncertainty of volatility shocks, i.e. the diffusive and jump components of the volatility process. Using the S&P 500 index requires strong assumptions about the price jumps, as in general they are only identifiable up to the conditional quadratic variation under the risk-neutral measure from variance swaps. In contrast, with model-free index dynamics, our estimation suffers from less potential misspecification, yet we can identify the term structure of variance, positive and downward volatility jumps, latent volatility factors, as well as a negative upper bound for total variance risk premia, which is informative enough to determine the sign of the premia. Moreover, provided additional assumptions on the structure of price jumps, along with the data on S&P 500 returns, our framework can also determine the exact amount of variance risk premia.

Our paper is also related to the model specification of index volatility dynamics, one of the central themes of empirical option pricing and financial econometrics. This strand of literature investigates the volatility dynamics through S&P 500 options, see e.g. Bakshi et al. (1997), Bates (2000), Pan (2002), Eraker (2004), and Broadie et al. (2007) for examples of affine jump diffusion models with variance risk premia using synthetic variance swaps for individual firms and index. In addition, Todorov (2010) focuses on the role of index jumps on variance risk premia using synthetic variance swaps in a semiparametric framework. Martin (2013) proposes a more robust relative of the variance swap and an alternative measure of market volatility nonparametrically.
stochastic volatility driven by one squared-root factor. Recent findings by Christoffersen et al. (2009) and Bates (2012) also suggest that models with two squared-root factors are essential in capturing the term structure of variance. Nevertheless, financial econometricians are in support of log-volatility models, potentially with jumps or even comprised purely of jumps, which fit the objective dynamics of volatility much better, see e.g. Barndorff-Nielsen and Shephard (2001), Chernov et al. (2003), and Todorov and Tauchen (2011). Indeed, log-volatility models naturally allow downward volatility jumps since they always guarantee the positivity of variance. Plus, log-volatility models are not restricted by a similar Feller’s condition for square-root processes, which is often binding empirically. Therefore, log-volatility models allow highly persistent volatility factors. The drawback of these log volatility models lies in their lack of tractability for derivative pricing. Nevertheless, we derive closed-form pricing formulae for variance swaps, and using these we can investigate the pricing implications of log-volatility models.

The paper is organized as follows. Section 2 presents our framework for variance swap modeling and pricing. Section 3 presents the canonical representation of admissible models. Section 4 discusses the statistical inference, followed by empirical results in Section 5. Section 6 concludes the paper. The Appendix provides all the technical details.

2 Variance Swap Valuation

A variance swap contract is an over-the-counter derivative in which the contract holder pays at maturity $t + \tau$ a fixed amount (variance swap rate) for the realized variance:

$$\frac{1}{\tau} \sum_{i=1}^{\lfloor \tau/\Delta \rfloor} \left( Y_{t+i\Delta} - Y_{t+(i-1)\Delta} \right)^2,$$

where $Y$ is the log-price of the underlying, i.e. S&P 500 index. Investors hedge variance risk by entering long positions of such contracts, earning negative variance risk premia. Variance swap trading has grown rapidly since the aftermath of LTCM turmoil in late 1990s. These over-the-counter contracts are more favorable than S&P 500 options for the purpose of volatility trading using medium- or low-frequency trading strategies, since market participants can express their views on volatility without having to do labor-intensive delta hedging.

As is well known, the realized variance converges (in probability) to the quadratic variation of $Y$, i.e. $[Y, Y]_{t, t+\tau}$, and modeling the quadratic variation is a common practice that facilitates the
variance swap pricing.\textsuperscript{7} Since there is no money changing hands at initiation of the trade, i.e. time $t$, the variance swap rate, under some risk neutral measure $Q$, is given by:\textsuperscript{8}

$$P(t, \tau) = 100 \times \frac{1}{\tau} E_t^Q \left\{ [Y, Y]_{t, t+\tau} \right\} = 100 \times \frac{1}{\tau} E_t^Q \left\{ \int_t^{t+\tau} \sigma_s^2 ds + \int_t^{t+\tau} \int_\mathbb{R} z^2 \nu_s^Q(dz) ds \right\}.$$ 

Apparently, the payoff of variance swaps depends on the underlying index $Y_t$ only through its risk-neutral quadratic variation, hence variance swaps contain much less information about the dynamics of $Y_t$, compared to European options. As a result, without making parametric assumptions, much less can be learned from swaps.\textsuperscript{9}

In contrast with the literature on pricing variance swaps, e.g. Amengual (2008), Egloff et al. (2010), Aït-Sahalia et al. (2012), and Carr et al. (2012), we model the spot variance (scaled by 100) as some function $f^Q$ of certain latent factor $X$:\textsuperscript{10}

$$100 \left( \sigma_s^2 + \int_\mathbb{R} z^2 \nu_s^Q(dz) \right) = f^Q(X_s).$$

This modeling strategy is similar in spirit to that of the term structure of interest rate models, where the short-rate $r_t$ is modeled as a linear function of latent factors. The difference, though, is that the spot variance not only contains the diffusive component which remains the same under changes of measures, but also includes the price jump component which depends on the risk-neutral measure. There are several reasons why modeling the whole variation together is advantageous. First, the index return dynamics remain model-free, allowing index jumps to have different intensity process from volatility jumps. This avoids spurious volatility jumps that may be falsely identified due to

\textsuperscript{7}The quadratic variation takes on the following form, see e.g. Protter (2004):

$$\frac{1}{\tau} [Y, Y]_{t, t+\tau} = \frac{1}{\tau} \left\{ \int_t^{t+\tau} \sigma_s^2 ds + \int_t^{t+\tau} \int_\mathbb{R} z^2 \mu(ds, dz) \right\},$$

where we model $Y$ as an Itô semimartingale. The explicit form of $Y$ can be written as

$$Y_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_\mathbb{R} z 1_{\{|z| \leq 1\}} (\mu(ds, dz) - \nu_s(dz) ds) + \int_0^t \int_\mathbb{R} z 1_{\{|z| > 1\}} \mu(ds, dz),$$

where $\mu(ds, dz)$ is a Poisson random measure with its compensator (stochastic intensity) $\nu_s(dz)$ being predictable and such that $\int_\mathbb{R} 1 \wedge z^2 \nu_s(dz)$ is locally integrable. This is the most general setup that excludes arbitrage, which has been discussed in e.g. Delbaen and Schachermayer (1994), Carr and Wu (2009), and Bollerslev and Todorov (2011).

\textsuperscript{8}Similar to the notation in the previous footnote, $\nu_t^Q(dz)$ denotes the compensator of price jumps under the risk neutral measure $Q$.

\textsuperscript{9}From cross-section of options, one can infer the prices of Arrow-Debreu securities, see e.g. Aït-Sahalia and Lo (1998) and Song and Xiu (2012).

\textsuperscript{10}The spot variance here refers to $d(Y_s, Y_t)^Q/ds$, where $(Y_s, Y_t)^Q$ is the conditional quadratic variation of $Y_s$, see, e.g. Protter (2004). For convenience, we use the term “spot variance” instead of “spot conditional quadratic variation”. We choose to model the spot conditional quadratic variation of returns instead of the instantaneous variance $\sigma_t^2$, as the former is naturally related to the pay-off structure of variance swaps. An alternative model is given by Appendix E.
index jumps. Moreover, the structure of price jumps under $Q$ is only identifiable up to the quadratic variation from variance swap rates. Hence, it is not possible to learn the risk-neutral distribution of price jumps in all cases, unless strong parametric assumptions are imposed to allow identification. Given the particular empirical questions we are after, there is no need to model each component of the quadratic variation separately.\(^\text{11}\) Finally, we can avoid the debate on the jump component of a variance swap payoff, as this becomes irrelevant here.\(^\text{12}\)

To ensure the tractability of this general class of models, we assume that the underlying $N$-dimensional factor $X_t$ follows a multivariate affine process, similar to the affine term structure models discussed in Dai and Singleton (2000), but allowing for jumps, e.g. in Duffie et al. (2000), Li and Song (2013) and Chen and Joslin (2012). Following their notations, the risk-neutral model of $X_t$ is written as:

$$
\frac{dX_t}{X_t} = (\Lambda^Q + K^Q X_t) dt + \Sigma \sqrt{S_t} dW^Q_t + dZ^Q_t,
$$

where $W^Q$ is an $N$-dimensional standard Brownian motion, and $S_t$ is a diagonal matrix in $\mathbb{R}^{N \times N}$ with $[S_t]_{i,i} = \alpha_i + \beta_i^T X_t$. $Z^Q_t$ is a pure jump martingale,\(^\text{13}\) which we can write down explicitly:

$$
Z^Q_t = \int_0^t \int_{\mathbb{R}^N} z(\mu^Q(ds, dz) - \nu^Q(X_s, dz) ds),
$$

where $\mu^Q$ is a Poisson random measure with compensator $\nu^Q$ that may depend on the latent factors in $X$. We assume that there exists $l_0^Q \in \mathbb{R}^+$ and $l_1^Q \in \mathbb{R}^N_+$, such that $\nu^Q(X_t, dz) = (l_0^Q + (l_1^Q)^T X_t) \nu^Q(dz)$. The specification of jumps nests almost all models in the literature, including standard Lévy models with $\nu^Q(X_s, dz) = \nu^Q(dz)$ being the Lévy measure, as well as jump diffusions, with time-varying jump intensity $l_0^Q + (l_1^Q)^T X_t$ and with $\nu^Q(dz)$ regulating jump sizes. Note that we define $Z_t$ as a compensated jump martingale, which simplifies the stationarity condition required in the canonical forms below.

We consider two types of models for $f^Q(X)$, which warrant a positive variance:

**Type I:** $f^Q(X) = \Pi_0^Q + \Pi_1^Q X + X^T \Pi_2^Q X$;  
**Type II:** $f^Q(X) = e^{\Pi_0^Q + \Pi_1^Q X}$.

\(^\text{11}\)That being said, learning about the exact variance risk premia requires such a joint model of the index and variance swaps, which we provide in Appendix E, whose variance dynamics is nested within our framework.

\(^\text{12}\)The variance swap contract pays realized variance, which is approximately equal to the quadratic variation. Nevertheless, practitioners may quote the price by building the replicating portfolio using S&P 500 options, following the same method as what the CBOE uses to construct the VIX. In that case, the payoff related to the jump part is $2 \int_{\mathbb{R}} (e^z - 1 - z \nu^Q(dz))$, which is slightly different from its quadratic variation $\int_\mathbb{R} z^2 \nu^Q(dz)$, but the difference is arguably small, see Jiang and Tian (2005). In the presence of large jumps, the discrepancy could be substantial, see e.g. Ait-Sahalia et al. (2012).

\(^\text{13}\)The martingale condition requires the mean measure to satisfy $\int_{\|z\| > 1} \|z\| \nu^Q(\cdot, dz) < \infty$, see e.g. Proposition 3.18 in Cont and Tankov (2004).
Type I models nest popular affine models and quadratic models in the literature, including those discussed in Bakshi et al. (1997), Bates (2000), Pan (2002), Eraker (2004), Eraker et al. (2003), Broadie et al. (2007), Bates (2012), and Carr et al. (2003) designed for pricing equity options, or Dai and Singleton (2000), Ahn et al. (2002), and Leippold and Wu (2002) for term structure of interest rate models. An overview of all these models is given by Singleton (2006). Unlike quadratic term structure of interest rate models, the underlying diffusion of the factor $X$ is not restricted to be constant, which allows richer dynamics for the term structure of variance.

Type II models are more common in the literature of financial econometrics. These models, as discussed in e.g. Nelson (1990), Chernov et al. (2003), Andersen et al. (2005), Todorov and Tauchen (2011), and Li and Xiu (2013), are known to fit volatility dynamics better, as they could achieve downward volatility jumps and slower mean-reverting rates or larger volatility of volatility without being restricted by the positivity constraint or Feller condition in a typical square-root process (Cox et al. (1985)) for volatility. Nevertheless, these models in general do not yield closed-form option pricing formulae, hence they are rarely used for $\mathcal{Q}$-measure dynamics.

Under this framework, the variance swap rate is given by:

$$P(t, \tau, X_t) = \left\{ \begin{array}{ll}
\frac{1}{\tau} \int_{t}^{t+\tau} \Pi_0^Q + \left(\Pi_1^Q\right)^T \nabla_u \Psi(s, t, u, X_t) \bigg|_{u=0} + \nabla_u^T \Pi_2^Q \nabla_u \Psi(s, t, u, X_t) \bigg|_{u=0} ds, & \text{for Type I models;} \\
\frac{1}{\tau} \int_{t}^{t+\tau} e^{\Pi_0^Q} \Psi(s, t, \Pi_1^Q, X_t) ds, & \text{for Type II models,}
\end{array} \right.$$  

where $\nabla_u = (\partial/\partial u_1, \ldots, \partial/\partial u_N)^T$ is a derivative operator. The derivation is based on the Fourier Transforms of tempered distributions. The details are given in Appendix A. To ensure the positivity of $f^Q$ for Type I models, we impose a constraint that $\Pi_0^Q - \frac{1}{2}(\Pi_1^Q)^T (\Pi_2^Q)^{-1} \Pi_1^Q \geq 0$ and that $\Pi_2^Q$ is a positive semi-definite matrix. For Type II models, no constraints are necessary.

These models are very flexible and general to model variance dynamics, so that we call them the Term Structure of Variance Models (TSVMs). Extending this model to allow a linear mixture payoff $f^Q$ of Type I and Type II is straightforward.

\section{Canonical Forms of TSVMs}

Our framework draws a great deal from the success of the affine term structure of interest rates models, see e.g. Dai and Singleton (2000), Piazzesi (2010), and Singleton (2006) for details. Indeed, variance is positive and non-tradable, like the short-rate. Moreover, variance swaps link to the spot variance in the same manner as certain fixed-income instruments are connected to the short-rate.
In light of this, we also model latent factors as affine diffusions with jumps. Nevertheless, we model volatility as a non-affine function of these latent factors, e.g. a quadratic or an exponential function, which warrants a positive variance. In contrast, most Gaussian term structure of interest rate models do not have a non-negative lower bound, in contradiction with the zero lower bound of interest rates. Ahn et al. (2002) and Leippold and Wu (2002) propose quadratic models, but their models require a constant diffusion matrix in order to obtain closed-form prices of bonds.

3.1 Canonical Forms

By analogy with the term structure of interest rate models, we propose canonical forms under a \( Q \)-measure that are identifiable up to invariant transformations. Such transformations refer to transformations and the rescaling of state and parameter vectors without changing the variance swap rates. We prefer canonical forms under \( Q \), since in general, \( P \)-measure dynamics are not restricted as long as there exists an equivalent change of measure. In addition to the standard canonical forms in the literature, e.g. Dai and Singleton (2000) and Ahn et al. (2002), we classify jumps, whose intensity process needs to be positive as well. These canonical models nest square-root factors, Ornstein-Uhlenbeck factors, pure-jump factors with state-dependent intensity, self-exciting jumps, Lévy jumps, etc.

For clarity, we do not allow pure jump factors below, but defer the discussion of those cases to Appendix C. Instead we focus on the canonical forms with all latent factors containing Brownian motions. As a result, we can recycle the notation \( \Lambda_m(N) \) by Dai and Singleton (2000), in which \( N \) is the number of state variables, while \( m \) is the number of independent linear combinations of those state variables that appear in the diffusion matrix, i.e., \( m = \text{rank}(B) \), where \( B = (\beta_1, \ldots, \beta_N) \). The state variables that appear in the diffusion matrix are non-negative. Moreover, the intensity can only load on positive factors, and jumps of the positive factors can only have positive sizes. The extended canonical forms in this case are easy to obtain.\(^{14}\)

\(^{14}\) For each \( m \), we partition \( X^T = (X_{m \times 1}^T, X_{(N-m) \times 1}^T)^T \). The extended canonical representation by Dai and Singleton (2000) takes a special form of equation (1), where for \( m > 0 \),

\[
K^Q = \begin{pmatrix}
K^Q_{m \times m} & 0_{m \times (N-m)} \\
K^Q_{(N-m) \times m} & K^Q_{(N-m) \times (N-m)}
\end{pmatrix},
\]

and \( K^Q \) is either upper or lower triangle for \( m = 0 \). In addition,

\[
\Lambda^Q = \begin{pmatrix}
\Lambda^Q_{m \times 1} \\
0_{(N-m) \times 1}
\end{pmatrix}, \quad \Sigma = I_N, \quad \alpha = \begin{pmatrix}
0_{m \times 1} \\
1_{(N-m) \times 1}
\end{pmatrix}, \\
B = \begin{pmatrix}
I_{m \times m} & B_{m \times (N-m)} \\
0_{(N-m) \times m} & 0_{(N-m) \times (N-m)}
\end{pmatrix}, \quad \ell_1^Q = \begin{pmatrix}
\ell_{1,m \times 1}^Q \\
0_{(N-m) \times 1}
\end{pmatrix},
\]
### 3.2 Risk Premia Specification

In general, risk premia can be chosen as completely affine, e.g. Dai and Singleton (2000), or essentially affine, e.g. Duffee (2002), or can be defined as the scaled difference between $\mathbb{P}$- and $\mathbb{Q}$-measures. As shown by Cheridito et al. (2007), the last procedure can also ensure lack of arbitrage as long as the existence and boundary non-attainment conditions are satisfied under both measures.

Next we give two examples of two-factor volatility models, which are used in the empirical section. Both models allow negative jumps in at least one of the factors. For convenience, we specify the risk premia with respect to the Brownian shocks and jumps to be affine functions of $\Sigma\sqrt{S}$ and $X$ respectively, so that the $\mathbb{P}$-dynamics remains affine.

### 3.3 Two-Factor Volatility Models

In the following, we provide two models with each allowing a two-factor structure and negative volatility jumps. Modeling volatility as a two-factor process is an established view from the literature. Engle and Rangel (2008) decompose volatility shocks into the short-run and long-run components, and relate the long-run part to business cycles in a comprehensive international setting. Adrian and Rosenberg (2008) also decompose equity volatility into similar components, and in addition relate the short-run component to market skewness risk with a cross-section of equity returns. Corradi et al. (2013) directly model the market volatility as a combination of business cycle factors and one additional latent factor, and find that their macro-factors explain the majority of volatility fluctuations. Christoffersen et al. (2009) also find two-factor volatility structure necessary to model S&P 500 options.

#### 3.3.1 Example: $A_{1}(2)$ Model

The $A_{1}(2)$ model specifies the dynamics of $X$ as:

\[
\begin{bmatrix}
    dX_{1t} \\ dX_{2t}
\end{bmatrix} = \begin{bmatrix}
    \lambda_1^Q \\ 0
\end{bmatrix} + \begin{bmatrix}
    \kappa_{11}^Q & 0 \\ \kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
    X_{1t} \\ X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
    \sqrt{X_{1t}} \\ \sqrt{1 + \beta_{21} X_{1t}}
\end{bmatrix} \begin{bmatrix}
    dW_{1t}^Q \\ dW_{2t}^Q
\end{bmatrix} + \begin{bmatrix}
    dZ_{1t}^Q \\ dZ_{2t}^Q
\end{bmatrix},
\]

where $X_1$ is a square-root factor, and $X_2$ is an Ornstein-Uhlenbeck factor. To guarantee admissibility and stationarity of the model, we require $\beta_{21} \geq 0$, $l_0 \geq 0$, $l_1 > 0$, $l_{12} = 0$, $\kappa_{11}^Q < 0$, and $\lambda_1^Q \geq 1/2$.

Moreover, $\bar{\nu}^Q(\mathbb{R}_m^m \times \mathbb{R}^{N-m}) = 0$. 

Jumps of $X_1$ and $X_2$ follow compound Poisson processes with independent jump sizes satisfying the exponential or double exponential distributions:

\[
\text{size of } Z_{1t}^Q \sim \exp(\beta_{1+}^Q), \quad \text{and} \quad \text{size of } Z_{2t}^Q \sim \begin{cases} 
\exp(\beta_{2+}^Q) & \text{with probability } q_2 \\
-\exp(\beta_{2-}^Q) & \text{with probability } 1-q_2.
\end{cases}
\]

Their intensity is specified as $l_0 + l_{11}X_{1t}$, hence the compensator is given by

\[
(l_0 + l_{11}X_{1t})\beta_{1+}^Q dt, \quad (l_0 + l_{11}X_{1t})(q_2\beta_{2+}^Q - (1-q_2)\beta_{2-}^Q) dt^T.
\]

For this model, we specify the dynamics under the objective measure $\mathbb{P}$ as

\[
\begin{bmatrix}
\frac{dX_{1t}}{dX_{2t}} \\
\end{bmatrix} = \begin{bmatrix}
\lambda_1^p \\
\lambda_2^p \\
\end{bmatrix} + \begin{bmatrix}
\kappa_{11}^p & 0 \\
\kappa_{21}^p & \kappa_{22}^p \\
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t} \\
\end{bmatrix} dt + \begin{bmatrix}
\sqrt{X_{1t}} \\
0 \\
\end{bmatrix} dt + \begin{bmatrix}
\sqrt{1 + \beta_{21}X_{1t}} \\
\sqrt{1 + \beta_{22}X_{1t}} \\
\end{bmatrix} \begin{bmatrix}
\frac{dW_{1t}^p}{dW_{2t}^p} \\
\frac{dZ_{1t}^p}{dZ_{2t}^p} \\
\end{bmatrix},
\]

which implies an affine specification of risk premia.\(^{15}\) This model is similar to the stochastic central tendency model discussed by Amengual (2008) and Mencia and Sentana (2012) for variance swaps and VIX derivatives, respectively.

### 3.3.2 Example: $A_0(2)$ Model

Another model that incorporates negative jumps into both factors can be specified as

\[
\begin{bmatrix}
\frac{dX_{1t}}{dX_{2t}} \\
\end{bmatrix} = \begin{bmatrix}
\kappa_{11}^Q & 0 \\
\kappa_{21}^Q & \kappa_{22}^Q \\
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t} \\
\end{bmatrix} dt + \begin{bmatrix}
\frac{dW_{1t}^Q}{dW_{2t}^Q} \\
\frac{dZ_{1t}^Q}{dZ_{2t}^Q} \\
\end{bmatrix},
\]

with $\kappa_{11}^Q < 0$, $\kappa_{22}^Q < 0$, $l_0 \geq 0$, and $l_{11} = l_{12} = 0$. As it turns out, both factors need to follow Ornstein-Uhlenbeck processes. Jumps follow compound Poisson processes with independent jump sizes following double exponential distributions:

\[
\text{size of } Z_{1t}^Q \sim \begin{cases} 
\exp(\beta_{1+}^Q), & q_1 \\
-\exp(\beta_{1-}^Q), & 1-q_1,
\end{cases}
\quad \text{and}
\]

\[
\text{size of } Z_{2t}^Q \sim \begin{cases} 
\exp(\beta_{2+}^Q), & q_2 \\
-\exp(\beta_{2-}^Q), & 1-q_2.
\end{cases}
\]

\(^{15}\)Writing $\eta_1\sqrt{X_1}$ and $\eta_2\sqrt{1 + \beta_{21}X_1}$ as the market price of risk corresponding to the Brownian shocks, we have

\[
\begin{align*}
\lambda_1^p &= \lambda_1^p + l_0(\beta_{1+}^p - \beta_{1+}^Q), \\
\lambda_2^p &= -\eta_2 + l_0(p_2\beta_{2+}^p - (1-p_2)\beta_{2+}^Q) - (q_2\beta_{2+}^Q - (1-q_2)\beta_{2-}^Q), \\
\kappa_{11}^p &= \kappa_{11}^Q - l_1(\beta_{1+}^Q - \beta_{1+}^p), \\
\kappa_{21}^p &= \kappa_{21}^Q - \eta_2\beta_{21}^Q + l_1((p_2\beta_{2+}^p - (1-p_2)\beta_{2+}^Q) - (q_2\beta_{2+}^Q - (1-q_2)\beta_{2-}^Q)),
\end{align*}
\]

where $\beta_{1+}^p$, $\beta_{2+}^p$, and $\beta_{2-}^p$ are jump sizes under objective measure $\mathbb{P}$, and $p_2$ is the mixture probability.
Their intensity is specified as \( l_0 \). For this model, we specify the dynamics under \( \mathbb{P} \) as

\[
\begin{bmatrix}
\frac{dX_{1t}}{dt} \\
\frac{dX_{2t}}{dt}
\end{bmatrix} = \left( \begin{bmatrix}
\lambda_1^P \\
\lambda_2^P
\end{bmatrix} + \begin{bmatrix}
\kappa_{11}^P & 0 \\
\kappa_{21}^P & \kappa_{22}^P
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} \right) dt + \begin{bmatrix}
\frac{dW_{1t}^P}{dt} \\
\frac{dW_{2t}^P}{dt}
\end{bmatrix} + \begin{bmatrix}
\frac{dZ_{1t}^P}{dt} \\
\frac{dZ_{2t}^P}{dt}
\end{bmatrix}.
\]

Similarly, the market price of risk is implicitly determined.\(^{16}\) Compared to the square-root factor, an Ornstein-Uhlenbeck factor is not restricted by the Feller condition, hence potentially allowing slower mean-reversion and larger volatility of volatility.

Other two factor volatility models including the \( \mathbb{A}_2(2) \) model, which is used extensively in the literature, is provided in Appendix C.

4 Likelihood Inference of TSVMs

This section contains details of our econometric inference, which readers may skip without any difficulty in understanding empirical results.

Our estimation strategy echoes with the estimation methods widely used in estimating dynamic term structure models for treasuries, where latent factors and parameters can be identified from bond prices alone. Recently, Ross (2012) discusses recovering risk premia from index options alone under certain assumptions using an equilibrium model. In the same spirit, our factors and parameters are identified from the cross-section of variance swaps through its likelihood using Markov Chain Monte Carlo (MCMC) methods, see e.g. Johannes and Polson (2010) for more details. We estimate four models in total including \( \mathbb{A}_0(2) \) and \( \mathbb{A}_1(2) \) of two different types, using MCMC that involves algorithms of Metropolis-Hasting within Gibbs.

We assume that there are observations available on \( n \) different variance swap rates over the period \([0, T]\) and that observations are recorded at a frequency \( \Delta \), so that the total number of time periods under consideration is \( n = T/\Delta \) with \( T \) being the length of the sample in years. For instance, \( \Delta = 1/252 \) denotes one trading day. Let \( Y \) denote the \( n \times p \) matrix of data.

For convenience, we introduce \( V \) and \( \Phi \) to summarize latent variables and parameters. Typically, \( V \) will contain the latent factors in \( X \) of the model as well as the remaining latent variables such as jump sizes and jump times, even though they do not enter into the pricing formulae. As for \( \Phi \), we split it into \((\Theta_M^T, \Theta_{H}^T, \Theta_F^T, \Theta_E^T)\). \( \Theta_M = (\Lambda^Q, vec^T(K^Q), \{\alpha_i, \beta_{ij}\}_{i=1}^m, \theta_2^Q)^T \), which contains the

\[\begin{align*}
\kappa_{11}^P &= \kappa_{11}^Q, \\
\kappa_{21}^P &= \kappa_{21}^Q, \\
\kappa_{22}^P &= \kappa_{22}^Q, \\
\lambda_1^P &= \eta_1 + l_0(\beta_{1+}^P - \beta_{1-}^Q), \\
\lambda_2^P &= \eta_2 + l_0 \left( (p_2 \beta_{2+}^P - (1-p_2)\beta_{2-}^Q) - (q_2 \beta_{2+}^Q - (1-q_2)\beta_{2-}^Q) \right).
\end{align*}\]

\(^{16}\) Writing \( \eta_1 \) and \( \eta_2 \) as the market price of risk corresponding to the Brownian shocks, we have
parameters determining the dynamics of the latent factors under the risk-neutral measure, with $\theta^Q_Z$ denoting the parameters governing the jump processes; $\Theta_M = (\Pi^Q_M, vec^T(\Pi^Q_M), vec^T(\Pi^Q_{\mathcal{E}}))^T$ includes the parameters defining $f^Q; \Theta_P = (\Lambda^P, vec^T(R^P))^T$ is implicitly defined once the market prices of risk, together with some elements of $\Theta_M$ are specified; and finally, given that there are more derivatives than sources of uncertainty in the theoretical model we allow pricing errors to avoid stochastic singularity, which are characterized by parameters summarized in $\Theta_E$. Specifically, we assume additive pricing errors $\varepsilon^j_{i\Delta}$ associated with derivative $j$ at time $i\Delta$ so that the observed price satisfies

$$Y^j_{i\Delta} = P(t, \tau_j, X_{i\Delta}; \Theta_M, \Theta_P) + \varepsilon^j_{i\Delta}$$

with $\varepsilon^j_{i\Delta} \sim N(0, s^2_j)$ and $\varepsilon^j_{i\Delta}$ is independent of $\varepsilon^h_{i\Delta}$ for $h \neq j$ and across time. $s^2_j$’s are stacked in $\Theta_E$. Notice that all prices are assumed to be observed with error in our framework and there is no need to assume that certain combinations of variance swap prices are perfectly observed.

We use a likelihood-based Bayesian approach for model estimation using the MCMC method. The purpose of MCMC sampling is to obtain a sample of parameters $\Phi = (\Theta_M^T, \Theta_{\Pi}^T, \Theta_{P}^T, \Theta_{E}^T)^T$ and latent variables $V$ from the their joint posterior density $p(V, \Phi|Y)$. Specifically, for a given model $M$, the posterior distribution summarizes the sample information regarding $\Phi$ and $V$,

$$p(V, \Phi|Y, M) \propto L(Y|V, \Phi, M) \cdot \mathcal{H}(V|\Phi, M) \cdot p(\Phi|M)$$

where $L(Y|V, \Phi, M)$ denotes the likelihood function, $\mathcal{H}(V|\Phi, M)$ is the density for the latent variables and $p(\Phi|M)$ is the prior density over the parameter vector $\Phi$.

We use a Gibbs sampling procedure to estimate these models. In essence, this amounts to reducing a complex problem, i.e. sampling from the joint posterior distribution, into a sequence of tractable ones, i.e. sampling from conditional distributions for a subset of the parameters conditional on all the other parameters, for which the literature already provides a solution. The Gibbs sampling procedure contains several blocks: two of them involving Metropolis-within-Gibbs steps. A brief description of them is next.

Inference is determined by sampling from the joint posterior distribution and latent states. The block algorithm in terms of posterior conditional distributions in the general model we consider is to sample sequentially from

| Latent factors | $p(X^{(g)}_t|X^{(g)}_{<t}, X^{(g-1)}_{>t}, V^{(g-1)}, \Phi^{(g-1)}, Y)$ |
| Q-measure parameters | $p(\theta^Q_M^{(g)}|X^{(g-1)}, V^{(g-1)}, \Phi^{(g-1)}, Y)$ |
| Pricing equation parameters | $p(\theta^P_{\Pi}^{(g)}|X^{(g-1)}, V^{(g-1)}, \Phi^{(g-1)}, Y)$ |
| P measure parameters | $p(\theta^P_{E}^{(g)}|X^{(g-1)}, V^{(g-1)}, \theta^P_{\mathcal{E}}^{(g-1)}, Y)$ |
Pricing error variances:  
\[ p(\theta_E^{(g)} | X^{(g-1)}, \Theta_M^{(g-1)}, \Theta_{II}^{(g-1)}, Y) \]

Jump processes:  
\[ p(V^{(g)} | X^{(g-1)}, V^{(g-1)}, \Theta_P^{(g-1)}, Y) \]

The supplemental Appendix D contains a detailed description of how we sample the relevant quantities for each of the sampling blocks. Below, we briefly summarize the conditional posterior distributions as well as the priors we use in our estimation strategy.

Table 1 presents the mean, standard deviation and 95% highest density region for these priors. The simulation results for Models \( A_0(2) \) and \( A_1(2) \) of two different types are provided in Tables 2 and 3. The parameters are calibrated from our empirical estimation.

## 5 Empirical Results

### 5.1 Data

We estimate four different specifications of models using daily data on variance swap rates on six different maturities (2, 3, 6, 9, 12 and 24 months) over the period January 4, 1996 to January 11, 2013. The number of daily observations is 4,278, excluding weekends and holidays. Due to restrictions from our data source, the sample is constructed as follows: it contains data on variance swaps quotes on 5 maturities (2, 3, 6, 12 and 24 months) from an anonymous U.S. bank over the period January 4, 1996 to March 30, 2007, whereas the the second dataset, which belongs to the same source covers the period starting from January 2, 2001 to January 11, 2013 with 4 maturities (3, 6, 9, and 12 months). Overall, we have an unbalanced panel of variance swaps over the past 17 years.

Figure 2 presents the variance swap rates for the different maturities over the whole sampling period. It is worth mentioning that, during the first half of the sample, they are characterized by a significantly higher market volatility which is due in part to the Asian, Russian and LTCM crises. After the “quiet” period that characterized the market between the years 2004 and 2007, we observe a sharp increase in volatility due to the 2007 - 08 financial meltdown, followed by two spikes related to the European sovereign debt crisis and the U.S. debt-ceiling confrontation.

Figure 3 highlights the changes in the slope of the variance term structure. For most of the sample, the variance is upward sloping, whereas in the middle and aftermath of crises, the term structure switches to a downward sloping shape, suggesting that volatility is expected to decrease towards its mean level. The fact that the term structure is not in perfect tandem with the variance level suggests the necessity to incorporate at least one additional factor that captures the term structure of variance.
We perform principal component analysis for the balanced panel with 1558 observations. The first three eigenvalues account for 97.80%, 99.69%, and 99.91% of total variations. The corresponding eigenvectors suggest that the first eigenvalue is related to level shifts in the variance curve while the second one captures changes in the slope of the curve. Nevertheless, the convexity effect seems negligible for variance swap data as the contribution of its corresponding principal component is tiny.

To understand what may drive the latent volatility factors, we conduct regression analysis using factors of economic fundamentals. We select two credit variables, including the daily TED spread, calculated as the difference between the three-month LIBOR and the three-month T-Bill interest rate and the default spread (DEF), calculated as the difference between the monthly Moody’s AAA and BAA corporate bond yield. We also obtain two monthly macro factors from the Federal Reserve’s website, the Chicago Fed National Activity Index (CFI), constructed from 85 monthly indicators of economic activity, and industrial production growth (IPG), as suggested by Pástor and Veronesi (2013) and Adrian and Rosenberg (2008). We also include the daily term spread (TERM), i.e. the difference between the yields on the 10-year and 3-month Treasury securities. We add one monthly liquidity factor (LIQ), the innovation of the aggregate liquidity from Pastor and Stambaugh (2003). To identify potential policy risk that may be related to volatility jumps, we add the policy news index (POL) constructed by Baker et al. (2013). Finally, we construct the market skewness factor as it is shown to be important for the short-run component by Adrian and Rosenberg (2008).

To explain volatility jumps, we construct measures of news surprises based on surveys of economists’ expectations on 18 economic indicators from Bloomberg. The detailed information about the categories, the announcement time, and the frequency of these news events are given in Table 4. We proxy news surprise as the differences between the actual release and the survey expectations:

\[
\text{News Surprise} = \frac{\text{Announced Quantity}_t}{\text{Median of Expectations}_t}.
\]

The news surprises of economic indicators are treated as the control variables since they are expected to produce jumps in S&P 500 returns other markets, e.g. Beechey and Wright (2009) and Faust and Wright (2009). To proxy the resolution of policy uncertainty, we use the schedules of FOMC and ECB meetings, as well as the speech schedules of the Federal Reserve’s Chairmen. Although the schedules are usually pre-announced and the target interest rates do not change often, the minutes, statements or press conferences after the FOMC meetings are informative about monetary policy decisions, and the information could be totally unpredictable.
5.2 Model Performance and Model Selection

Our principal component analysis suggests choosing two factors. The variation in the term structure also suggests at least two factors: one for the short-end of the curve and the other one for the long-end or the slope, intuitively. In light of evidence of negative jumps highlighted in Figure 1, we estimate two TSVMs: \( \Lambda_1(2) \) and \( \Lambda_0(2) \) of Types I and II, each of which allows for negative jumps through at least one Ornstein-Uhlenbeck factor.\(^{17}\)

In Tables 5 and 6, we report the posterior means and standard deviations of the parameter vectors \( \Theta_M, \Theta_H, \Theta_P, \) and \( \Theta_E \) for \( \Lambda_1(2) \) and \( \Lambda_0(2) \) models of Type II.\(^{18}\) Parameters are defined in annual terms following the convention in the empirical option pricing literature.

We first discuss the estimates of \( \Theta_M \) and \( \Theta_P \). As can be seen from the \( \kappa^{Q}_1 \)'s estimates, the first factor \( X_1 \) mean-reverts much faster than the second factor \( X_2 \) does. \( \kappa^{Q}_{11} \) is closer to the values found in the option pricing literature under the pricing measure. Also, the mean reversion parameter of \( X_2 \) under both measures is very low, around 0.2, implying that shocks to \( X_2 \) have a half life of several years. Moreover, for both models, positive jump sizes are larger under \( Q \) than under \( P \), while negative jump sizes are smaller in magnitude under \( Q \). This indicates that both types of jumps are priced, and that negative jumps have positive risk premia. The lower panel of Table 5 contains the corresponding summary statistics of the posterior distribution of the pricing equation parameters \( \Theta_H \). Not surprisingly, these parameters are estimated with high precision given that they are identified from prices.

We then analyze the properties of pricing error variances \( \Theta_E \), from which we can intuitively learn about the statistical performance of different models. Ideally, better models tend to produce smaller pricing errors. Figure 4 presents the median and interquartile range that characterize the pricing error variances. Again, we find a similar performance across four models.

Figures 5 and 6 provide the time series of all factors. More interestingly, the extracted factors share similar patterns across models and types. Although the levels of these factors are not the same due to differences in models, the similarity suggests that the extracted factors are very robust.

From the above analysis, it seems difficult to determine which model performs better, as their \( Q \)-measure fittings all look good in sample. Now we compare their out of sample performance of

\(^{17}\)Since the \( \Lambda_2(2) \) model does not allow any negative volatility jumps, we omit it for reasons of space. The empirical fitting is similar to that of Aït-Sahalia et al. (2012).

\(^{18}\)For reasons of space, we omit the parameter estimates of Type I models. Their performance is equally good for \( Q \)-measure fitting, as shown below, but Type II models are more parsimonious and easier to interpret. Type II models fit \( P \)-measure dynamics better, as it is clear from Figure 1 that the absolute changes of the squared VIX display much stronger heteroskedasticity compared to the percentage changes.
model-fitting using the VIX. The out-of-sample study here is on the cross-section, instead of time-
series, as is common for models with latent factors, see e.g. Piazzesi (2010). The VIX is constructed
by the CBOE using option portfolios, which often coincide with how variance swap contract writers
hedge their risk exposure. As a result, it is expected that the time-series of the squared VIX (scaled
by 100) and the model-predicted 1-month swap rates present very similar patterns, although as Ait-
Sahalia et al. (2012) point out, the difference between the squared VIX and variance swap rates is
related to the higher order moments of price jumps. The results are shown in Figure 7. Indeed,
the out-of-sample performance compared to the VIX is almost identical across these models, with
correlations as high as 0.89 for \( A_0(2) \) of Type II and \( A_1(2) \) of Type I, followed by \( A_1(2) \) of Type II
and \( A_0(2) \) of Type I, whose correlations are 0.88 and 0.85, respectively. We hereafter focus on Type
II models as they are more parsimonious and easy to interpret.

Having witnessed a tight competition among the \( Q \)-measure performance of these models, we then
move on to the time series of the estimated spot variance, which can be decomposed into jumps and
Brownian shocks. We decompose changes of estimated spot variances for \( A_1(2) \) and \( A_0(2) \) models in
Figures 8 respectively, which shed light on some new evidence on model selection among two-factor
volatility models.

The changes of the variance appear very similar to changes of the squared VIX in Figure 1, but
the decomposition is strikingly different. Perhaps not surprisingly, although the \( A_1(2) \) model is able
to capture negative jumps in one of its factors, the Ornstein-Uhlenbeck factor, this factor turns out
to be slow-mean reverting and highly persistent, which cannot accommodate jumps that perhaps
only affect short-run volatility levels. As a result, several significantly downward volatility changes
are attributed to Brownian shocks, as the square-root factor does not permit negative jumps. The
fact that the Ornstein-Uhlenbeck factor is selected by the data to be more persistent than the square-
root factor is not surprising, as the Feller condition prevents the mean-reversion speed from being
smaller than a certain threshold, which would be binding if the square-root factor was selected to be
the long-run one. In contrast, the \( A_0(2) \) model can accommodate jumps in both the short-run and
long-run factors, so that those short-run jumps missed by \( A_1(2) \) are captured.

Although the \( A_1(2) \) and \( A_2(2) \) (the \( A_{2,0}(2,0) \) model in Appendix C) models are more popular
than the \( A_0(2) \) model in the literature, we provide evidence that advocates the latter, as the former
two models cannot accommodate downward volatility jumps, especially those that only affect the
short-run volatility component. In addition, the Ornstein-Uhlenbeck factor seems to be a better
choice than the square-root factor when it comes to modeling persistent processes. We hence employ
the \( A_0(2) \) model in the following empirical study.
5.3 Where do Downward Volatility Jumps Come from?

There are several intriguing findings about volatility jumps from our best-fitting $A_0(2)$ model. First of all, we find that downward volatility jumps are as common as positive ones, and that they are often associated with resolutions of policy uncertainty. Apart from the three news headlines mentioned in the introduction, we highlight 23 of the largest downward jumps in Table 7, whose magnitudes exceed 20%. We also provide the corresponding changes of the S&P 500 index, most of which are, not surprisingly, positive, showing the so-called leverage effect. From this table, we find that the majority of large downward volatility jumps are caused by changes of current monetary policy or indications about future monetary policy, despite few jumps being relevant to fiscal policy, all of which may help comfort investors.

We now explain what the latent volatility factors are before figuring out how jumps are related to them. We conduct regression analysis trying to link the identified latent factors with known economic risk factors. We consider one-by-one simple regressions of each factor $X_i$, sampled at the end of each month from 1996 - 2012, on the innovation of each covariate given by Section 5.1, as well as the lagged value of $X_i$:

$$X_{i,t} = \beta_0 + \beta_1 Z_{j,t} + \beta_2 X_{i,t-1} + \varepsilon_t. \quad (4)$$

with $Z_{j,t}$ being the innovation of the $j$th covariate. For POL and TERM, we use ARIMA(1,1,0) innovations, as the Dicker Fuller tests fail to reject the unit-roots in our sample period. For IPG, we use the AR(3) innovation, following Adrian and Rosenberg (2008). For the rest of the covariates, we use AR(1) innovations. The results are identical when using other regression specifications.

We also consider a multiple time-series regression for all the innovations of the covariates plus the lagged $X_i$:

$$X_{i,t} = \beta_0 + \beta_1 \text{DEF}_t + \beta_2 \text{TED}_t + \beta_3 \text{TERM}_t + \beta_4 \text{LIQ}_t + \beta_5 \text{POL}_t +$$

$$+ \beta_6 \text{SKEW}_t + \beta_7 \text{ExM}_t + \beta_8 \text{IPG}_t + \beta_9 \text{CFI}_t + \beta_{10} X_{i,t-1} + \varepsilon_t. \quad (5)$$

Tables 8 and 9 provide regression results for $X_1$ and $X_2$, respectively. It is suggested by Table 8 that the time variation of short-run volatility factor $X_1$ can be explained by credit risk, liquidity risk, policy news, in addition to the excess returns. The signs of each coefficient agree with the intuition that short-run volatility rises if risk or uncertainty increases. When stacking these covariates into the multiple regression, policy news, excess market returns, and lagged values of $X_1$ subsume the rest of the covariates. As to the second volatility factor $X_2$, the default risk, term premia, and excess market return becomes significant with all covariates included. The AR(1) coefficient confirms that
$X_2$ is much more persistent than $X_1$. It is worth mentioning that our business cycle variables are not significant, potentially for two reasons. First, the sample period is as short as 17 years, which does not accommodate many business cycles. Secondly, the longest maturity of our variance swaps is 2 years, so that the “long” term factor extracted here may be regarded as the “median” term in macroeconomics, so that business cycle variables are less important.

To understand how volatility jumps interact with these factors, we regress the magnitudes of positive and negative volatility jumps onto the magnitude of macroeconomic news surprises for each volatility factor, respectively:

$$\text{Positive/Negative jump size of } X_{j,t} = \beta_0^{+/-} + \sum_{i=1}^{21} \beta_i^{+/-} |s_{i,t}| + \varepsilon_{j,t}^{+/-}, \text{ for } j = 1, 2. \quad (6)$$

where coefficients with $+/-$ correspond to regressions with positive and negative jumps respectively, and $s_{i,t}$ is the $i$-th news surprise at time $t$. If there is no such news event on day $t$, $s_{i,t}$ is set to 0. All news surprises are rescaled so that they all have the time-series standard error being equal to 1. The results are provided in Table 10. We find that negative volatility jumps in $X_1$ are mainly caused by FOMC meetings and Federal Reserve Chairmen’s speeches, whereas other volatility jumps are responsive to surprising news about employment, consumer spending, and national output. This conforms with our conjecture and earlier event studies that negative volatility jumps are highly correlated with the resolution of policy uncertainty. In addition, such jumps mostly affect the short-term volatility level.

### 5.4 Downward Volatility Jumps and the Federal Reserve Protection Premia

Finally, we investigate the pricing implication of downward volatility jumps. Comparing the estimates in Tables 5, positive volatility jumps have larger magnitudes under the $Q$-measure than under the $P$-measure, whereas negative jumps have smaller magnitudes under $Q$. This indicates that positive volatility jumps have negative price of risk, whereas negative jumps have positive price of risk. As a result, the total variance risk premia are overestimated, if downward volatility jumps are excluded.

To gauge the economic importance of downward jumps, we compare the variance risk premia implied from the $A_0(2)$ and $A_1(2)$ models, as the latter fails to include downward volatility jumps. We define variance risk premia in the same way as introduced in Carr and Wu (2009), Todorov (2010), and Bollerslev and Todorov (2011):

$$\text{VRP}(t, \tau) = \frac{1}{\tau} \left\{ \mathbb{E}_t^{P} \left( \int_t^{t+\tau} f^P(X_s) ds \right) - \mathbb{E}_t^{Q} \left( \int_t^{t+\tau} f^Q(X_s) ds \right) \right\}.$$
We assume that \( f^P = f^Q \),\(^{19}\) hence the exact variance risk premia can be calculated based on our estimates. We compare the time series of the variance risk premia in Figure 9 for both models.

Figure 9 suggests that variance risk premia are negative and countercyclical, i.e. they become even more negative during a crisis. For example, the lower troughs in the figure are associated with the 97-98 Asian crisis, the dot-com bubble, the recent financial meltdown, the European and U.S. debt crises, suggesting that investors require more compensation for bearing variance risk during difficult times. More importantly, the risk premia is 50-100% overestimated by the \( A_1(2) \) model, which cannot accommodate downward volatility jumps in the short-run.\(^{20}\) This further indicates that downward volatility jumps are compensated with positive risk premia, and that they contribute an important source of compensation for variance risk, most of which are due to the protection offered by the Federal Reserve.

### 6 Conclusion

Motivated by recent news headlines on the dramatic changes of the VIX following the announcements of policy makers, we conduct a systematic investigation of the sudden declines of market volatility. We find downward volatility jumps as common as positive ones, and that the majority of them are caused by FOMC announcements and the speeches of Federal Reserve Chairmen, showing the impact of Central Bank intervention, whereas a small portion of downward volatility jumps are responsive to surprising news about employment, consumer spending, and national output. This conforms with earlier event studies that negative volatility jumps are highly correlated with the resolution of policy uncertainty. Moreover, we find that such jumps mostly affect the short-term volatility level.

Our results indicate that positive volatility jumps have negative price of risk, whereas negative jumps have positive price of risk. In other words, the protection offered by the Federal Reserve is indeed priced by investors. As a result, ignoring downward volatility jumps leads to serious model misspecifications, which in turn leads to an exaggeration of the total variance risk premia by a large extent.

\(^{19}\)In general \( f^P \) is smaller than \( f^Q \) due to the difference in the jump component of variance risk premia. And jumps of S&P 500 dynamics are compensated with positive risk premia. Therefore,

\[
\int_t^{t+\tau} \int_R z \nu^P_t(dz) \leq \int_t^{t+\tau} \int_R z \nu^Q_t(dz) \iff \int_t^{t+\tau} f^P(X_s)ds \leq \int_t^{t+\tau} f^Q(X_s)ds.
\]

So, we have an upper bound for the variance risk premia in general:

\[
\text{VRP}(t, \tau) \leq \frac{1}{T} \left\{ \mathbb{E}_t^P \left( \int_t^{t+\tau} f^Q(X_s)ds \right) - \mathbb{E}_t^Q \left( \int_t^{t+\tau} f^Q(X_s)ds \right) \right\}.
\]

As can be seen from Figure 9, this upper bound is negative, which conforms with negative total variance risk premia found in the literature.

\(^{20}\)We find the same result from the comparison between \( A_0(2) \) and \( A_2(2) \).
extent.

In order to model downward volatility jumps, we introduce a new non-affine modeling framework which extends the classification and characterization of term structure models to allow jumps. Our canonical models nest square-root factors, Ornstein-Uhlenbeck factors, pure-jump factors with state-dependent intensity, self-exciting jumps, Levy jumps, etc. We find the log-type volatility model, which has been favored by financial econometricians in the past, with two Ornstein-Uhlenbeck factors and double exponential jumps yields the best performance in fitting the data.

Another distinction between our paper and the previous work lies in our econometric strategy, as we only use variance swap rates. With a model-free index dynamics, our estimation suffers from less potential misspecification, yet we can identify the term structure of variance, positive and downward volatility jumps, latent volatility factors, as well as a negative upper bound for total variance risk premia, which is informative enough to determine the sign of the premia.

Our modeling framework opens several avenues for empirical asset pricing. For instance, log-type volatility models may deserve further investigation regarding their asset pricing performance and the implications for exchange-traded derivatives, such as S&P 500 options or VIX derivatives. In addition, our framework can be used for modeling jumps in the treasury markets.
Appendix

A Variance Swap Pricing

Recall that since $X$ is affine, the generalized conditional characteristic function (GCCF) of $X_s$ is defined below, for any $s \geq t$ with $t$ fixed:

$$
\Psi(s, t, u, X_t) = \mathbb{E}_t^Q \left[e^{u^\top X_s}\right],
$$

where $u \in \mathbb{C}^N$. There exists a closed-form formula for the GCCF function given by Duffie et al. (2000):

$$
\log\left(\Psi(s, t, u, X_t)\right) = A(s - t, u) + B(s - t, u)^\top X_t,
$$

where $A$ and $B$ satisfy the following ordinary differential equations (ODEs):

$$
\dot{B} = (K^Q)^\top B + \frac{1}{2} \sum_{i=1}^m (\Sigma^Q B_i^2 \beta_i + i^Q \phi(B)),
$$

$$
\dot{A} = (\lambda^Q)^\top B + \frac{1}{2} \sum_{i=1}^m (\Sigma^Q B_i^2 \alpha_i + i^Q \phi(B)),
$$

where $B(t) = u$, $A(t) = 0$, and for any $h \in \mathbb{C}^N$,

$$
\phi(h) = \int_{\mathbb{R}^N} (e^{h^\top z} - 1 - h^\top z) \tilde{\nu}(dz).
$$

Denote the transition density of the process $X$ as $p(X_s|s - t, X_t)$, and let $u = -i\nu$ in $\Psi$ with $\nu \in \mathbb{R}^N$, we have

$$
\mathbb{E}_t^Q\left(f^Q(X_s)\big|X_t = x\right) = \int_{\mathbb{R}^N} f^Q(x')p(x'|s - t, x)dx'
$$

$$
= \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} f^Q(x')e^{i\nu^\top x'}\Psi(s, t, -i\nu, x)dx'd\nu.
$$

We then take advantage of the tempered distributions (see e.g. Kanwal (2004)) to simplify the integral with respect to $x'$. A similar idea has been used in Chen and Joslin (2012). Consider the Type I model first. Note that

$$
\int_{\mathbb{R}^N} (\Pi_0^Q + (\Pi_1^Q)^\top x' + x'^\top (\Pi_2^Q)x')e^{i\nu^\top x'}dx' = (2\pi)^N \left(\Pi_0^Q - i(\Pi_1^Q)^\top \nabla_v - \nabla_v(\Pi_2^Q)\nabla_v\right)\delta(v),
$$

where $\delta(\cdot)$ is Dirac delta that satisfies $\int_{\mathbb{R}^N} \delta(v)dv = 1$, and $\int_{\mathbb{R}^N} \delta(v)g(v)dv = g(0)$ for any test function $g$. Therefore, by direction calculations we obtain

$$
\mathbb{E}_t^Q\left(f^Q(X_s)\big|X_t = x\right) = \int_{\mathbb{R}^N} \left(\Pi_0^Q - i(\Pi_1^Q)^\top \nabla_v - \nabla_v(\Pi_2^Q)\nabla_v\right)\delta(v)\Psi(s, t, -i\nu, x)dv
$$
\[ = \Pi_0^Q + (\Pi_1^Q)^	op \nabla_u \Psi(s, t, u, x) \bigg|_{u=0} + \nabla_u \Pi_2^Q \nabla_u \Psi(s, t, u, x) \bigg|_{u=0}. \]

For the second type, we have similarly
\[
\int_{\mathbb{R}^N} e^{\Pi_0^Q + (\Pi_1^Q)^	op x'} e^{iv^	op x'} dx' = (2\pi)^N e^{\Pi_0^Q} \delta(v - i\Pi_1^Q),
\]
so that we can derive
\[
\mathbb{E}^Q \left( f^Q(X_s) \bigg| X_t = x \right) = \int_{\mathbb{R}^N} e^{\Pi_0^Q} \delta(v - i\Pi_1^Q) \Psi(s, t, -iv, x) dv = e^{\Pi_0^Q} \Psi(s, t, \Pi_1^Q, x).
\]

The pricing formulae for variance swaps follow immediately. Note that we have applied properties of tempered distributions to simply the calculations, all of which can be found in Kanwal (2004). The same technique has been applied by Amengual and Xiu (2012) for pricing VIX derivatives.

## B Invariant Transformations

We extend Dai and Singleton (2000) and Ahn et al. (2002) to provide invariant transformations of the general model for both types of specifications. These transformations may lead to alternative specification without altering the price of variance swaps. We summarize the state factors, Brownian motions, jumps, and parameter vectors in \( \theta \):

\[
\theta = \left( X_t, W_t^Q, Z_t^Q, \Lambda^Q, K^Q, \Sigma, \{\alpha_i, \beta_i\}_{1 \leq i \leq m}, \nu^Q(\cdot, dz), \Pi_0^Q, \Pi_1^Q, \Pi_2^Q \right).
\]

There are 4 classes of admissible transformations in total:

**An Affine Transformation** \( \mathcal{T}_A \) refers to \( \mathcal{T}_A X_t = \mathcal{V} + \mathcal{L} X_t \), where \( \mathcal{V} \) is an \( N \times 1 \) vector and \( \mathcal{L} \) is an \( N \times N \) nonsingular matrix. As a result, \( \mathcal{T}_A \theta \) is defined below. The transformation is the same for both types of specification, with \( \Pi_2 \) set at 0 in the case of the second type.

\[
\mathcal{T}_A \theta = \left( \begin{array}{c} \mathcal{V} + \mathcal{L} X_t, \mathcal{W}_t^Q, \mathcal{L} Z_t^Q, \mathcal{L} \Lambda - \mathcal{L} K \mathcal{L}^{-1} \mathcal{V}, \mathcal{L} K \mathcal{L}^{-1}, \mathcal{L} \Sigma, \\
\{\alpha_i - \beta_i \mathcal{L}^{-1} \beta_i, \mathcal{L}^{-1} \beta_i \}_{1 \leq i \leq m}, \nu^Q(\mathcal{L}^{-1}(\cdot + \mathcal{V}), \mathcal{L} dz), \\
\Pi_0^Q - (\Pi_1^Q)^\top \mathcal{L}^{-1} \mathcal{V} + \mathcal{V}^\top \mathcal{L}^{-1} (\Pi_2^Q) \mathcal{L}^{-1} \mathcal{V}, \mathcal{L}^{-1} \mathcal{V} - 2 \mathcal{L}^{-1} \Pi_2^Q \mathcal{L}^{-1} \mathcal{V}, \mathcal{L}^{-1} \Pi_2^Q \mathcal{L}^{-1} \mathcal{V} \end{array} \right).
\]

**An Orthonormal Rotation** \( \mathcal{T}_O \) refers to an affine transformation on the Brownian factor \( W_t^Q \) such that \( \mathcal{T}_O W_t^Q = O W_t^Q \), where \( O \) is an orthonormal matrix satisfying \( O^\top O = OO^\top = I_{N \times N} \).

\[
\mathcal{T}_O \theta = \left( X_t, O W_t^Q, Z_t^Q, \Lambda^Q, K^Q, \Sigma O^\top, \{\alpha_i, \beta_i\}_{1 \leq i \leq m}, \nu^Q(\cdot, dz), \Pi_0^Q, \Pi_1^Q, \Pi_2^Q \right).
\]

**A Diffusion Rescaling** \( \mathcal{T}_D \) rescales the diagonal elements of \( S_t \) by a nonsingular diagonal matrix \( D \) in \( \mathbb{R}^{N \times N} \). That is,

\[
\mathcal{T}_D \theta = \left( X_t, W_t^Q, Z_t^Q, \Lambda^Q, K^Q, \Sigma D^{-1}, \{D_i^2 \alpha_i, D_i^2 \beta_i\}_{1 \leq i \leq m}, \nu^Q(\cdot, dz), \Pi_0^Q, \Pi_1^Q, \Pi_2^Q \right).
\]

**A Permutation** \( \mathcal{T}_P \) alters the order of state variables, which has no observable effect.
C  Families of Admissible Affine Diffusions with Jumps

Here we discuss the canonical forms for general TSVMs. Each model of this class is assigned to a family $A_{m,j}(N, J)$, in which $N$ is the number of Brownian state variables, $J$ is the number of pure jump factors, while $m$ and $j$ are the number of independent linear combinations of those state variables that are positive, respectively. In the absence of pure jump factors, we can recycle the notation $A_m(N)$ in Dai and Singleton (2000).

C.1  General Canonical Forms

For each $m$, we partition $X^T = (X_{m\times 1}^T, X_{j\times 1}^T, X_{(N-m)\times 1}^T, X_{(J-j)\times 1}^T)^T$. The canonical representation takes a special form of equation (1), where for $m > 0$,

$$K^Q = \begin{pmatrix}
K_{m\times m}^Q & K_{m\times j}^Q & 0_{m\times (N-m)} & 0_{m\times (J-j)} \\
K_{j\times m}^Q & K_{j\times j}^Q & 0_{j\times (N-m)} & 0_{j\times (J-j)} \\
K_{(N-m)\times m}^Q & K_{(N-m)\times j}^Q & K_{(N-m)\times (N-m)}^Q & 0_{(N-m)\times (J-j)} \\
K_{(J-j)\times m}^Q & K_{(J-j)\times j}^Q & K_{(J-j)\times (N-m)}^Q & K_{(J-j)\times (J-j)}^Q 
\end{pmatrix},$$

and $K_{(N-m)\times (N-m)}^Q$ and $K_{(J-j)\times (J-j)}^Q$ is either the upper or lower triangle for $m = 0$ or $j = 0$, respectively. In addition,

$$\Lambda^Q = \begin{pmatrix}
\Lambda_{m\times 1}^Q \\
0_{j\times 1} \\
0_{(N+J-m-j)\times 1}
\end{pmatrix}, \quad \Sigma = \begin{pmatrix}1_{N\times N} \\ 0_{J\times J} \end{pmatrix}, \quad \alpha = \begin{pmatrix}0_{(m+j)\times 1} \\ 1_{(N-m)\times 1} \\ 0_{(J-j)\times 1}
\end{pmatrix},$$

$$B = \begin{pmatrix}
I_{m\times m} & 0_{m\times j} & B_{m\times (N-m)} & 0_{m\times (J-j)} \\
0_{j\times m} & 0_{j\times j} & B_{j\times (N-m)} & 0_{j\times (J-j)} \\
0_{(N-m)\times m} & 0_{(N-m)\times j} & 0_{(N-m)\times (N-m)} & 0_{(N-m)\times (J-j)} \\
0_{(J-j)\times m} & 0_{(J-j)\times j} & 0_{(J-j)\times (N-m)} & 0_{(J-j)\times (J-j)}
\end{pmatrix}, \quad \tau_1^Q = \begin{pmatrix}
\tau_{1,m\times 1}^Q \\
\tau_{1,j\times 1}^Q \\
0_{(N-m)\times 1} \\
0_{(J-j)\times 1}
\end{pmatrix},$$

with restrictions such that for $1 \leq i \neq k \leq m$, $(m+1) \leq s \neq t \leq m+j$ and $1 \leq j \leq N + J$,

$$K_{i,k}^Q \geq 0, \quad K_{s,t}^Q \geq 0, \quad \text{Re}(\text{Eigen}(K^Q)) < 0, \quad \Lambda_i^Q - \int_{\mathbb{R}} z_i \bar{\nu}^Q(z) dz \geq \frac{1}{2}, \quad \Lambda_{s}^Q \geq 0,$$

$$B_{ij} \geq 0, \quad B_{sj} \geq 0, \quad \tau_{1,i,j}^Q \geq 0, \quad \tau_{1,s}^Q = 1 \text{ or } 0, \quad \tau_0^Q \prod_{i=1,s}^{m+j} \tau_{1,s}^Q \neq 0.$$

Moreover, $\bar{\nu}^Q(\mathbb{R}_-^{m+j} \times \mathbb{R}^{N+J-m-j}) = 0$. 

25
C.2 Single Factor Models

The dynamics of the state variable may take one of four distinct forms. In the \( A_{0,0}(1, 0) \) model, we have

\[
dX_{1t} = \kappa_1^Q X_{1t} dt + dW_{1t}^Q + dZ_t^Q
\]

with \( \lambda_1^Q = 0 \). That is, jumps are of Lévy type, and potentially negative in size. Pricing VIX or variance swaps with \( f^Q \) of Type II under this model have been discussed in Todorov and Tauchen (2011) and Li and Xiu (2013).

With respect to the \( A_{1,0}(1, 0) \) model, we have

\[
dX_{1t} = (\lambda_1^Q + \kappa_1^Q X_{1t}) dt + \sqrt{X_{1t}} dW_{1t}^Q + dZ_t^Q,
\]

with \( \lambda_1^Q \geq 0 \). This is the most common model in the literature, employed for modeling the variance of the S&P 500 index, see e.g. Eraker et al. (2003), Eraker (2004), and Broadie et al. (2007). When jumps are absent, this model is equivalent to the standard \( A_1(1) \) model or Heston Model, see e.g. Heston (1993) and Pan (2002).

As for the \( A_{0,1}(0, 1) \) model, we have

\[
dX_{1t} = \kappa_1^Q X_{1t} dt + dZ_t^Q,
\]

with \( \lambda_1^Q = 1, \) and \( \kappa_1^Q < 0 \). Jumps are of a compound Poisson type with positive sizes and intensity being affine in \( X_{1t} \). With \( X_{1t} \) being interpreted as the intensity, this model is a self-exciting or Hawkes process, see e.g. Hawkes (1971), Aït-Sahalia et al. (2010), and Xiu (2012).

Finally, for the \( A_{0,0}(0, 1) \) model, we have

\[
dX_{1t} = \kappa_1^Q X_{1t} dt + dZ_t^Q,
\]

with \( \lambda_1^Q = 0 \). Jumps are of a Lévy type with potential negative sizes. This is a pure-jump type of model for volatility suggested by Todorov and Tauchen (2011).

C.3 Two-Factor Models

Below we introduce two-factor models. We have discussed \( A_{1,0}(2, 0) \) and \( A_{0,0}(2, 0) \) models in the main text. Here we focus on the remaining models.

The dynamics of the state variables in the \( A_{2,0}(2, 0) \) model are

\[
\begin{bmatrix}
  dX_{1t} \\
  dX_{2t}
\end{bmatrix} = \begin{bmatrix}
  \lambda_1^Q \\
  \lambda_2^Q
\end{bmatrix} + \begin{bmatrix}
  \kappa_{11}^Q & \kappa_{12}^Q \\
  \kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
  X_{1t} \\
  X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
  \sqrt{X_{1t}} & 0 \\
  0 & \sqrt{X_{2t}}
\end{bmatrix} \begin{bmatrix}
  dW_{1t}^Q \\
  dW_{2t}^Q
\end{bmatrix} + \begin{bmatrix}
  dZ_{1t}^Q \\
  dZ_{2t}^Q
\end{bmatrix},
\]
with \( t_0 > 0 \) and \( t_{12}^Q > 0 \). Jumps in \( Z_{1t} \) and \( Z_{2t} \) cannot be negative. This model has been applied to fit variance swap prices by Egloff et al. (2010) and Aït-Sahalia et al. (2012). We do not use this model, as volatility cannot jump downwards.

For pure jump models, the \( \mathcal{A}_{0,0}(1,1) \) model is

\[
\begin{bmatrix}
dX_{1t} \\
dX_{2t}
\end{bmatrix} = \begin{bmatrix}
\kappa_{11}^Q & \kappa_{12}^Q \\
\kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
\sqrt{1 + \beta_{12} X_{2t}^W} \, dW_{1t}^Q \\
0
\end{bmatrix} + \begin{bmatrix}
dZ_{1t}^Q \\
dZ_{2t}^Q
\end{bmatrix},
\]

with \( t_0 > 0 \), \( t_{11}^Q = 0 \), and \( t_{12}^Q = 0 \).

The \( \mathcal{A}_{0,1}(1,1) \) model is specified as

\[
\begin{bmatrix}
dX_{1t} \\
dX_{2t}
\end{bmatrix} = \begin{bmatrix}
\kappa_{11}^Q & \kappa_{12}^Q \\
\kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
\sqrt{1 + \beta_{12} X_{2t}^W} \, dW_{1t}^Q \\
0
\end{bmatrix} + \begin{bmatrix}
dZ_{1t}^Q \\
dZ_{2t}^Q
\end{bmatrix},
\]

with \( t_0 > 0 \), \( t_{11}^Q = 0 \), \( t_{12} = 1 \) or 0, and jumps in \( Z_{21}^Q \) are positive.

The \( \mathcal{A}_{1,0}(1,1) \) model is

\[
\begin{bmatrix}
dX_{1t} \\
dX_{2t}
\end{bmatrix} = \begin{bmatrix}
\lambda_1^Q & 0 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
\kappa_{11}^Q & \kappa_{12}^Q \\
\kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
\sqrt{X_{1t}} \, dW_{1t}^Q \\
0
\end{bmatrix} + \begin{bmatrix}
dZ_{1t}^Q \\
dZ_{2t}^Q
\end{bmatrix},
\]

with \( t_0 > 0 \), \( t_{11}^Q \geq 0 \), \( t_{12} = 0 \), and jumps in \( Z_{11}^Q \) are positive.

The \( \mathcal{A}_{1,1}(1,1) \) model is

\[
\begin{bmatrix}
dX_{1t} \\
dX_{2t}
\end{bmatrix} = \begin{bmatrix}
\lambda_1^Q & 0 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
\kappa_{11}^Q & \kappa_{12}^Q \\
\kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
\sqrt{X_{1t} + \beta_{12} X_{2t}^W} \, dW_{1t}^Q \\
0
\end{bmatrix} + \begin{bmatrix}
dZ_{1t}^Q \\
dZ_{2t}^Q
\end{bmatrix},
\]

with \( t_0 > 0 \), \( t_{11}^Q \geq 0 \), \( t_{12} = 0 \) or 1, and both jumps in \( Z_{11}^Q \) and \( Z_{21}^Q \) are positive.

Also, the \( \mathcal{A}_{0,0}(0,2) \) model is

\[
\begin{bmatrix}
dX_{1t} \\
dX_{2t}
\end{bmatrix} = \begin{bmatrix}
\kappa_{11}^Q & \kappa_{12}^Q \\
\kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
dZ_{1t}^Q \\
dZ_{2t}^Q
\end{bmatrix},
\]

with \( t_0^Q > 0 \), \( t_{11}^Q = 0 \), and \( t_{12}^Q = 0 \).

For the \( \mathcal{A}_{0,1}(0,2) \) model, we have

\[
\begin{bmatrix}
dX_{1t} \\
dX_{2t}
\end{bmatrix} = \begin{bmatrix}
\kappa_{11}^Q & 0 \\
\kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
dZ_{1t}^Q \\
dZ_{2t}^Q
\end{bmatrix},
\]

with \( t_0^Q \geq 0 \), \( t_{11}^Q = 0 \) or 1, \( t_0^Q t_{11}^Q \neq 0 \), and \( t_{12}^Q = 0 \). Jumps in \( Z_{11}^Q \) are positive.

For the \( \mathcal{A}_{0,2}(0,2) \) model, we have

\[
\begin{bmatrix}
dX_{1t} \\
dX_{2t}
\end{bmatrix} = \begin{bmatrix}
\kappa_{11}^Q & \kappa_{12}^Q \\
\kappa_{21}^Q & \kappa_{22}^Q
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t}
\end{bmatrix} dt + \begin{bmatrix}
dZ_{1t}^Q \\
dZ_{2t}^Q
\end{bmatrix},
\]

with \( t_0^Q \geq 0 \), \( t_{11}^Q = 0 \) or 1, \( t_{12}^Q = 0 \) or 1, and \( t_0^Q t_{11}^Q t_{12}^Q \neq 0 \). Jumps in both \( Z_{11}^Q \) and \( Z_{21}^Q \) are positive.
D Likelihood Inference in Detail

Below we give a more detailed description of some of the Gibbs blocks used in the posterior simulator.

D.1 Time Discretization and Additional Notation

For the purpose of concreteness in this section we focus on model $h_1(2)$. As mentioned above in section 3.3.1,

\[
\begin{bmatrix}
    dX_{1t} \\
    dX_{2t}
\end{bmatrix}
= 
\begin{bmatrix}
    \lambda_1^P & \lambda_2^P \\
    \lambda_1^F & \lambda_2^F
\end{bmatrix}
+ 
\begin{bmatrix}
    \kappa_{11}^P & 0 \\
    \kappa_{21}^P & \kappa_{22}^P
\end{bmatrix}
\begin{bmatrix}
    X_{1t} \\
    X_{2t}
\end{bmatrix}
\] 

\[
+ 
\begin{bmatrix}
    \sqrt{X_{1t}} & 0 \\
    0 & \sqrt{1 + \beta_{21} X_{1t}}
\end{bmatrix}
\begin{bmatrix}
    dW_{1t}^P \\
    dW_{2t}^P
\end{bmatrix}
+ 
\begin{bmatrix}
    dZ_{1t}^P \\
    dZ_{2t}^P
\end{bmatrix}.
\]

A time discretization of the model with time interval $\Delta$ yields

\[
X_{1i\Delta} - X_{1(i-1)\Delta} = \left[ \lambda_1^P + \kappa_{11}^P X_{1(i-1)\Delta} \right] \Delta + \sqrt{\Delta} \epsilon_{1i\Delta} + J_{1i\Delta}^P n_{i\Delta},
\]

\[
X_{2i\Delta} - X_{2(i-1)\Delta} = \left( \lambda_2^P + \kappa_{21}^P X_{1(i-1)\Delta} + \kappa_{22}^P X_{2(i-1)\Delta} \right) \Delta + \sqrt{1 + \beta_{21} X_{1(i-1)\Delta}} \Delta \epsilon_{2i\Delta} + J_{2i\Delta}^P n_{i\Delta},
\]

where $n_{i\Delta}$ denotes the jump time indicator that takes the value one if there is a jump on that day, using our definition of market prices of risk,

\[
\lambda_1^P = - \eta_1 + l_0 \left( \beta_{1+}^P - \beta_{1+}^Q \right),
\]

\[
\lambda_2^P = - \eta_2 + l_0 \left\{ p_2 \beta_{2+}^P - (1 - p_2) \beta_{2-}^P \right\} - \left( q_2 \beta_{2+}^Q - (1 - q_2) \beta_{2-}^Q \right),
\]

\[
\kappa_{11}^P = \kappa_{11}^P - \eta_1 + l_{11} \left( \beta_{1+}^P - \beta_{1+}^Q \right),
\]

\[
\kappa_{21}^P = \kappa_{21}^P - \eta_2 + l_{11} \left\{ p_2 \beta_{2+}^P - (1 - p_2) \beta_{2-}^P \right\} - \left( q_2 \beta_{2+}^Q - (1 - q_2) \beta_{2-}^Q \right),
\]

\[
\kappa_{22}^P = \kappa_{22}^Q,
\]

and where $\epsilon_{1i\Delta}$ and $\epsilon_{2i\Delta}$ are standard normal variates with zero correlation, $J_{1i\Delta}^P$ and $J_{2i\Delta}^P$ are Gamma and mixture of Gammas described above. We write $J_{i\Delta}^P = (J_{1i\Delta}^P, J_{2i\Delta}^P)$.

Conditional on jump times and sizes, we have that under the objective measure $P$ the dynamics of the jump-adjusted process satisfy

\[
\dot{X}_{1i\Delta} = b_{10} + b_{11} X_{1(i-1)\Delta} + \sqrt{X_{1(i-1)\Delta}} \Delta \epsilon_{1i\Delta},
\]

\[
\dot{X}_{2i\Delta} = b_{20} + b_{21} X_{1(i-1)\Delta} + b_{22} X_{2(i-1)\Delta} + \sqrt{1 + \beta_{21} X_{1(i-1)\Delta}} \Delta \epsilon_{2i\Delta},
\]

where

\[
\dot{X}_{1i\Delta} = X_{1i\Delta} - J_{1i\Delta}^P n_{i\Delta},
\]

\[
\dot{X}_{2i\Delta} = X_{2i\Delta} - J_{2i\Delta}^P n_{i\Delta},
\]

\[
b_{10}^P = \lambda_1^P \Delta, \quad b_{20}^P = \lambda_2^P \Delta,
\]

\[
\lambda_1^P, \lambda_2^P, \lambda_1^F, \lambda_2^F, \kappa_{11}^P, \kappa_{21}^P, \kappa_{22}^P, \kappa_{11}^Q, \kappa_{21}^Q, \kappa_{22}^Q,
\]

\[
b_{10}^P, b_{20}^P, b_{11}, b_{21}, b_{22}.
\]
\[ b_{11} = 1 - \kappa_{11}^P \Delta, \quad b_{21} = \kappa_{21}^P \Delta, \quad b_{22} = 1 - \kappa_{22}^P \Delta, \]

are the basis to compute \( p(X_{i\Delta} | X_{(i-1)\Delta}, V_{\Delta}; \Theta_M, \Theta_P) \).

### D.2 Latent Factors

The conditional posterior for the spot variance is not known in closed form. To sample from it, we use a Gaussian approximation to the transition density \( p(X_{i\Delta} | X_{(i-1)\Delta}, \Phi) \) corresponding to the Euler discretization of the process.\(^{21}\) At the \( g \)-th iteration of the sampler, we then draw from its conditional posterior using a random-walk Metropolis algorithm with normal proposal density with mean and variance computed as in Proposition 2 of Eraker (2001) but taking into account the presence of jumps. The acceptance rate of this step is in the 20-30% range for all models.

### D.3 \( \mathbb{Q} \)-Measure Parameters \( \Theta_M \) and \( \Theta_{\Pi} \)

Conditional on jump sizes, jump times, spot variance, short-run variance level, and \( (\Theta_{\Pi}, \Theta_P, \Theta_E) \), the posterior of \( \Theta_M \) is proportional to (2). Since this conditional distribution is nonstandard, it is sampled using a Metropolis step with a normal source density centered at the current draw and covariance matrix proportional to the Hessian of \( L(Y | V, \Phi, M) \cdot \mathcal{H}(V | \Phi, M) \) at the peak of \( \Theta_M \). The Hessian was computed by concentrating the latent variables and remaining parameters on their posterior means from a preliminary run of the algorithm. An analogous, but simpler procedure since \( \mathcal{H}(V | \Phi, M) \) does not appear in the conditional posterior allows us to draw \( \Theta_{\Pi} \). The acceptance rate of this step is around 30% for all four models. The priors are relatively uninformative but still impose the relevant constraints.

### D.4 \( \mathbb{P} \)-Measure Parameters \( \Theta_P \) and Pricing Error Variances \( \Theta_E \)

A similar procedure to the one mentioned above can be used to sample \( \Theta_P \). In practice, however, since those parameters do not depend on variance swap rates once we condition on \( V \), it is often the case that the conditional posterior distribution is available and therefore one can sample from it directly. The same comment applies to the variances of pricing errors as long as one chooses appropriately both the pricing error distributions and priors.

As for \( \beta_{1+}^P \), recall \( J_{\Pi\Delta}^P \sim \text{Exponential}(\beta_{1+}^P) \), so that conditional on \( J_{\Pi \Delta}^P \), and setting a conjugate prior for \( \beta_{1+}^P \), say \( \pi_{\beta_{1+}^P}(\beta_{1+}) \sim \text{InvGam}(\delta_{\beta_{1+}^1}, \delta_{\beta_{1+}^2}) \), where \( \mathbb{P} \) is omitted for convenience, we have

\(^{21}\)Eraker et al. (2003) show that discretization performs well with daily data. Alternatively, one could introduce a set of \( h - 1 \) auxiliary data points in between of each pair of sampled latent variables and integrate them out of the likelihood function by MCMC.
that
\[ p(\beta_1+|\ldots) \propto p(j|\ldots, \beta_{1+}, X, n) \cdot \pi_{\beta_{1+}}(\beta_{1+}) \propto \beta_{1+}^{-N_J} \exp\left( -\frac{1}{\beta_{1+}} \sum_{i=1}^{N_J} J_{i1\Delta}^{p} \right) \cdot \beta_{1+}^{-(\delta_{\beta_{1+}+1})} \exp\left( -\delta_{\beta_{1+}+2}/\beta_{1+} \right), \]
where \( N_J = \sum_{i=1}^{N} n_i \Delta; \) that is \( \beta_{1+}^{p} \ldots \sim \text{invGam}(\delta_{\beta_{1+}+1}^{*}, \delta_{\beta_{1+}+2}^{*}) \) with \( \delta_{\beta_{1+}+1}^{*} = N_J + \delta_{\beta_{1+}+1} \) and \( \delta_{\beta_{1+}+2}^{*} = \delta_{\beta_{1+}+2} + \sum_{i=1}^{N_J} J_{i1\Delta}^{p}. \)

Similarly, we proceed with \( \beta_{2+}^{p} \) and \( \beta_{2-}^{p} \), but using the appropriate sample sizes \( N_{J_+} = \sum_{i=1}^{N} n_i \Delta 1\{j_{i2\Delta}^{p} > 0\} \) and \( N_{J_-} = N_J - N_{J_+} \) and with \( \pi_{\beta_{2+}}(\beta_{2+}) \sim \text{invGam}(\delta_{\beta_{2+}+1}^{*}, \delta_{\beta_{2+}+2}^{*}) \) and \( \pi_{\beta_{2-}}(\beta_{2-}) \sim \text{invGam}(\delta_{\beta_{2-}+1}^{*}, \delta_{\beta_{2-}+2}^{*}) \) being the corresponding priors.

As for \( p \), assuming a Beta prior \( \text{Beta}(\delta_{p_1}, \delta_{p_2}) \), the conditional posterior of \( p \) is simply \( \text{Beta}(\delta_{p_1}^{*}, \delta_{p_2}^{*}) \) with \( \delta_{p_1}^{*} = \delta_{p_1} + N_{J_+}^{*} \) and \( \delta_{p_2}^{*} = \delta_{p_2} + N_{J_-}^{*}. \)

Conditional on \( X \), \( \beta_{1+}^{p}, \beta_{2+}^{p}, \beta_{2-}^{p}, p \), \( \Theta_M \) and jump times and sizes, using the jump adjusted processes \( \hat{X}_{1,i\Delta} = X_{1,i\Delta} - J_{1,i\Delta}^{p} n_i \Delta \) and \( \hat{X}_{2,i\Delta} = X_{2,i\Delta} - J_{2,i\Delta}^{p} n_i \Delta \), we can sample \( \kappa_{11}^{p}, \kappa_{22}^{p} \) and \( \kappa_{21}^{p} \) using results from the standard normal linear model with known variance and without intercept. Specifically, rewrite the discretized SDEs as
\[
\begin{bmatrix}
  x_{1i} \\
  x_{2i}
\end{bmatrix}
= \begin{bmatrix}
  \kappa_{11}^{p} & 0 \\
  \kappa_{21}^{p} & \kappa_{22}^{p}
\end{bmatrix}
\begin{bmatrix}
  x_{1(i-1)\Delta} \\
  x_{2(i-1)\Delta}
\end{bmatrix}
+ \begin{bmatrix}
  \epsilon_{1i\Delta} \\
  \epsilon_{2i\Delta}
\end{bmatrix},
\]
with
\[
x_{1i} = \frac{\hat{X}_{1,i\Delta} - X_{1(i-1)\Delta} - \lambda_{11}^{p} \Delta}{\sqrt{\beta_{11} X_{1(i-1)\Delta} \Delta}}, \quad x_{2i} = \frac{\hat{X}_{2,i\Delta} - X_{2(i-1)\Delta} - \lambda_{22}^{p} \Delta}{\sqrt{\alpha_{2} + \beta_{21} X_{1(i-1)\Delta} \Delta}},
\]
where we have used
\[
\begin{align*}
\hat{X}_{1,i\Delta} &= \lambda_{11}^{p} \Delta + b_{11} X_{1(i-1)\Delta} + \sqrt{\beta_{11} X_{1(i-1)\Delta} \Delta} \epsilon_{1i\Delta}, \\
\hat{X}_{2,i\Delta} &= \lambda_{22}^{p} \Delta + b_{21} X_{1(i-1)\Delta} + b_{22} X_{2(i-1)\Delta} + \sqrt{\alpha_{2} + \beta_{21} X_{1(i-1)\Delta} \Delta} \epsilon_{2i\Delta},
\end{align*}
\]
with
\[
b_{11} = 1 - \kappa_{11}^{p} \Delta, \quad b_{21} = \kappa_{21}^{p} \Delta \quad \text{and} \quad b_{22} = 1 - \kappa_{22}^{p} \Delta.
\]

In this context, prior information can be easily introduced through dummy observation priors (Gaussian conjugate priors) and the standard normal linear model can be used to sample from.

### D.5 Jump Sizes and Times

Sampling these components depends heavily on the model in hand. In the context of finite activity jump processes, there exists an extensive methodological and applied literature that discusses and applies MCMC methods for Poisson jump processes (see Johannes and Polson (2010) for a recent
survey). More recently, closing the gap between pure Brownian and pure Poisson processes, Li et al. (2008) develop MCMC methods for continuous-time models with stochastic volatility and infinite activity Lévy processes such as the Variance-Gamma model of Madan et al. (1998) or the log-stable process of Carr and Wu (2003).

Specifically, in our application the jump indicator, \( n_i \Delta \), is a binary random variable (taking on 0 or 1). To compute the Bernoulli probability, we use the conditional density of increments to volatility and returns to get that

\[
p(n_i \Delta = 1|y, V, \Phi, Y) \propto p(X_i \Delta | X_{(i-1)\Delta}, V_i \Delta, \Phi) \cdot p(n_i \Delta = 1|X_{(i-1)\Delta}).
\]

Thus, it is sufficient to find the probability \( \Pr(n_i \Delta = 1|V, \Phi, Y) \), which again simplifies to

\[
\Pr(n_i \Delta = 1|V, \Phi, Y) = \sum_{s=0}^{1} \frac{p(X_i \Delta | X_{(i-1)\Delta}, n_i \Delta = 1, J_{i\Delta}^p, \Phi) \cdot p(n_i \Delta = 1|X_{(i-1)\Delta})}{p(X_i \Delta | X_{(i-1)\Delta}, n_i \Delta = s, J_{i\Delta}^p, \Phi) \cdot p(n_i \Delta = s|X_{(i-1)\Delta})},
\]

where \( p(X_i \Delta | X_{(i-1)\Delta}, n_i \Delta = s, J_{i\Delta}^p, \Phi) \) is a bivariate normal distribution with mean and covariance matrix

\[
\begin{pmatrix}
\lambda_1^p \Delta + b_{11} X_{1(i-1)\Delta} + s \cdot J_{1i\Delta}^p \\
\lambda_2^p \Delta + b_{21} X_{1(i-1)\Delta} + b_{22} X_{2(i-1)\Delta} + s \cdot J_{2i\Delta}^p
\end{pmatrix}
\]

and

\[
\begin{pmatrix}
X_{1(i-1)\Delta} & 0 \\
0 & 1 + \beta_{21} X_{1(i-1)\Delta}
\end{pmatrix},
\]

and \( p(n_i \Delta = 1|X_{(i-1)\Delta}) = (l_0 + l_{11} X_{1(i-1)\Delta}) \Delta \). Not surprisingly, the conditional posterior of jump times does not depend on the option prices directly since option prices do themselves not depend on the jump indicator. The same is true for the conditional distribution of jump sizes \( J_{1\Delta}^p \) and \( J_{2\Delta}^p \) as prices depend only on the states \( X_t \). Specifically, the conditional posterior for the jump sizes to stock return and volatility, \( p(J_{i\Delta}^p | n_i \Delta = 1, \Phi, Y) \), which, by Bayesian rule is proportional to

\[
p(X_i \Delta | X_{(i-1)\Delta}, n_i \Delta = s, J_{i\Delta}^p, \Phi) \cdot p(J_{1i\Delta}^p | n_i \Delta = s, \Phi) \cdot p(J_{2i\Delta}^p | n_i \Delta = s, \Phi),
\]

where the first term is the bivariate normal distribution specified above with \( s = 1 \) and the second and third terms are the discrete scale mixture of two Gammas and Gamma random variates from the initial specification. To find the conditional posterior, it is convenient to do it separately, say

\[
p(X_{1i\Delta} | X_{1(i-1)\Delta}, n_i \Delta = 1, J_{1i\Delta}^p, \Phi) \cdot p(J_{1i\Delta}^p | n_i \Delta = 1, \Phi),
\]

and

\[
p(X_{2i\Delta} | X_{1(i-1)\Delta}, X_{2(i-1)\Delta}, n_i \Delta = 1, J_{2i\Delta}^p, \Phi) \cdot p(J_{2i\Delta}^p | n_i \Delta = 1, \Phi).
\]

Then, completing the square of the exponent of

\[
\frac{1}{\sqrt{2\pi} \beta_{11} X_{1(i-1)\Delta}} \exp \left( -\frac{(X_{1i\Delta} - \lambda_1^p \Delta - b_{11} X_{1(i-1)\Delta} - J_{1i\Delta}^p)^2}{2 X_{1(i-1)\Delta} \beta_{11}^2} \right) \cdot \frac{1}{\beta_{1+}^2} \exp \left( -\frac{J_{1i\Delta}^p}{\beta_{1+}} \right)
\]
all three terms as a function of \( J_{1\Delta}^p \)

\[
\frac{\beta_{1+}^p (X_{1\Delta} - \lambda_{1}^p \Delta - b_{11} X_{1(i-1)\Delta}) - X_{1(i-1)\Delta} \beta_{11}^p}{\beta_{1+}^p X_{1(i-1)\Delta} \beta_{11}^p} J_{1\Delta}^p
\]

\[
- \frac{(J_{1\Delta}^p)^2 + (X_{1\Delta} - \lambda_{1}^p \Delta - b_{11} X_{1(i-1)\Delta})^2}{2X_{1(i-1)\Delta} \beta_{11}^p}
\]

leads to a truncated normal \( TN(\mu_{1+}^*, \sigma_{1+}^2; J_{1\Delta}^p > 0) \) with

\[
\mu_{1+}^* = \frac{\beta_{1+}^p (X_{1\Delta} - \lambda_{1}^p \Delta - b_{11} X_{1(i-1)\Delta}) - X_{1(i-1)\Delta} \beta_{11}^p}{\beta_{1+}^p}
\]

and variance \( \sigma_{1+}^2 = X_{1(i-1)\Delta} \beta_{11}^p \).

Similarly, completing the square of \( y_{i\Delta} \mid V_{i\Delta}, V_{(i-1)\Delta} \),

\[
\frac{1}{\sqrt{2\pi (1 + \beta_{21} X_{1(i-1)\Delta}) \Delta}} \exp \left\{ - \frac{[X_{2i\Delta} - \lambda_{2}^p \Delta - b_{21} X_{1(i-1)\Delta} - b_{22} X_{2(i-1)\Delta} - J_{2i\Delta}^p]^2}{2(1 + \beta_{21} X_{1(i-1)\Delta}) \Delta} \right\}
\]

\[
\cdot \left[ \frac{p}{\beta_{2+}^p} \exp \left( - \frac{J_{2i\Delta}^p}{\beta_{2+}^p} \right) + \frac{1 - p}{\beta_{2-}^p} \exp \left( - \frac{J_{2i\Delta}^p}{\beta_{2-}^p} \right) \right]
\]

yields a discrete scale mixture of truncated normals with mixing variate that takes a positive value with mean

\[
\mu_{2+}^* = \frac{\beta_{2+}^p [X_{2i\Delta} - \lambda_{2}^p \Delta - b_{21} X_{1(i-1)\Delta} - b_{22} X_{2(i-1)\Delta}] - (1 + \beta_{21} X_{1(i-1)\Delta}) \Delta}{\beta_{2+}^p}
\]

and variance \( \sigma_{2+}^2 = (\alpha_{2} + \beta_{21} X_{1(i-1)\Delta}) \Delta \) with probability \( p \), and with probability \( 1 - p \), a negative value with mean

\[
\mu_{2-}^* = \frac{\beta_{2-}^p [X_{2i\Delta} - \lambda_{2}^p \Delta - b_{21} X_{1(i-1)\Delta} - b_{22} X_{2(i-1)\Delta}] + (1 + \beta_{21} X_{1(i-1)\Delta}) \Delta}{\beta_{2-}^p}
\]

and the same variance \( \sigma_{2+}^2 \). In short, if \( s \in \{0, 1\} \), with \( \Pr(s = 1) = p \), then

\[
J_{2i\Delta}^p = s \cdot TN(\mu_{2+}^*, \sigma_{2+}^2; J_{2i\Delta}^p > 0) + (1 - s) \cdot TN(\mu_{2-}^*, \sigma_{2+}^2; J_{2i\Delta}^p < 0).
\]

Finally, when \( n_{1\Delta} = 0 \), the conditional posteriors of \( J_{1\Delta}^p \) and \( J_{2i\Delta}^p \) are the prior implied by the model assumptions, as the data provide no information about the jump sizes.

**E A Joint Model of S&P 500 Dynamics and Variance Swaps**

Although we focus on the variance dynamics and develop a partial framework for modeling variance swaps without using S&P 500 index, our framework can be extended to jointly model the index and its variance dynamics. We specify the Q-measure dynamics of the log price \( Y_t \) in the following:

\[
dY_t = \mu_t dt + \sigma_t dB_t^Q + J_t \xi dN_t^Q,
\]

32
\[
\sigma_t^2 = \exp \left\{ \Pi_0 + \Pi_1^T X_t \right\} + \Pi_2 + \Pi_3^T X_t + X_t^T \Pi_4 X_t,
\]
\[
dX_t = (\Lambda^Q + K^Q X_t)dt + \Sigma \sqrt{S_t} dW_t^Q + J^Q X dN_t^Q,
\]

where \( B^Q_t \) and \( W^Q_t \) are correlated Brownian motions, and \( N^Q_t \) is a Poisson process with intensity \( \lambda_t = \lambda_0 + \lambda T X_t \). The jump size can follow any distribution which does not depend on \( X_t \). Under this model, the spot quadratic variation \( f^Q(X_t) \) is a linear combination of Type I and Type II, with a closed-form variance swap pricing formula immediately following from Appendix A. Our Bayesian MCMC method extends naturally to the inference here. The empirical investigation of this general model is left for future research.
References


Baumohl, B. (2010), *The Secrets of Economic Indicators*, Prentice Hall.


<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
<th>Mean</th>
<th>Stdev</th>
<th>HPD 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1^q$, $\kappa_{11}^q$, $\kappa_{21}$, $\kappa_{22}^q$, $\beta_{21}$</td>
<td>$\text{Gamma}(2, 2)$</td>
<td>4.000</td>
<td>(2.827)</td>
<td>[0.485, 11.137]</td>
</tr>
<tr>
<td>$\beta_{1+}^q$, $\beta_{1-}^q$, $\beta_{2+}^q$, $\beta_{2-}^q$</td>
<td>$\text{Inv.Gamma}(3, 2)$</td>
<td>1.000</td>
<td>(1.000)</td>
<td>[0.277, 3.233]</td>
</tr>
<tr>
<td>$q_1^q$, $q_2^q$</td>
<td>$\text{Uniform}(0, 1)$</td>
<td>0.500</td>
<td>(0.289)</td>
<td>[0.025, 0.975]</td>
</tr>
<tr>
<td>$l_0$, $l_{11}$</td>
<td>$\text{Gamma}(1, 1)$</td>
<td>1.000</td>
<td>(1.000)</td>
<td>[0.025, 3.687]</td>
</tr>
<tr>
<td>$\Pi_0^q$, $\Pi_1^q$, $\Pi_2^q$, $\Pi_{11}^q$, $\Pi_{12}^q$, $\Pi_{22}^q$</td>
<td>$\text{N}(0, 6)$</td>
<td>0.000</td>
<td>(6.000)</td>
<td>[-9.795, 9.795]</td>
</tr>
<tr>
<td>$\lambda_1^p$, $\lambda_2^p$</td>
<td>$\text{N}(0.01, 0.01)$</td>
<td>0.010</td>
<td>(0.010)</td>
<td>[-0.01, 0.03]</td>
</tr>
<tr>
<td>$\kappa_{11}^p$, $\kappa_{12}^p$, $\kappa_{22}^p$</td>
<td>$\text{Gamma}(1, 1)$</td>
<td>1.000</td>
<td>(1.000)</td>
<td>[0.025, 3.687]</td>
</tr>
<tr>
<td>$\beta_{1+}^p$, $\beta_{1-}^p$, $\beta_{2+}^p$, $\beta_{2-}^p$</td>
<td>$\text{Inv.Gamma}(3, 2)$</td>
<td>1.000</td>
<td>(1.000)</td>
<td>[0.277, 3.233]</td>
</tr>
<tr>
<td>$s_1^2$, $s_2^2$, $s_3^2$, $s_4^2$, $s_5^2$, $s_6^2$</td>
<td>$\text{Gamma}(1, 1)$</td>
<td>1.000</td>
<td>(1.000)</td>
<td>[0.287, 1.385]</td>
</tr>
</tbody>
</table>

**Table 1: Prior Distribution**

Note: This table presents mean, standard deviation, and 95% highest prior density region for the priors we use to implement our estimation procedure. $\text{Gamma}(\alpha, \beta)$ denotes a gamma random variate with parameters $\alpha$ and $\beta$, such that its mean is $\alpha \beta$ and variance $\alpha \beta^2$; $\text{Inv.Gamma}(\alpha, \beta)$ denotes an inverse gamma random variate with parameters $\alpha$ and $\beta$ such that if $X \sim \text{Inv.Gamma}(\alpha, \beta)$ then $X^{-1} \sim \text{Gamma}(\alpha, \beta^{-1})$; $\text{Uniform}(a, b)$ denotes a random variable which is uniformly distributed on the interval $[a, b]$; and $\text{N}(\mu, \sigma^2)$ denotes a Gaussian random variate with mean $\mu$ and variance $\sigma^2$. Details on models and parameterizations are given in Section 3.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>$A_0(2)$ - Type I</th>
<th>$A_1(2)$ - Type I</th>
<th>$A_0(2)$ - Type II</th>
<th>$A_1(2)$ - Type II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True Bias Stdev</td>
<td>True Bias Stdev</td>
<td>True Bias Stdev</td>
<td>True Bias Stdev</td>
</tr>
<tr>
<td>$\lambda^Q_1$</td>
<td>9.000 0.030 0.042</td>
<td>-3.500 0.003 0.035</td>
<td>-4.000 -0.002 0.085</td>
<td>6.000 0.027 0.064</td>
</tr>
<tr>
<td>$\kappa^Q_{11}$</td>
<td>-3.500 -0.032 0.044</td>
<td>-2.000 0.025 0.014</td>
<td>1.200 0.052 0.125</td>
<td>-5.000 -0.004 0.025</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>-0.300 0.002 0.002</td>
<td>-0.200 -0.017 0.020</td>
<td>-0.250 0.002 0.007</td>
<td>-0.200 -0.001 0.005</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>0.250 0.001 0.001</td>
<td>0.500 0.000 0.000</td>
<td>0.250 0.001 0.001</td>
<td>0.200 0.000 0.001</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>1.000 -0.024 0.017</td>
<td>0.500 0.000 0.000</td>
<td>0.250 0.001 0.001</td>
<td>0.200 0.000 0.001</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>0.400 -0.008 0.005</td>
<td>0.050 0.000 0.000</td>
<td>0.500 0.000 0.000</td>
<td>0.300 0.000 0.000</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>0.400 -0.008 0.005</td>
<td>0.400 -0.003 0.027</td>
<td>0.250 0.000 0.001</td>
<td>0.150 0.000 0.000</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>0.030 -0.001 0.001</td>
<td>0.050 0.000 0.000</td>
<td>0.500 0.001 0.001</td>
<td>0.030 0.000 0.000</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>0.500 0.000 0.007</td>
<td>0.500 -0.007 0.008</td>
<td>0.500 0.001 0.006</td>
<td>0.500 -0.001 0.002</td>
</tr>
<tr>
<td>$\beta^Q_{21}$</td>
<td>1.500 0.016 0.039</td>
<td>1.500 -0.017 0.024</td>
<td>0.750 0.008 0.011</td>
<td>0.500 0.000 0.000</td>
</tr>
<tr>
<td>$l_0$</td>
<td>12.000 0.000 0.001</td>
<td>0.800 -0.001 0.001</td>
<td>1.500 -0.017 0.024</td>
<td>0.750 0.008 0.011</td>
</tr>
<tr>
<td>$l_{11}$</td>
<td>5.000 0.004 0.013</td>
<td>1.500 0.009 0.006</td>
<td>1.500 0.003 0.012</td>
<td>1.250 -0.005 0.004</td>
</tr>
<tr>
<td>$l_{11}$</td>
<td>5.000 0.006 0.016</td>
<td>0.600 -0.014 0.022</td>
<td>0.500 0.001 0.001</td>
<td>0.400 0.000 0.000</td>
</tr>
<tr>
<td>$l_{11}$</td>
<td>8.000 0.014 0.027</td>
<td>0.600 -0.012 0.007</td>
<td>0.400 0.000 0.000</td>
<td>0.400 0.000 0.000</td>
</tr>
<tr>
<td>$l_{11}$</td>
<td>0.100 0.004 0.006</td>
<td>0.200 0.001 0.005</td>
<td>0.400 0.000 0.000</td>
<td>0.400 0.000 0.000</td>
</tr>
<tr>
<td>$l_{11}$</td>
<td>1.200 0.002 0.005</td>
<td>0.080 -0.003 0.003</td>
<td>0.400 0.000 0.000</td>
<td>0.400 0.000 0.000</td>
</tr>
</tbody>
</table>

Table 2: Simulation Results for $\Theta_M$ and $\Theta_H$

Note: This table provides a summary of an Monte Carlo simulation exercise with 100 replications for the two examples of two-factor volatility models that we introduce in section 3, namely the $A_0(2)$ and $A_1(2)$ for both Type I and Type II specifications. We report true values, bias and standard deviations across the simulations for $\Theta_M$, the parameters determining the dynamics of the latent factors under the risk-neutral measure, and $\Theta_H$, the parameters defining $f^\phi$. Details on models and parameterizations are given in Section 3.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\lambda_0^{a}(2)$ - Type I</th>
<th>$\lambda_1^{a}(2)$ - Type I</th>
<th>$\lambda_0^{a}(2)$ - Type II</th>
<th>$\lambda_1^{a}(2)$ - Type II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>Bias</td>
<td>Stdev</td>
<td>True</td>
</tr>
<tr>
<td>$\lambda_0^a$</td>
<td>-0.600</td>
<td>0.015</td>
<td>0.287</td>
<td>-0.030</td>
</tr>
<tr>
<td>$\lambda_1^a$</td>
<td>0.150</td>
<td>-0.160</td>
<td>0.267</td>
<td>1.000</td>
</tr>
<tr>
<td>$\kappa_{11}^P$</td>
<td>-3.000</td>
<td>0.130</td>
<td>0.343</td>
<td>1.000</td>
</tr>
<tr>
<td>$\kappa_{21}^P$</td>
<td>1.000</td>
<td>-0.203</td>
<td>0.195</td>
<td>-0.300</td>
</tr>
<tr>
<td>$\beta_1^{P+}$</td>
<td>0.150</td>
<td>0.036</td>
<td>0.034</td>
<td>-0.200</td>
</tr>
<tr>
<td>$\beta_1^{P-}$</td>
<td>0.450</td>
<td>0.022</td>
<td>0.100</td>
<td>0.250</td>
</tr>
<tr>
<td>$\beta_2^{P+}$</td>
<td>0.050</td>
<td>-0.331</td>
<td>0.015</td>
<td>0.450</td>
</tr>
<tr>
<td>$\beta_2^{P-}$</td>
<td>0.250</td>
<td>0.223</td>
<td>0.045</td>
<td>0.500</td>
</tr>
<tr>
<td>$s_2^M$</td>
<td>0.150</td>
<td>0.005</td>
<td>0.005</td>
<td>0.150</td>
</tr>
<tr>
<td>$s_3^M$</td>
<td>0.010</td>
<td>0.001</td>
<td>0.000</td>
<td>0.010</td>
</tr>
<tr>
<td>$s_4^M$</td>
<td>0.050</td>
<td>0.002</td>
<td>0.001</td>
<td>0.050</td>
</tr>
<tr>
<td>$s_5^M$</td>
<td>0.050</td>
<td>0.000</td>
<td>0.002</td>
<td>0.050</td>
</tr>
<tr>
<td>$s_6^Y$</td>
<td>0.010</td>
<td>0.001</td>
<td>0.000</td>
<td>0.010</td>
</tr>
<tr>
<td>$s_7^Y$</td>
<td>0.100</td>
<td>0.000</td>
<td>0.004</td>
<td>0.500</td>
</tr>
</tbody>
</table>

Table 3: Simulation Results for $\Theta_P$ and $\Theta_E$

Note: This table provides a summary of an Monte Carlo simulation exercise with 100 replications for the two examples of two-factor volatility models that we introduce in section 2, namely the $\lambda_0^{a}(2)$ and $\lambda_1^{a}(2)$ for both Type I and Type II specifications. We report true values, bias and standard deviations across the simulations for the additional parameters that characterize the $\mathbb{P}$-dynamics, $\Theta_P$, and the pricing error variances, summarized in $\Theta_E$. Details on models and parameterizations are given in Section 3.
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Category</th>
<th>Frequency</th>
<th>Release Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment Rate</td>
<td>Employment</td>
<td>Monthly</td>
<td>8:30 am First Friday of each month</td>
</tr>
<tr>
<td>ADP Employment Change</td>
<td>Employment</td>
<td>Monthly</td>
<td>8:15 am - Two days before Employment situation</td>
</tr>
<tr>
<td>Initial Jobless Claims</td>
<td>Employment</td>
<td>Weekly</td>
<td>8:30 am every Thursday</td>
</tr>
<tr>
<td>Personal Income</td>
<td>Consumer Spending and Confidence</td>
<td>Monthly</td>
<td>8:30 am 4 weeks after end of reported month</td>
</tr>
<tr>
<td>Personal Spending</td>
<td>Consumer Spending and Confidence</td>
<td>Monthly</td>
<td>8:30 am 4 weeks after end of reported month</td>
</tr>
<tr>
<td>Advance Retail Sales</td>
<td>Consumer Spending and Confidence</td>
<td>Monthly</td>
<td>8:30 am 2 weeks after end of reported month</td>
</tr>
<tr>
<td>Consumer Confidence</td>
<td>Consumer Spending and Confidence</td>
<td>Monthly</td>
<td>10:00 am - Last Tuesday of month being surveyed</td>
</tr>
<tr>
<td>GDP</td>
<td>National Output and Inventories</td>
<td>Quarterly</td>
<td>8:30 am - Final week of Jan Apr Jul Oct</td>
</tr>
<tr>
<td>Durable Goods Orders</td>
<td>National Output and Inventories</td>
<td>Monthly</td>
<td>8:30 am three to four weeks after the end of reporting month</td>
</tr>
<tr>
<td>ISM Manufacturing</td>
<td>National Output and Inventories</td>
<td>Monthly</td>
<td>10:00 am First Business day after reporting month</td>
</tr>
<tr>
<td>Chicago PMI</td>
<td>National Output and Inventories</td>
<td>Monthly</td>
<td>10:00 am First Business day of month being covered</td>
</tr>
<tr>
<td>Empire State Manufacturing</td>
<td>National Output and Inventories</td>
<td>Monthly</td>
<td>8:30 am around 15th of month being reported</td>
</tr>
<tr>
<td>Business Inventories</td>
<td>National Output and Inventories</td>
<td>Monthly</td>
<td>10:00 am released six weeks after the month ends</td>
</tr>
<tr>
<td>Production and Utilization</td>
<td>National Output and Inventories</td>
<td>Monthly</td>
<td>9:15 am released the 15th of the following month</td>
</tr>
<tr>
<td>New Residential Sales</td>
<td>Housing and Construction</td>
<td>Monthly</td>
<td>8:30 am released two to three weeks following month being covered</td>
</tr>
<tr>
<td>FOMC Meeting</td>
<td>Central Bank</td>
<td>Eight Times</td>
<td>2:15 pm of day of conclusion of FOMC meeting</td>
</tr>
<tr>
<td>Federal Reserver Chairman Speech</td>
<td>Central Bank</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td>ECB Governing Council Meeting</td>
<td>Central Bank</td>
<td>Monthly</td>
<td>N.A.</td>
</tr>
<tr>
<td>CPI</td>
<td>Prices, Productivity, Wages</td>
<td>Monthly</td>
<td>8:30 am second or third week following month being covered</td>
</tr>
<tr>
<td>PPI</td>
<td>Prices, Productivity, Wages</td>
<td>Monthly</td>
<td>8:30 am two or three weeks after month ends</td>
</tr>
<tr>
<td>Employment Cost Index</td>
<td>Prices, Productivity, Wages</td>
<td>Quarterly</td>
<td>8:30 am - Last Thursday of Jan Apr Jul Oct</td>
</tr>
</tbody>
</table>

**Table 4: Economic Indicators**

Note: In this table, we report the details of 21 macro news announcements or central bank events used in Section 5.1. All times are reported in Eastern Time. Source: Bloomberg and the book by Baumohl (2010) on economic indicators.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\Lambda_0(2)$</th>
<th>$\Lambda_1(2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>$\lambda_1^0$</td>
<td>-4.464 (0.035)</td>
<td>[-4.531,-4.391]</td>
</tr>
<tr>
<td>$\kappa_{11}^0$</td>
<td>1.532 (0.042)</td>
<td>[1.445,1.611]</td>
</tr>
<tr>
<td>$\kappa_{21}^0$</td>
<td>-0.207 (0.003)</td>
<td>[-0.214,-0.201]</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.386 (0.013)</td>
<td>[0.366,0.410]</td>
</tr>
<tr>
<td>$\beta_{1r}^0$</td>
<td>0.091 (0.001)</td>
<td>[0.088,0.093]</td>
</tr>
<tr>
<td>$\beta_{1s}^0$</td>
<td>0.049 (0.004)</td>
<td>[0.044,0.055]</td>
</tr>
<tr>
<td>$\beta_{2r}^0$</td>
<td>0.467 (0.017)</td>
<td>[0.444,0.507]</td>
</tr>
<tr>
<td>$\beta_{2s}^0$</td>
<td>10.039 (0.008)</td>
<td>[10.025,10.054]</td>
</tr>
<tr>
<td>$l_0$</td>
<td>0.400 (0.050)</td>
<td>[0.339,0.508]</td>
</tr>
<tr>
<td>$\Pi_0^\theta$</td>
<td>1.196 (0.011)</td>
<td>[1.177,1.215]</td>
</tr>
<tr>
<td>$\Pi_1^\theta$</td>
<td>0.478 (0.003)</td>
<td>[0.474,0.483]</td>
</tr>
</tbody>
</table>

Table 5: Posterior Estimates of $\Theta_M$ and $\Theta_{II}$ for Type II models

Note: This table presents the posterior estimates for $\Theta_M$, the parameters determining the dynamics of the latent factors under the risk-neutral measure, and $\Theta_{II}$, the parameters defining $f^Q$ across all models. We report the mean, the standard deviation, and the 95% highest posterior density intervals for the $\Lambda_0(2)$ and $\Lambda_1(2)$ for the Type I specification. We use daily data on variance swaps from January 4, 1996 to January 11, 2013. The number of daily observations is 4,278, excluding weekends and holidays. Details on models and parameterizations are given in Section 3.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>$A_0(2)$</th>
<th>$A_1(2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>$\lambda_1^p$</td>
<td>-0.093 (0.266)</td>
<td>[-0.614,0.427]</td>
</tr>
<tr>
<td>$\lambda_2^p$</td>
<td>1.017 (0.261)</td>
<td>[0.518,1.531]</td>
</tr>
<tr>
<td>$\kappa_{11}$</td>
<td>-4.448 (0.235)</td>
<td>[-4.913,-3.987]</td>
</tr>
<tr>
<td>$\kappa_{21}$</td>
<td>-0.267 (0.086)</td>
<td>[-0.433,-0.098]</td>
</tr>
<tr>
<td>$\beta_{1+}^p$</td>
<td>0.178 (0.019)</td>
<td>[0.144,0.219]</td>
</tr>
<tr>
<td>$\beta_{1-}^p$</td>
<td>0.117 (0.010)</td>
<td>[0.099,0.138]</td>
</tr>
<tr>
<td>$\beta_{2+}^p$</td>
<td>0.141 (0.017)</td>
<td>[0.112,0.177]</td>
</tr>
<tr>
<td>$\beta_{2-}^p$</td>
<td>0.444 (0.020)</td>
<td>[0.406,0.477]</td>
</tr>
<tr>
<td>$\beta_{2+}^p$</td>
<td>0.199 (0.019)</td>
<td>[0.164,0.241]</td>
</tr>
<tr>
<td>$\beta_{2-}^p$</td>
<td>0.161 (0.018)</td>
<td>[0.13,0.2]</td>
</tr>
<tr>
<td>$s_{2M}^2$</td>
<td>0.146 (0.004)</td>
<td>[0.138,0.154]</td>
</tr>
<tr>
<td>$s_{3M}^2$</td>
<td>0.005 (0.000)</td>
<td>[0.004,0.005]</td>
</tr>
<tr>
<td>$s_{6M}^2$</td>
<td>0.046 (0.001)</td>
<td>[0.044,0.048]</td>
</tr>
<tr>
<td>$s_{9M}^2$</td>
<td>0.063 (0.002)</td>
<td>[0.060,0.067]</td>
</tr>
<tr>
<td>$s_{1Y}^2$</td>
<td>0.004 (0.000)</td>
<td>[0.003,0.004]</td>
</tr>
<tr>
<td>$s_{2Y}^2$</td>
<td>0.058 (0.002)</td>
<td>[0.055,0.062]</td>
</tr>
</tbody>
</table>

**Table 6: Posterior Estimates of $\Theta_P$ and $\Theta_E$ for Type II models**

Note: This table presents the posterior estimates for the parameters that characterize the $P$-dynamics, $\Theta_P$, and the pricing error variances $\Theta_E$. We report the mean, the standard deviation, and the 95% highest posterior density intervals for the $A_0(2)$ and $A_1(2)$ for the Type I specification. We use daily data on variance swaps from January 4, 1996 to January 11, 2013. The number of daily observations is 4,278, excluding weekends and holidays. Details on the models and parameterizations are given in Section 3.
<table>
<thead>
<tr>
<th>Date</th>
<th>ΔVariance</th>
<th>ΔReturn</th>
<th>News</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/18/98</td>
<td>-0.373</td>
<td>0.013</td>
<td>Clinton Admits To “Wrong” Relationship with Lewinsky and FOMC’s Decision to Leave Interest Rates Unchanged</td>
</tr>
<tr>
<td>09/01/98</td>
<td>-0.722</td>
<td>0.035</td>
<td>Federal Reserve to Add Money to the Banking System with Repo</td>
</tr>
<tr>
<td>09/08/98</td>
<td>-0.526</td>
<td>0.021</td>
<td>Federal Reserve Chairman Alan Greenspan’s Statement that a Rate Cut might be Forthcoming</td>
</tr>
<tr>
<td>09/23/98</td>
<td>-0.344</td>
<td>0.027</td>
<td>Federal Reserve Chairman Alan Greenspan Testimony before the Committee on the Budget, U.S. Senate</td>
</tr>
<tr>
<td>10/20/98</td>
<td>-0.253</td>
<td>-0.007</td>
<td>Three Big US Banks Delivered Better-Than-Expected Earnings and Bullish Mood After Fed Rate Cut Previous Week</td>
</tr>
<tr>
<td>08/11/99</td>
<td>-0.266</td>
<td>0.008</td>
<td>The Release of the Federal Reserve’s Beige Book Shows that US Economy Remains Strong</td>
</tr>
<tr>
<td>01/07/00</td>
<td>-0.500</td>
<td>0.031</td>
<td>Unemployment Report Shows the Lowest Unemployment Rate in the past 30 Years</td>
</tr>
<tr>
<td>03/16/00</td>
<td>-0.266</td>
<td>0.037</td>
<td>Release of Inflation Remains Tame Enough to Keep the Federal Reserve from Tightening Credit</td>
</tr>
<tr>
<td>04/17/00</td>
<td>-0.373</td>
<td>0.032</td>
<td>Treasury Secretary Lawrence H. Summers Statement that Fundamentals of Economy are in Place</td>
</tr>
<tr>
<td>10/19/00</td>
<td>-0.241</td>
<td>0.018</td>
<td>Fed’s Greenspan Gives Keynote Speech at Cato Institute and Jobless Claim Drop by 7,000 in Latest Week</td>
</tr>
<tr>
<td>01/03/01</td>
<td>-0.282</td>
<td>0.052</td>
<td>The Greenspan Fed Announcement of a Surprise, Inter-Meeting Rate Cut</td>
</tr>
<tr>
<td>05/17/05</td>
<td>-0.275</td>
<td>0.019</td>
<td>Treasury Secretary John Snow’s Call on China to Take an Intermediate Step in Revaluing its Currency</td>
</tr>
<tr>
<td>06/15/06</td>
<td>-0.549</td>
<td>0.017</td>
<td>Fed Chairman Ben Bernanke’s Speech on Inflation Expectations within Historical Ranges.</td>
</tr>
<tr>
<td>06/29/06</td>
<td>-0.295</td>
<td>0.016</td>
<td>The FOMC Statement to Raise Its Target for the Federal Funds Rate by 25 Basis Points</td>
</tr>
<tr>
<td>09/18/07</td>
<td>-0.415</td>
<td>0.024</td>
<td>The Federal Open Market Committee Decided Today to Lower its Target for the Federal Funds Rate by 50 Basis Points</td>
</tr>
<tr>
<td>10/14/08</td>
<td>-0.489</td>
<td>-0.048</td>
<td>Secretary of the Treasury Henry Paulson and President Bush Separately Announced Revisions to the TARP Program</td>
</tr>
<tr>
<td>10/20/08</td>
<td>-0.426</td>
<td>0.033</td>
<td>Chairman Ben S. Bernanke Testimony before the Committee on the Budget, U.S. House of Representatives</td>
</tr>
<tr>
<td>10/28/08</td>
<td>-0.313</td>
<td>0.075</td>
<td>Market Expectation on Federal Reserve Rate Cut the Next Day</td>
</tr>
<tr>
<td>11/13/08</td>
<td>-0.328</td>
<td>0.062</td>
<td>President Bush’s Speech on Financial Crisis</td>
</tr>
<tr>
<td>05/10/10</td>
<td>-0.647</td>
<td>0.003</td>
<td>European Policy Makers Unveiled an Unprecedented Emergency Loan Plan</td>
</tr>
<tr>
<td>08/09/11</td>
<td>-0.433</td>
<td>0.046</td>
<td>FOMC Statement Explicitly Stating a Duration for an Exceptionally Low Target Rate</td>
</tr>
<tr>
<td>10/27/11</td>
<td>-0.245</td>
<td>0.034</td>
<td>European Union Leaders Made a Bond Deal to Fix the Greek Debt Crisis</td>
</tr>
<tr>
<td>12/31/12</td>
<td>-0.432</td>
<td>0.025</td>
<td>Obama, Senate Republicans Reach Agreement on the “Fiscal Cliff” Issue</td>
</tr>
</tbody>
</table>

**Table 7: Policy News Potentially Associated with Estimated Volatility Jumps**

Note: In this table, we report the potential events in the last column that may lead to the 23 largest negative volatility jumps in sample. The first column is the date of the event. The second column shows changes in estimated spot variance, whereas the third column is the returns of the index on the corresponding days.
<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEF</td>
<td>0.279***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.098</td>
<td>(0.079)</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.079)</td>
<td></td>
</tr>
<tr>
<td>TED</td>
<td></td>
<td>0.188*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.079</td>
<td>(0.078)</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TERM</td>
<td></td>
<td></td>
<td>-0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.016</td>
<td>(0.035)</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIQ</td>
<td></td>
<td></td>
<td></td>
<td>-0.533**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.143</td>
<td>(0.164)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.221)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.004***</td>
<td></td>
<td></td>
<td></td>
<td>0.002***</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SKEW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.016</td>
<td></td>
<td></td>
<td>0.03</td>
<td>(0.023)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.030)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.026***</td>
<td></td>
<td></td>
<td>-0.024***</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.019</td>
<td></td>
<td>-0.032</td>
<td>(0.035)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.036)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
<td>0.049</td>
<td>(0.027)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>0.781***</td>
<td>0.788***</td>
<td>0.793***</td>
<td>0.779***</td>
<td>0.806***</td>
<td>0.792***</td>
<td>0.800***</td>
<td>0.791***</td>
<td>0.792***</td>
<td>0.816***</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.034)</td>
<td>(0.037)</td>
<td>(0.034)</td>
<td>(0.030)</td>
<td>(0.034)</td>
<td>(0.040)</td>
<td>(0.035)</td>
<td>(0.035)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>adj.R²</td>
<td>0.625</td>
<td>0.632</td>
<td>0.617</td>
<td>0.627</td>
<td>0.659</td>
<td>0.617</td>
<td>0.727</td>
<td>0.618</td>
<td>0.617</td>
<td>0.748</td>
</tr>
</tbody>
</table>

**Table 8: Regression Results on Factor $X_1$**

Note: In this table we report results on regression analysis to relate our volatility factors to several variables on economic fundamentals at the monthly frequency. Each column reports the results of estimating a linear regression of the posterior mean of the volatility factor $X_1$ on its lagged value (AR) and the innovation of the corresponding variable for each row. The last column corresponds to the multiple regression that includes all the explanatory variables we consider: TED spread, default spread (DEF), Chicago Fed National Activity Index (CFI), industrial production growth (IPG), term spread (TERM), monthly liquidity factor (LIQ), policy news index (POL), market skewness (SKEW), and excess market returns (ExM), see Section 5.1 for more information on their definitions. The coefficients corresponding to the constants are omitted from the regressions.
Table 9: Regression Results on Factor $X_2$

Note: In this table we report results on regression analysis to relate our volatility factors to several variables on economic fundamentals at the monthly frequency. Each column reports the results of estimating a linear regression of the posterior mean of the volatility factor $X_2$ on its lagged value (AR) and the innovation of the corresponding variable for each row. The last column corresponds to the multiple regression that includes all the explanatory variables we consider: TED spread, default spread (DEF), Chicago Fed National Activity Index (CFI), industrial production growth (IPG), term spread (TERM), monthly liquidity factor (LIQ), policy news index (POL), market skewness (SKEW), and excess market returns (ExM), see Section 5.1 for more information on their definitions. The coefficients corresponding to the constants are omitted from the regressions.
<table>
<thead>
<tr>
<th>News / Events</th>
<th>Jumps of $X_1$</th>
<th>Jumps of $X_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>0.624***</td>
<td>0.016</td>
</tr>
<tr>
<td>ADP Employment Change</td>
<td>0.252</td>
<td>-0.181</td>
</tr>
<tr>
<td>Initial Jobless Claims</td>
<td>-0.112</td>
<td>0.023</td>
</tr>
<tr>
<td>Personal Income</td>
<td>-0.46</td>
<td>-0.065</td>
</tr>
<tr>
<td>Personal Spending</td>
<td>0.485*</td>
<td>-0.101</td>
</tr>
<tr>
<td>Advance Retail Sales</td>
<td>-0.067</td>
<td>-0.094</td>
</tr>
<tr>
<td>Consumer Confidence</td>
<td>-0.329</td>
<td>-0.088</td>
</tr>
<tr>
<td>GDP</td>
<td>0.41</td>
<td>-0.024</td>
</tr>
<tr>
<td>Durable Goods Orders</td>
<td>0.219</td>
<td>-0.13</td>
</tr>
<tr>
<td>ISM PMI</td>
<td>-0.098</td>
<td>0.197</td>
</tr>
<tr>
<td>Chicago PMI</td>
<td>0.752***</td>
<td>-0.102</td>
</tr>
<tr>
<td>Empire State Manufacturing</td>
<td>0.13</td>
<td>0.247</td>
</tr>
<tr>
<td>Business Inventory</td>
<td>0.159</td>
<td>-0.197</td>
</tr>
<tr>
<td>Capacity Utilization</td>
<td>-0.186</td>
<td>0.343*</td>
</tr>
<tr>
<td>New Residential Sales</td>
<td>0.084</td>
<td>-0.017</td>
</tr>
<tr>
<td>FOMC Meeting</td>
<td>-0.309</td>
<td>0.775***</td>
</tr>
<tr>
<td>Fed Chairman Speech</td>
<td>-0.152</td>
<td>0.203*</td>
</tr>
<tr>
<td>ECB</td>
<td>0.165</td>
<td>-0.129</td>
</tr>
<tr>
<td>CPI</td>
<td>-0.155</td>
<td>-0.207</td>
</tr>
<tr>
<td>PPI</td>
<td>0.113</td>
<td>0.098</td>
</tr>
<tr>
<td>Employment Cost Index</td>
<td>-0.216</td>
<td>-0.11</td>
</tr>
</tbody>
</table>

Table 10: Regression Results on Volatility Jumps

Note: In this table, we report the regressions of the magnitudes of jumps in each volatility factor onto the magnitudes of news shocks, defined in Section 5.1. All the shocks are standardized to have variance equal to 1.
Figure 1: Negative Jumps in the VIX

Note: In this figure, we highlight three days corresponding to the following media headlines: VIX, Vstoxx Drop by Records as Stocks Soar on Europe’s Emergency Loan Plan. - Bloomberg, Monday May 10, 2010; VIX Index Driven to Second-Biggest Percentage Drop (- 27%) on Fed’s Rate Statement. - Bloomberg, Tuesday Aug 09, 2011; The CBOE Volatility Index, or the VIX, Wall Street’s favored measure of anxiety, posted its biggest one-day decline since August 2011, as lawmakers closed in on a deal to avert the “fiscal cliff.” - Reuters, Monday Dec 31, 2012.
Figure 2: Times Series of Variance Swap Rates

Note: In this figure we show the time series of the variance swap rates with 6 different maturities from January 4, 1996 to January 11, 2013. The maximum number of daily observations is 4,278, excluding weekends and holidays. Since we have an unbalanced panel, different maturities may have different number of observations, which are reported in the legend.
Figure 3: Term Structure of Variance Swap Rates

Note: In this figure we plot the term structure of variance swap rates. The upper and lower panels plot 
\[
\frac{P(t, 1/2) - P(t, 1/4)}{P(t, 1/4)} \quad \text{and} \quad \frac{P(t, 1) - P(t, 1/4)}{P(t, 1/4)},
\]
respectively, where \( P(t, \tau) \) denotes the variance swap rate at time \( t \) of a contract with time to maturity \( \tau \). Positive values reflect an upward sloping term structure while the opposite slope is implied by negative values.
Figure 4: Posterior Distributions of the Pricing Errors

Note: This figure reports the posterior estimates of the pricing error standard deviation parameters for both $A_0(2)$ and $A_1(2)$ models with Type I and II specifications through box-plots respectively. As usual, the central boxes describe the first and third quartiles of the sampling distributions, as well as their median. The maximum length of the whiskers is one interquartile range. We use daily data on variance swaps from January 4, 1996 to January 11, 2013. The number of daily observations is 4,278, excluding weekends and holidays. Details on the models and parameterizations are given in Section 3.
Figure 5: Factors of Type I Models

Note: This figure reports the posterior estimates of the two latent factors for both $A_0(2)$ and $A_1(2)$ models with Type I specification. Blue solid lines plot the posterior mean of the factors while red dotted lines plot the 95% highest posterior density intervals distributions. We use daily data on variance swaps from January 4, 1996 to January 11, 2013. The number of daily observations is 4,278, excluding weekends and holidays. Details on the models and parameterizations are given in Section 3.
Note: This figure reports the posterior estimates of the two latent factors for both $A_0(2)$ and $A_1(2)$ models with Type II specification. Blue solid lines plot the posterior mean of the factors while red dotted lines plot the 95% highest posterior density intervals distributions. We use daily data on variance swaps from January 4, 1996 to January 11, 2013. The number of daily observations is 4,278, excluding weekends and holidays. Details on the models and parameterizations are given in Section 3.
Figure 7: Out of Sample Performance

Note: This figure compares the estimated 1-month variance swap rates with the VIX across models and types over the entire sample period. The red solid line denotes the VIX from the CBOE, whereas the blue dash-dotted line is calculated based on the $Q$-parameters estimated from the variance swap rates with time-to-maturity at least 2 months.
Figure 8: Decomposition of Spot Variance of $A_0(2)$ and $A_1(2)$ Models of Type II

Note: In this figure we decompose the spot variance into its continuous and jump components based for both $A_0(2)$ (left) and $A_1(2)$ (right) models with Type II specification. The first panels plot the estimated percent changes of spot variance, followed by volatility jumps on the second panels, and the remaining Brownian shocks on the last panels. All these components are extracted based on the above models. The red circles correspond to the three aforementioned events in Figure 1.
Figure 9: Variance Risk Premia Implied from $A_0(2)$ and $A_1(2)$ Models of Type II

Note: In this figure we calculate the term structure of variance risk premia using $A_0(2)$ (left) and $A_1(2)$ (right) models with Type II specification. The red solid lines, the blue dashed lines, and the green dash-dotted lines correspond to the 1-month, 6-month, and 2-year maturities, respectively. Note that the scales of the y-axis are different for the two plots.